March 2019

**Machine Learning, Game Theory Algorithms, and Medium Access Protocols for 5G and Internet-of-Thing (IoT) Networks**

Mohamed Elkourdi

*University of South Florida, elkourdi@mail.usf.edu*

Follow this and additional works at: [https://scholarcommons.usf.edu/etd](https://scholarcommons.usf.edu/etd)

Part of the [Electrical and Computer Engineering Commons](https://scholarcommons.usf.edu/etd)

**Scholar Commons Citation**


https://scholarcommons.usf.edu/etd/7782

This Dissertation is brought to you for free and open access by the Graduate School at Scholar Commons. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Scholar Commons. For more information, please contact scholarcommons@usf.edu.
Machine Learning, Game Theory Algorithms, and Medium Access Protocols

for 5G and Internet-of-Thing (IoT) Networks

by

Mohamed Elkourdi

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy
Department of Electrical Engineering
College of Engineering
University of South Florida

Major Professor: Richard D. Gitlin, Sc.D.
    Nasir Ghani, Ph.D.
    Ismail Uysal, Ph.D.
    Srinivas Katkoori, Ph.D.
    Gabriel Arrobo, Ph.D.

Date of Approval:
    February 21, 2019

Keywords: Recurrent Neural Network, Coordinated Multipoint Transmission, Successive Interference Cancellation, Non-Orthogonal Multiple Access Schemes

Copyright © 2019, Mohamed Elkourdi
DEDICATION

To my beloved parents, Hesham and Sana
ACKNOWLEDGMENTS

First and foremost, I would like to thank my advisor Dr. Richard D. Gitlin for his continuous guidance, encouragement and support. I am deeply indebted to him for giving me the opportunity to work with him and be part of his research group in the Innovations in Wireless Information Networking Laboratory (iWINLAB). I consider Dr. Gitlin as my mentor from whom I have learned a lot throughout my Ph.D. All the lessons I learned from Dr. Gitlin have helped me to grow both professionally and personally.

Also, I would like to thank Dr. Nasir Ghani, Dr. Ismail Uysal, Dr. Srinivas Katkoori and Dr. Gabriel Arrobo for their feedback and serving in my committee and for Dr. Yao Liu for chairing my Ph.D. defense. Gratitude is extended to all the faculty members in Electrical and Computer Engineering (ECE) department at University of South Florida (USF), especially Dr. Nasir Ghani for his motivating and enjoyable talks. Also, I would like to acknowledge Dr. Paris Wiley for his help and support during my teaching assistantship.

I’m also very grateful for the Electrical Engineering department staff, Jessica Procko, Kristen Brandt and Diana Hamilton for their assistance with administrative work.

Also, one of the greatest gifts that I have earned from my Ph.D. journey is the friendships that I have made. I'm very thankful to all my friends and colleagues whom I met during my time at USF. Thanks goes to my friends Asim Mazin, Derar Hawatmeh, Majdi Ababneh, Mohamed Hafez, Ammar Amouri, Dilranjan Wickramasuriya, Ertugrul Guvenkaya, Z. Esat Ankarali, Anas Tom, M. Harun Yilmaz, Nabeel Sulieman, Calvin Perumalla, Yang Liu, Saud Aldosarry, Faeik Al Rabee, Mohammed Alrowaily, Alla Abdulla, Mohamed Abdin, Berker Pekoz, Abdullah
Qaroot, Adam Aboalroub, and Ali Fatih. I wish all of you the best in your future lives and I hope that our paths cross again in the future.

Many thanks to my cousins Khaled, and Peter Seller and his wife Elizabeth for their hospitality and helping me when I first arrived to Tampa. They definitely made it easy for me when I first moved to Florida. I also want to thank my dear siblings Tariq Elkourdi, Yousef Elkourdi and Farah Elkourdi for their encouragement and supporting me during the hard times and celebrating with me during the good times.

Last, but not least, I am blessed with a wonderful family, to whom I dedicated this dissertation. I want to express my utmost appreciation to my parents Sana Haj Ali and Hesham Elkourdi. Their unconditional love and support was the primarily reason that I was able to overcome the challenges and difficulties that I faced. I will be forever indebted to them for all they have done to help me reach the place that I am at today.
TABLE OF CONTENTS

LIST OF TABLES .................................................................................................................................................. iii

LIST OF FIGURES .................................................................................................................................................... iv

ABSTRACT ............................................................................................................................................................... vii

CHAPTER 1: INTRODUCTION ............................................................................................................................... 1
  1.1 Massive Machine-Type Communication (mMTC) ......................................................................................... 2
  1.2 Enhanced Mobile Broadband (eMBB) ........................................................................................................... 3
  1.3 Low Latency Communication (LLC) .............................................................................................................. 6
  1.4 Contributions and Organization of This Dissertation ................................................................................... 8

CHAPTER 2: LITERATURE REVIEW .......................................................................................................................... 11
  2.1 Topic 1: Medium Access Protocols for Internet of Things (IoT) Networks .................................................. 11
    2.1.1 Internet of Things (IoT) ......................................................................................................................... 11
    2.1.2 Medium Access Control (MAC) Protocols ............................................................................................. 13
      2.1.2.1 Contention-Free Protocols ............................................................................................................. 13
      2.1.2.2 Contention-Based Protocols ........................................................................................................... 14
        2.1.2.2.1 Overview of ALOHA-Based Protocols ..................................................................................... 14
        2.1.2.2.2 An Overview on CSMA/CA Protocol ...................................................................................... 18
    2.2 Topic 2: Deep Machine Learning in 5G Virtual Cell Networks .................................................................... 19
    2.3 Topic 3: Cell Selection/User Association in 5G Heterogenous Networks ....................................................... 22

CHAPTER 3: SLOTTED ALOHA NOMA (SAN) PROTOCOL FOR IoT NETWORKS ................................................. 24
  3.1 Introduction ....................................................................................................................................................... 24
  3.2 Overview on Non-Orthogonal Multiple Access (NOMA) ............................................................................ 25
  3.3 Flexible Frame Structure for the SAN Protocol ............................................................................................. 29
  3.4 Detecting the Number of Active IoT Devices ................................................................................................. 31
  3.5 Simulation Results .......................................................................................................................................... 34
  3.6 Slotted ALOHA-NOMA with MIMO Beam-Forming (BF-SAN) ..................................................................... 37

CHAPTER 4: OPTIMIZATION OF 5G VIRTUAL-CELL BASED COORDINATED MULTIPROTOCOL NETWORKS USING DEEP MACHINE LEARNING ................................................................. 41
  4.1 Introduction ....................................................................................................................................................... 41
  4.2 CoMP Management Based on Recurrent Neural Network ........................................................................... 43
  4.3 Simulation Results .......................................................................................................................................... 47
  4.4 Performance Analysis and Discussion ........................................................................................................... 51
CHAPTER 5: TOWARDS LOW LATENCY IN 5G HETNETS: A BAYESIAN CELL
SELECTION / USER ASSOCIATION APPROACH ..........................................................53
  5.1 Introduction ........................................................................................................53
  5.2 System Model .....................................................................................................53
  5.3 Problem Formulation and Proposed Bayesian Game Approach ......................56
  5.4 Simulation Results and Analysis .........................................................................60

CHAPTER 6: CONCLUSIONS .........................................................................................65
  6.1 Slotted ALOHA Non-Orthogonal Multiple Access (SAN) Protocol for IoT
      Networks ..............................................................................................................65
  6.2 Recurrent Neural Networks for Proactive Mobility Management in 5G
      Virtual-Cell Networks .........................................................................................66
  6.3 Bayesian Game Approach for Cell Selection / User Association in 5G HetNet ....67

REFERENCES ...........................................................................................................69

APPENDIX A: COPYRIGHT PERMISSIONS ...............................................................74
LIST OF TABLES

Table 4.1: Samples of RSS values for different nodes .....................................................................45
Table 5.1: Payoff matrix ..................................................................................................................60
Table 5.2: Simulation parameters ..................................................................................................61
LIST OF FIGURES

Figure 1.1 Illustration of coordinated medium access among multiple devices: time-
division multiple access (TDMA), frequency-division multiple access
(FDMA) and code-division multiple access (CDMA), where B is the total
bandwidth and T is the total transmission time .................................................. 2

Figure 1.2 Evolution of a classic cellular wireless network to a virtual cell network .......... 5

Figure 1.3 Evolution of RAN architecture to C-RAN architecture that consists of
distributed remote radio heads (RRHs) and a centralized baseband
units (BBUs) ............................................................................................................... 7

Figure 1.4 Fog radio access network (F-RAN) consisting of fog low power nodes
(F-LPNs) each provided with limited local storage, BBU pool and real
time computation ..................................................................................................... 7

Figure 1.5 Outline of the dissertation including the research topics and the names of
chapters .................................................................................................................... 8

Figure 2.1 The use case of an IoT managed smart home ................................................ 12

Figure 2.2 Schematic diagram of ALOHA protocol with M IoT devices ......................... 15

Figure 2.3 Schematic diagram of slotted ALOHA (SA) protocol with M IoT devices........ 16

Figure 2.4 Schematic diagram of reservation ALOHA (R-ALOHA) protocol ................. 17

Figure 2.5 An illustration of slotted ALOHA with successive interference
cancellation (SIC) receiver at the IoT gateway ..................................................... 18

Figure 2.6 Reference signal received power (RSRP)-based cell selection
approach used in long term evolution (LTE) systems ............................................. 23

Figure 2.7 Cell range expansion (CRE) approach ....................................................... 23

Figure 3.1 Orthogonal multiple access (OMA) vs. non-orthogonal multiple
access schemes ........................................................................................................ 26

Figure 3.2 OMA vs. NOMA capacity regions ................................................................ 27
Figure 3.3 Superposition coding for quadrature phase shift keying (QPSK) constellation points of devices A and B .................................................................28

Figure 3.4 Successive interference cancellation (SIC) receiver .................................................................29

Figure 3.5 The proposed SAN protocol frame structure ........................................................................31

Figure 3.6 Probability of detection as a function of SNR ........................................................................34

Figure 3.7 Throughput of slotted ALOHA-NOMA (SAN) with 2 and 3 optimum power levels vs slotted ALOHA as a function of probability of transmission ........................................................................35

Figure 3.8 Throughput as a function of optimum power levels for different number of attempts (2, 3 and 5) using SAN protocol ........................................................................36

Figure 3.9 Average channel access delay as a function of probability of transmission for the SAN protocol and the well-known slotted ALOHA protocol ........................................................................37

Figure 3.10 Throughput as a function of probability of transmission for BF-SAN, SAN and conventional slotted ALOHA protocols ........................................................................38

Figure 3.11 Throughput of BF-SA-NOMA VS. ALOHA-NOMA for different values of probability of transmission when $M = 60$, $m = 3$, $k = 2$ and 3 attempts, $N_f = 2$ and $N_t = 2$ ........................................................................39

Figure 3.12 Average delay of BF-SA-NOMA VS. ALOHA-NOMA for different values of probability of transmission when $M = 60$, $m = 3$, $k = 2$ and 3 attempts, $N_f = 2$ and $N_t = 2$ ........................................................................40

Figure 4.1 Adaptively enabling / disabling coordinated multipoint (CoMP) transmission according to the trajectory of a user in the network .................................................................42

Figure 4.2 A road network with eight base stations (BSs) (crosses) and a mobile node (circle) .................................................................................................................................44

Figure 4.3 (a) RNN structure ..................................................................................................................47

Figure 4.4 Gated recurrent unit (GRU) architecture ..................................................................................47

Figure 4.5 True versus predicted received signal strength (RSS) values for two different UEs .................................................................................................................................49

Figure 4.6 The CDF of enabling virtual cells using the GRU-RNN predictive model .........................50

Figure 4.7 Convergence of the training model calculated based on MSE loss function ..................50
Figure 5.1 Two-tier 5G HetNet consisting of a radio access controller (RAC), high-power node (HPN) and fog-low power nodes (F-LPNs).................................55

Figure 5.2 CCDFs of achieved latency for delay-sensitive traffic.................................................62

Figure 5.3 CDFs of proper association for delay-sensitive traffic..................................................63

Figure 5.4 CDFs of proper association for delay-tolerant traffic....................................................63
ABSTRACT

In the first part of this dissertation, a novel medium access protocol for the Internet of Thing (IoT) networks is introduced. The Internet of things (IoT), which is the network of physical devices embedded with sensors, actuators, and connectivity, is being accelerated into the mainstream by the emergence of 5G wireless networking. This work presents an uncoordinated non-orthogonal random-access protocol, which is an enhancement to the recently introduced slotted ALOHA-NOMA (SAN) protocol that provides high throughput, while being matched to the low complexity requirements and the sporadic traffic pattern of IoT devices. Under ideal conditions it has been shown that slotted ALOHA-NOMA (SAN), using power-domain orthogonality, can significantly increase the throughput using SIC (Successive Interference Cancellation) to enable correct reception of multiple simultaneous transmitted signals. For this ideal performance, the enhanced SAN receiver adaptively learns the number of active devices (which is not known a priori) using a form of multi-hypothesis testing. For small numbers of simultaneous transmissions, it is shown that there can be substantial throughput gain of 5.5 dB relative to slotted ALOHA (SA) for 0.07 probability of transmission and up to 3 active transmitters.

As a further enhancement to SAN protocol, the SAN with beamforming (BF-SAN) protocol was proposed. The BF-SAN protocol uses beamforming to significantly improve the throughput to 1.31 compared with 0.36 in conventional slotted ALOHA when 6 active IoT devices can be successfully separated using 2×2 MIMO and a SIC (Successive Interference Cancellation) receiver with 3 optimum power levels. The simulation results further show that the proposed protocol achieves higher throughput than SAN with a lower average channel access delay.
In the second part of this dissertation a novel Machine Learning (ML) approach was applied for proactive mobility management in 5G Virtual Cell (VC) wireless networks. Providing seamless mobility and a uniform user experience, independent of location, is an important challenge for 5G wireless networks. The combination of Coordinated Multipoint (CoMP) networks and Virtual-Cells (VCs) are expected to play an important role in achieving high throughput independent of the mobile’s location by mitigating inter-cell interference and enhancing the cell-edge user throughput. User-specific VCs will distinguish the physical cell from a broader area where the user can roam without the need for handoff, and may communicate with any Base Station (BS) in the VC area. However, this requires rapid decision making for the formation of VCs. In this work, a novel algorithm based on a form of Recurrent Neural Networks (RNNs) called Gated Recurrent Units (GRUs) is used for predicting the triggering condition for forming VCs via enabling Coordinated Multipoint (CoMP) transmission. Simulation results show that based on the sequences of Received Signal Strength (RSS) values of different mobile nodes used for training the RNN, the future RSS values from the closest three BSs can be accurately predicted using GRU, which is then used for making proactive decisions on enabling CoMP transmission and forming VCs.

Finally, the work in the last part of this dissertation was directed towards applying Bayesian games for cell selection/user association in 5G Heterogenous networks to achieve the 5G goal of low latency communication. Expanding the cellular ecosystem to support an immense number of connected devices and creating a platform that accommodates a wide range of emerging services of different traffic types and Quality of Service (QoS) metrics are among the 5G’s headline features. One of the key 5G performance metrics is ultra-low latency to enable new delay-sensitive use cases. Some network architectural amendments are proposed to achieve the 5G ultra-low
latency objective. With these paradigm shifts in system architecture, it is of cardinal importance to rethink the cell selection / user association process to achieve substantial improvement in system performance over conventional maximum signal-to-interference plus noise ratio (Max-SINR) and Cell Range Expansion (CRE) algorithms employed in Long Term Evolution-Advanced (LTE-Advanced). In this work, a novel Bayesian cell selection / user association algorithm, incorporating the access nodes capabilities and the user equipment (UE) traffic type, is proposed in order to maximize the probability of proper association and consequently enhance the system performance in terms of achieved latency. Simulation results show that Bayesian game approach attains the 5G low end-to-end latency target with a probability exceeding 80%
CHAPTER 1: INTRODUCTION

Fifth Generation (5G) wireless networks are envisioned as a holistic system that supports a wide range of applications with challenging specifications. The applications which will be supported by the 5G cellular network are categorized, according to their requirements, into the following three main use cases, massive Machine-Type Communication (mMTC), enhanced Mobile Broadband (eMBB) and Low-Latency Communication (LLC). The mMTC use case is necessary for supporting the envisioned 5G Internet of Things (IoT) scenarios, where a massive number of IoT devices with low complexity and limited power are connected over a shared wireless medium. Designing scalable medium access control (MAC) protocols that can meet the high throughput requirements and are compatible with low complexity and limited power specifications IoT devices is critical for mMTC. Also, to meet the high data rates and capacity requirements expected from eMBB, the concept of Virtual-Cell (VC) or “no cell” has been proposed as a key enabler for 5G new radio (NR) systems [1], [2]. In the Virtual-Cell concept a group of cooperating base-stations (BSs) is adaptively and dynamically formed so that the user can achieve a uniform Quality of Service (QoS) independent of its location in the network. Last, but certainly not least, one of the most challenging 5G objectives within the Low-Latency Communication (LLC) use-case is to meet the extremely low 5G latency requirement of 1 millisecond for some applications.

Notably, the infrastructure and technologies of the current 4G wireless networks will not be able to handle these requirements. Therefore, to meet the envisioned 5G objectives, it is
important to explore new technologies and paradigms for future 5G wireless networks. Next, the research areas investigated in this dissertation and a summary of our contributions are described.

### 1.1. Massive Machine-Type Communication (mMTC)

The rapid growth of both the number of connected devices and the data volume that is expected to be associated with Internet of Things (IoT) applications, has increased the popularity of Machine-to-Machine (M2M) type communication within the Fifth Generation (5G) wireless communication systems [3]. Hence, it is necessary to rethink the medium access control (MAC) protocol to match the massive amount of devices accessing the shared medium and the limited-power and low complexity requirements of IoT devices. As illustrated in Figure 1.1, coordinated medium access protocols, such as time-division multiple access (TDMA), frequency-division multiple access (FDMA), or code-division multiple access (CDMA) are contention-free protocols via ensuring that resources are exclusively scheduled for each device.

![Figure 1.1](image_url)

Figure 1.1. Illustration of coordinated medium access among multiple devices: time-division multiple access (TDMA), frequency-division multiple access (FDMA) and code-division multiple access (CDMA), where B is the total bandwidth and T is the total transmission time.

However, the capacity of these protocols is not scalable and cannot provide sufficient throughput to meet the demands for a massive number of IoT devices due to the empty slots and the scheduling overhead required for the connection establishment before each transmission.
Therefore, uncoordinated random access protocols have attracted lots of attention in the standards as a possible method for supporting a massive number of M2M communication of short packets with a low signaling overhead [4], [5].

On the other hand, uncoordinated medium access protocols such as ALOHA and slotted ALOHA (SA) only perform well in small networks, and cannot provide sufficient throughput in networks with large numbers of IoT devices transmitting simultaneously over the shared medium due to collisions at the IoT gateway. The same issues exist in IEEE 802.11, which is based on carrier-sense multiple access with collision avoidance (CSMA/CA) where collisions can still happen due to the hidden and exposed terminal problems. Moreover, CSMA/CA is not compatible with the limited-power requirements of IoT devices since it requires continuous channel monitoring.

1.2. Enhanced Mobile Broadband (eMBB)

The Fifth Generation (5G) of mobile networks using Millimeter Wave (mmWave) technology is anticipated to deliver a substantial increase in the rates of data traffic over the cellular network, as much as 10 Gbps compared to 100 Mbps in 4G networks [6]. For instance, such applications as streaming Ultra-High Definition (UHD) video, Augmented Reality (AR) and Virtual Reality (VR), which have emerged under the 5G Enhanced Mobile Broadband (eMBB) use case, require very high throughput rates everywhere even at the cell edges (i.e., providing a uniform user experience) [6].

However, degradation in the throughput likely will occur, particularly in cellular systems with a frequency reuse of unity, as the user equipment (UE) moves towards the cell edge due to several factors such as the path loss and the interference from neighboring cells. This degradation
in the perceived throughput could greatly undermine the quality of a real-time applications that requires a very high throughput.

Furthermore, in order to maintain the connectivity of a UE with the network, a handover (HO) process takes place by changing the association of the UE to another cell with a better signal quality. During the HO process, the UE might also experience a degradation in the quality of service as a result of the HO delay. Thus, providing seamless mobility and a reliable quality of service anywhere in the network for mobile users is an important challenge in 5G cellular networks.

Many solutions have been proposed in the 3GPP standards, such as using self-organizing networks (SON) in order to enhance the cell-edge user throughput via mitigating the inter-cell interference. For instance, an Inter-Cell Interference Coordination (ICIC) technique [7] proposed in Long Term Evolution (LTE), Release 8, and the enhanced ICIC (eICIC) technique proposed [8] in LTE-Advanced (LTE-A), Releases 10 and 11, exploit the interference information exchanged between adjacent cells via the X2 backhaul interface to improve the cell edge user’s throughput by adaptively muting resources that cause strong interference and assigning different frequencies and powers for the cell edge users. On the other hand, these techniques can result of a degradation in the overall system throughput due to restricting the usage of radio resources in time and frequency.

Coordinated multipoint (CoMP) transmission [9] is another proposed solution that is also based on cooperation among cells. However, unlike the previous techniques where the UE communicates with only one base station, CoMP can allow cell edge users to communicate with multiple BSs (forming a CoMP cooperating set) and hence improve the throughput of cell edge
users and the overall network. CoMP was first standardized in Long Term Evolution-Advanced (LTE-A), Releases 11 and 12.

5G systems are envisioned to further leverage CoMP technology for providing a uniform user experience anywhere in the network. In contrast to the cell-centric approach used in traditional cellular networks with fixed cell coverage areas, a new “cellular” paradigm is emerging, which is a user-centric approach, where the UE is associated with multiple BSs and creates a Virtual Cell (VC) [10] that is adapted according to the mobility of UEs. For instance, using coordinated joint transmission-CoMP (JT-CoMP) formation of VCs in the downlink (DL) can help to attain a uniform user experience for mobile users via improving the throughput of cell-edge users and minimizing the number of hard handovers. The concept of a Virtual Cell is illustrated in Figure 1.2. Virtual Cells break the traditional concept of cell boundaries, via coordinating the data transmission among multiple BSs to jointly serve the users and eliminate the interference from neighboring BSs.

--- Signal
--- Interference
cell boundaries

Coordinated Transmission

Central Unit (CU)

virtual cell

Figure 1.2. Evolution of a classic cellular wireless network to a virtual cell network.
1.3. **Low Latency Communication (LLC)**

The Fifth-Generation (5G) networks are expected to support a broad domain of emerging services with various Quality of Service (QoS) requirements, i.e., from narrow-bandwidth, delay-sensitive services to bandwidth-hungry, delay-tolerant services. To address these disparate services, several solutions have recently been proposed. For instance, Coordinated Multi-Point (CoMP) \[11\] is envisioned as one of the prominent 5G solutions by improving the service for cell edge users suffering from high levels of interference particularly in multitier networks. CoMP has been standardized in LTE-Advanced (Release.11) \[12\].

Moreover, 5G is anticipated to witness an increase in the heterogeneity and density of access nodes (ANs) (e.g. ultra-dense networks (UDNs)) \[13\] as a solution to cope with the tremendous growth in the number of devices connected to the network and consequently boost the system’s capacity \[14\]. However, in UDNs the multi-tier interference becomes more severe.

Hence, the radio resource allocation process should be carried out in a central unit (CU), while taking into account the bigger network picture \[15\]. For instance, \[16\] proposes exploiting the temporal and spatial traffic fluctuations in the network to reduce the interference levels by turning off some ANs with low or no traffic loads.

Recently, the Cloud-Radio Access Network (C-RAN) architecture has attracted attention as a key enabler in implementing interference avoidance and cancellation algorithm for flexible multi-tier 5G networks \[17\], as shown in Figure1.3. Despite the C-RAN’s centralization throughput gain, which is achieved by pooling the computational resources, there is a challenge regarding the delay-sensitive traffic due to the large network latency accruing from transferring the data through the core network and the Internet backbone.
Figure 1.3. Evolution of RAN architecture to C-RAN architecture that consists of distributed remote radio heads (RRHs) and a centralized baseband units (BBUs).

To satisfy the low latency requirements of delay-sensitive traffic, some computing processes, application servers and caching capabilities can be migrated from the cloud to the edge of the network, as illustrated in Figure 1.4, as an evolved architecture [18]. This architecture is commonly referred to as a Fog-Radio Access Network (F-RAN) [19], [20]. Hence, some of the traffic, in particular delay-sensitive traffic, will be served at the FRAN nodes and do not need to travel to the core network and the Internet. This results in a significant reduction in the network latency [21].

Figure 1.4. Fog radio access network (F-RAN) consisting of fog low power nodes (F-LPNs) each provided with limited local storage, BBU pool and real time computation.
1.4. Contributions and Organization of This Dissertation

The research topics that are studied in this dissertation along with the titles of the main chapters are illustrated in Figure 1.5.

Figure 1.5. Outline of the dissertation including the research topics and the names of chapters.

The contributions presented in this dissertation can be summarized as follows:

- A novel medium access protocol dubbed Slotted ALOHA NOMA (SAN) is designed for large scale IoT networks [22], [23]. The SAN protocol is easy to implement, energy efficient, scalable and compatible with the low complexity requirements of IoT devices. The synergistic combination of Slotted ALOHA (SA) protocol with NOMA and Successive Interference Cancellation (SIC) receivers was demonstrated to significantly improve the throughput performance with respect to the conventional SA protocol. The SAN protocol uses a form of multiple hypothesis testing for determining the number of active IoT devices in the medium. Knowing the number of active IoT devices at the IoT gateway is essential in order to optimize the SIC power levels to distinguish between the signals from different IoT devices that are transmitting on the same time and frequency. SAN has a substantial throughput gain of 5.5 dB relative to SA, when using 3 optimum power levels and a total of 60 IoT devices with 0.07 probability of
transmission. We compared the performance of SAN with the CSMA/CA protocol when deployed in large scale IoT networks [24]. The SAN protocol demonstrates superior throughput performance over CSMA/CA. Moreover, in [25], [26] we addressed the problem of channel access delay in SAN protocol by reducing the probability of collision and consequently lowering the average back-off delay via using Multiple Input Multiple Output-Beam Forming (MIMO-BF).

- Deep machine learning is used for proactive mobility management in 5G wireless networks. Proactively predicting the triggering conditions to enable/disable Coordinated Multipoint (CoMP) transmission and Virtual-Cell (VC) selection is investigated [26], [27]. A Recurrent Neural Network (RNN) based on the Gated Recurrent Units (GRUs) mechanism is applied to the feature vectors of the Received Signal Strength (RSS) values from all the Base Stations (BSs) in the network to make an accurate prediction about the RSS values in the future. The proposed algorithm is a promising approach to achieving the goal of providing a uniform user experience everywhere in a network with dynamic formation of VCs. Moreover, considering the very low-latency requirements of next-generation 5G networks, the proposed algorithm helps the mobility management function facilitate VC formations.

- A novel game theoretical approach based on Bayesian games is created for cell selection/user association in 5G Heterogeneous Networks (HetNets) [28]. The utility functions of the user equipment and the network are defined based on the achievable data rate, the average Round Trip Time (RTT) delay, and the access node’s traffic preferences. The proposed Bayesian game algorithm provides a significant improvement in terms of the probability of proper association (as a function of traffic type) and the achieved latency compared with Maximum-Signal to Interference to Noise Ratio (Max-SINR) criterion and Cell Range Expansion (CRE) approach. Such a methodology can be quite important in achieving the 5G low latency objective.
This dissertation is organized as follows. Chapter 2 presents a literature review of the different research areas in this dissertation. A novel medium access protocol for IoT networks and simulation results for SAN and SAN with Multiple Input Multiple Output Beam-Forming (MIMO-BF) are presented in Chapter 3. An innovative algorithm based on RNN-GRUs for predicting the triggering conditions for enabling/disabling CoMP mode in 5G wireless networks is presented and evaluated in Chapter 4. A Bayesian game approach is studied in Chapter 5 for low latency communication in 5G HetNets. Chapter 6 concludes the dissertation and provides some future research directions.
CHAPTER 2: LITERATURE REVIEW

In this chapter, we present a literature review and some of the necessary background information needed for our research. In Section 2.1, the concept of the Internet of Things (IoT) and the associated challenges are described along with some of the fundamental medium access schemes and Random Access (RA) protocols. In Section 2.2, some of the state-of-the-art work on applying Machine Learning (ML) techniques are reviewed with a focus on application to future wireless networks. In Section 2.3, some of the current cell selection/user association approaches are described. Two of the proposed Radio Access Network (RAN) paradigms for 5G networks are also presented.

2.1. Topic 1: Medium Access Protocols for Internet of Things (IoT) Networks

2.1.1 Internet of Things (IoT)

The term Internet of Things (IoT) has gained a lot of popularity in recent years, as a consequent to the anticipated increase in the number of devices connected to the Internet. According to Ericsson, the number of connected devices to the internet is expected to reach 28 billion devices by 2021 [29]. IoT technology transforms physical objects “things” from traditionally being inactive to interactive and accessible from anywhere through the internet [30], [31], [32], [33]. As shown in Figure 2.1 smart homes are one of the main applications of IoT for an improved personal lifestyle. For instance, in smart homes it becomes possible by interacting with the smart refrigerators to send a message whenever a food item needs to be replenished, switching on the lights, adjusting the temperature, closing the garage door, and controlling other smart appliances remotely. The data generated by these smart home appliances will be transferred
through the IoT gateway to the Internet and then eventually to a smart home cloud server. Realizing the smart home vision can only be achieved through designing a medium access control (MAC) protocol to provide efficient connectivity for large number of devices. There are considerations that need to be taken into account when designing medium access control (MAC) protocols for this class of IoT networks.

First, these smart home IoT devices generate payloads that are often smaller than the typical signaling overhead. Secondly, these IoT devices are energy-limited, since some of these IoT devices are battery powered sensor nodes and have low-complexity in terms of computation and processing capabilities. Moreover, one of the important considerations in designing MAC protocols for IoT networks is scalability, which can be defined as the feasibility of supporting massive number of devices using the same network and being able to accommodate devices newly joining the network without a large increase in complexity. This massive number of connected devices and the sporadic nature of their data traffic can lead to a severe congestion in medium access channels. Hence, this severe congestion will increase the delay for the channel access and will increase the number of retransmissions which eventually causes an increase in energy expenditure too.

Figure 2.1. The use case of an IoT managed smart home.
2.1.2 Medium Access Control (MAC) Protocols

In the IoT literature, the IoT devices are referred to as machines, and the communication among these machines is referred to as Machine-to-Machine (M2M) communication or Machine Type Communication (MTC). The MAC protocols used for M2M communication in IoT networks can mainly be classified into two main categories as contention-free and contention-based protocols.

2.1.2.1 Contention-Free Protocols

One of the solutions for MAC protocols in M2M IoT networks is contention-free medium access protocols. In contention-free MAC protocols, collisions are avoided by ensuring that the resources are allocated exclusively to IoT devices, which can be achieved either via fixed or dynamic assignment strategies. Although fixed resource assignment strategies as in TDMA, FDMA or CDMA can avoid collisions, they are not scalable and inefficient. For instance, in TDMA the total transmission time is divided into time slots and each device can only transmit during its assigned time slot, and it is impossible to reallocate time slots assigned for one device to other devices if the time slot is not used. With dynamic assignment strategies, the IoT devices can access the medium on demand. The controller (IoT gateway) sends a control-overhead message asking each IoT device if it has data to send, to reserve the transmission medium for only one device at a time. If there is no data to be sent, the controller sends the control message again for the next IoT device, and so on.

Therefore, contention-free MAC protocols cannot provide high throughput to meet the demands of large numbers of IoT devices due to control overhead and the unused empty slots. For those reasons, contention-free MAC protocols are not scalable, though they can prevent collisions.
2.1.2.2 Contention-Based Protocols

Contestion-based MAC protocols (or uncoordinated random access schemes) have attracted lots of attention as a possible method for making massive M2M communication possible with a low signaling overhead [34], [35]. Contention based protocols such as Carrier Sense Multiple Access/Collison Avoidance (CSMA/CA) used in IEEE 802.11 perform well only for small networks and they do not have sufficient throughput for large scale networks due to the frequency of collisions. The same issues are valid for the familiar ALOHA and slotted ALOHA (SA) protocols. Moreover, CSMA/CA is energy inefficient, since it requires continuous channel monitoring and there is a significant overhead due to control packets that are not compatible with the limited battery requirements of IoT devices. Furthermore, slotted ALOHA protocols or their variants such as diversity slotted ALOHA (SA) [36] are not scalable approaches, because they require tight synchronization of the IoT devices in time domain, which can belong to different providers within the framework of heterogenous IoT networking.

2.1.2.2.1 Overview of ALOHA-Based Protocols

The ALOHA protocol, which is also referred to as pure ALOHA, was proposed by Abramson, from the University of Hawaii, as Random Access (RA) protocol for providing wireless connectivity between computers at different islands in Hawaii [37]. In wireless networks operating using the ALOHA protocol, the devices in the network access the shared wireless medium and transmit packets of different durations, independently and asynchronously. In ALOHA systems, a transmitted packet is received and decoded successfully when no other packets are received within the duration of the packet. When packets collide, and data is lost when two or more of the transmitted packets interfere with each other, as illustrated in Figure. 2.2. The IoT gateway sends an acknowledgement via a feedback channel, when the packets are received and decoded.
successfully. In case an IoT device did not receive an acknowledgement from the IoT gateway within a certain time frame, it resends the same packet after “backing off” for a randomly selected period of time.

Figure 2.2. Schematic diagram of ALOHA protocol with M IoT devices. The areas under the red dashed-lines correspond to the overlapping portions of transmitted packets from different IoT devices.

The inter-arrival times of the new packets and the retransmission times of the collided packets are modeled as independent processes that are exponentially distributed. It is worth mentioning that ALOHA protocol has a small probability of collision when the number of active IoT devices in the network and the duration of packets are small. In particular, for an average packet duration of $\tau$ seconds and an arrival rate of $\lambda$ packets per second, the throughput can be expressed as $\lambda\tau e^{-2\lambda\tau}$. Then, the maximum throughput (collision-free) can be calculated as $\frac{1}{2}e^{-1} \approx 0.18$. Therefore, ALOHA protocol is not a suitable solution when the number of active IoT devices grows massively.

Later, Slotted ALOHA (SA) was then proposed by Roberts and Abramson [38], [39]. In SA the time is divided into equal time slots that are equal to the packet duration $\tau$. The IoT devices can start transmission only at the beginning of any time slot, as shown in Figure 2.3. The maximum throughput of SA is $e^{-1} \approx 0.368$, which is double the maximum throughput of ALOHA protocol.
However, SA requires strict synchronization which requires additional overhead compared with ALOHA protocol.

Figure 2.3. Schematic diagram of slotted ALOHA (SA) protocol with M IoT devices. The areas under the red dashed-lines correspond to the overlapping portions of transmitted packets from different IoT devices.

Some efforts have been done to improve the efficiency of SA protocol. In [40], Reservation ALOHA (R-ALOHA) is introduced as an enhancement to the SA protocol. R-ALOHA protocol, as shown in Figure 2.4, has two modes of operation: contention mode (un-reserved) and contention-free mode (reserved). In the contention mode the time is divided into small time sub-slots. During the contention mode, the IoT devices that have packets ready for transmission, first transmit short reservation requests using SA protocol over the time sub-slots. An IoT device can request one time slot or more within the reservation request sent. Then, IoT devices needs to wait for an Acknowledgement (ACK) sent by the IoT gateway to inform each IoT device on where to allocate their packets in time. Following the contention mode, the system switches to contention-free mode where the time is divided into equally spaced time slots.
The premise which the ALOHA and the Slotted-ALOHA protocols were built on is that when two or more packets are transmitted during the same time slot, they will interfere (collide) with each other and the data will get lost. Therefore, in order to have a successful transmission, only one device should be transmitting at each time slot. In [36], the Contention Resolution Diversity Slotted ALOHA (CRDSA) protocol uses successive interference cancellation (SIC) receiver at the IoT gateway to resolve the collisions that might happen when two packets or more are transmit during the same time slot. In CRDSA, the IoT devices transmit replicas of their packets at different time slots. A pointer is included within each of the transmitted packets, indicating the location of the transmitted replicas in time.

![Figure 2.4. Schematic diagram of reservation ALOHA (R-ALOHA) protocol.](image)

For instance, as illustrated in Figure 2.5a, there are three packets transmitted from the IoT devices 1, 2 and 3 during the first, third and fourth time-slots. Only the packet transmitted from the second IoT-device (gray packet) is received successfully (collision-free) during the fourth time-slot. After successfully detecting the packet transmitted by the second IoT device, the IoT gateway then reads the pointer sent at the header of the packet to find at what time-slot the replica is been transmitted.

The IoT gateway uses this information to cancel the interference from the gray packet in the first time-slot. Assuming that the gray packet can be perfectly cancelled in the first time-slot (collision resolved), then the green packet can be decoded at the first time slot, as illustrated in
Figure 2.5b. Likewise, the blue packet can be decoded in the third time-slot, as shown in Figure 2.5c, at the cost of delay.

Figure 2.5. An illustration of slotted ALOHA with successive interference cancellation (SIC) receiver at the IoT gateway.

2.1.2.2 An Overview on CSMA/CA Protocol

The Carrier Sense Multiple Access/Collision Avoidance (CSMA/CA) protocol was proposed as a mechanism to avoid collisions in random access protocols. The collision avoidance comes at the expense of energy inefficiency and the complexity of the IoT devices. The IoT devices keep sensing the wireless medium before making any transmission as precautionary measure to minimize collisions in the system. Moreover, in CSMA/CA the back-off time (the waiting time before attempting to transmit again) follows an exponential distribution. The retransmissions are spaced in time, according to an exponential distribution and as a function of the number of
collisions, in order to minimize the chances for another collisions. In particular, the retransmissions process should happen less-frequently (become more spaced in time) when the collisions start occurring more.

However, as mentioned above, CSMA/CA is energy inefficient, since it requires continuous channel monitoring and there is a significant overhead due to control packets that are not compatible with the limited battery requirements of IoT devices. CSMA/CA also suffers from hidden-terminal and exposed terminal problems [41]. Moreover, CSMA/CA performs like other contention-based protocols that perform well only for small networks and do not have sufficient throughput for large scale networks due to the frequency of collisions.

Our contribution in this topic is a novel medium access protocol that significantly increases the throughput by 5.5 dB in comparison with Slotted ALOHA protocol that makes it suitable for large scale IoT networks. Moreover, the proposed MAC layer protocol is easy to implement, energy efficient, scalable and compatible with the low complexity requirements of IoT devices.

2.2. Topic 2: Deep Machine Learning in 5G Virtual Cell Networks

Recently, there has been an increased interest for applying Machine Learning (ML) techniques to future 5G wireless networks. Unlike traditional deterministic rule-based optimization techniques, ML algorithms are real-time optimization algorithms that are adaptable and can cope with the ever-changing wireless environment. In a nutshell, ML techniques aim to deduce the best decisions to be made by the network via analyzing the prior collected data. ML techniques can be categorized as supervised learning, unsupervised learning, and reinforcement learning. In supervised learning the machine is trained using labeled input data with the goal of learning the function that maps an input to an output. Whereas, in unsupervised learning the input data is unlabeled, and the job of the machine is to extract features and identify patterns in the
dataset to label the data by itself. Reinforcement learning is also another category of ML, where the machine learns by try and error through its interactions with the environment. In reinforcement learning the "machine" receives rewards for performing positive actions and penalties for performing negative actions. Unlike, the unsupervised learning which is used to identify patterns in unlabeled data, the goal in reinforcement learning is to find a sequence of actions that can maximize the total cumulative reward.

Recently, there have been proposals using ML in predicting the next base station (BS) a user will be associated with. In particular, ML is used for mobility prediction that is utilized for making optimal hand-over (HO) decisions in wireless networks. In [42], the authors used Support Vector Machines (SVMs) for predicting the next cell ID based on certain features. They started first by using a short sequence of previously visited cell IDs and achieved an accuracy of 78%. However, the accuracy of classification did not improve much after the length of the sequence was increased to more than three. Then after the authors added more features such as the time duration that a user was associated with a particular cell, users GPS coordinates and IDs, the classification accuracy did not improve and remained in the range between 77 - 78%. Although, their approach was able to predict where a user would be associated with next, it was unable to predict the when this transition would happen. Hence, they further used undirected graphical models for predicting the next cell ID and the time a user would be associated with. They were still only able to achieve an accuracy of 78%.

In [43], Tkacik and Kordik have used a Neural Turing Machine (NTM) to predict the mobility of criminal suspects. NTMs utilize a memory matrix with selective write and read operations in order to remember data and have dynamic state. The authors validated their proposed method both with synthetic data generated for seven people and real data recorded from six people.
over a period of three months. They defined their own accuracy metric and showed that it outperformed a k-Nearest Neighbor (kNN) classifier. The highest accuracy they obtained in predicting locations was just over 82%, which was for one specific user.

In [44], Luo et al. used a Nonlinear Autoregressive Exogenous Model (NARX) based on a Neural Network (NN) for HO prediction. The received signal strength (RSS) values and the delays between Access Points (APs) are fed to an NN with 12 hidden layers. The NN was able to predict a point which was close to the optimal HO location. Javed et al. [45] used signal strength variation, the past HO rate, number of cells with RSS values above a certain threshold (active set) and active set update rate as features for the AdaBoost algorithm with Decision Stumps to predict the occurrence of a HO. Liou and Huang [46] used the past three location coordinates as features for a Neural Location Indicator for predicting the node’s mobility and inferring HOs. Anagnostopoulos et al. [47] used the last four previous cell IDs a user had visited, in four-element feature vectors, for predicting the user’s next cell ID. They have compared the performance of k-Nearest Neighbors (kNN), Naïve Bayes, Bagging and a vote of kNN, J48 Decision Trees, rule-based learner (JRip) and AdaBoost. In [48], Anagnostopoulos et al. incorporated transition time slots into feature vectors. However, the obtained accuracy on the next cell predictions was very similar to [47].

However, none of the above-proposed ML techniques consider a mobile node to be concurrently associated with multiple base stations (BSs) in creating Virtual Cells (VCs). In [49], Wickramasuriya et al. used sequences of Received Signal Strength (RSS) values as features for training a Recurrent Neural Network (RNN) classifier based on Long Short-Term Memory (LSTM) to predict the next base station (BS) that a mobile node will be associated with and the optimal VC topology according to the mobility of users. While the authors achieved an accuracy
of 98 %, they did not consider predicting the optimal triggering conditions for enabling/disabling coordinated multipoint (CoMP) transmission based on the mobility of users - which is the goal of our work.

Our contribution in this topic is a novel algorithm for proactively predicting the optimal triggering conditions that can be used to dynamically enable/disable CoMP transmission instead of relying on passive approaches. The proposed algorithm is based on using Recurrent Neural Networks (RNNs) with a Gated Recurrent Units (GRUs) mechanism.

2.3. **Topic 3: Cell Selection/User Association in 5G Heterogenous Networks**

As a natural result of the expected increase of heterogeneity in 5G networks, as we earlier discussed in CHAPTER 1, the user equipment (UE) will have a large range of connectivity options with different requirements such as power consumption, latency budget, and the achievable data rate. To best exploit the available opportunities, both the UE’s traffic type and the capabilities of available Access Nodes (ANs) should be taken into account at the cell selection/user association stage. As illustrated in Figure 2.6, the process of cell selection/user association in traditional cellular systems is based on the AN that can provide the highest signal-to-interference-plus-noise ratio (SINR) [50]. However, this approach is generally not optimum in multi-tier networks with diverse traffic. To alleviate this problem, [51] introduces the cell range expansion (CRE) approach for Low Power Nodes (LPNs) via a biasing method such that the High-Power Node (HPN) transmit power is reduced on a group of sub-carriers in order to enable better coverage on the same group of sub-carriers for an overlaid LPN, as shown in Figure 2.7. Another approach for cell association is to employ user-perceived rate considering the SINR and the network load [52].

However, these papers did not take into account the different types of traffic. Indeed, the selection and association procedure based only on the maximum SINR and CRE criteria might
degrade the performance of the system from a latency perspective (i.e. when UEs of delay-tolerant traffic get associated with Fog-Low Power Nodes (F-LPNs) or when UEs of delay-sensitive traffic are associated with a HPN). Our contribution in this topic is a novel Bayesian cell selection / user association algorithm, incorporating the access nodes capabilities and the user equipment (UE) traffic type, in order to maximize the probability of proper association and consequently enhance the system performance in terms of achieved latency.

Figure 2.6. Reference signal received power (RSRP)-based cell selection approach used in long term evolution (LTE) systems.

Figure 2.7. Cell range expansion (CRE) approach.
CHAPTER 3: SLOTTED ALOHA NOMA (SAN) PROTOCOL FOR IoT NETWORKS

3.1. Introduction

Non-Orthogonal Multiple Access (NOMA) has emerged as a promising technology in 5G networks for many applications [53], [54]. The proposed Slotted ALOHA NOMA (SAN) protocol is a synergistic combination of the low complexity Slotted ALOHA (SA) protocol with the high throughput feature of NOMA [22], [23]. The main bottleneck of SA systems is the low throughput caused by the high number of collisions, which can be addressed by NOMA. In SAN the signaling overhead is reduced in the detection phase of the proposed protocol where the number of active IoT devices are detected by the gateway using a form of multiple hypotheses testing, which is further explained in Section 3.4.

The SAN protocol is also an energy efficient protocol due to the fact that a Successive Interference Cancellation (SIC) receiver can resolve collisions under certain conditions [55], [56], and thus minimizes retransmissions and significantly increases the conventional Slotted ALOHA (SA) throughput. The SAN protocol can be suitable for various scenarios where many IoT devices are transmitting simultaneously on the same frequency with different power levels to an IoT

---

1 This chapter is an enhanced version of


gateway since the received signals can be reliably detected via use of Successive Interference Cancellation (SIC) receiver.

The structure of this chapter is as follows. In Section 3.2, an overview of Power Domain Non-Orthogonal Multiple Access (PD-NOMA) is presented. The frame structure of the proposed SAN protocol is discussed in Section 3.3 and Section 3.4 describes detecting the number of active IoT devices using a form of multiple hypothesis testing [22], [23]. In Section 3.5, simulation results are presented to show the superiority of the proposed SAN method regarding throughput with respect to Slotted ALOHA (SA). Also, additional simulation results are presented that analyze the throughput of the SAN protocol as function of the number optimum power levels and compare the SAN protocol with conventional SA in terms of channel access delay. Lastly, Section 3.6 presents a further enhancement of the SAN protocol using “MIMO” Beam-Forming (BF) to reduce the channel access delay [25]. Furthermore, simulation results are performed in Section 3.6 to show that SAN with beamforming outperforms both SAN and Slotted ALOHA protocols in terms of throughput and channel access delay without excessively increasing the SIC power levels.

3.2. Overview on Non-Orthogonal Multiple Access (NOMA)

The Internet of Things (IoT) is a key part in the future 5G network. Therefore, it was essential to rethink the medium access (MA) schemes to meet the requirements of high spectrum efficiency while supporting a massive number of connected devices. Non-Orthogonal Multiple Access (NOMA) is a promising scheme that provides better spectrum efficiency using Successive Interference Cancellation (SIC) at the receiver [57].

It is worth mentioning that NOMA in the power domain has been proposed for inclusion in Long Term Evolution-Advanced (LTE-A) for the downlink as Multiuser Superposition Transmission (MUST) [58]. As shown in Figure 3.1, unlike Orthogonal Multiple Access (OMA)
schemes (such as TDMA, FDMA and CDMA), NOMA can improve the capacity of the system by simultaneously serving multiple devices over the same time/frequency resources using different power levels that are separated at the receiver using a SIC processor.

Orthogonal Multiple Access (OMA) is a well-known technique that is free of inter-device interference within each resource block. On the other hand, NOMA enhances the system capacity compared with OMA through sharing the same resource block by multiple devices. The enhancement in system capacity is achieved by exploiting the differences in channels between devices. However, there is no "free-lunch", this enhancement is obtained at the cost of employing algorithms at the receiver to iteratively cancel the interference from other devices.

Moreover, using NOMA provides each device a transmission bandwidth that is relatively wider compared to OMA. In Figure 3.2 the capacity comparison using NOMA and OMA techniques is illustrated when the channels for the two devices A and B are different. We can see that the boundary of achievable rate pairs with NOMA is outside the OMA capacity region.

![Diagram](image)

Figure 3.1. Orthogonal multiple access (OMA) vs. non-orthogonal multiple access schemes. $\beta_A$ and $\beta_B$ are the fractions of the total power for devices A and B. $P_A$ and $P_B$ represents the allocated power levels for devices A and B respectively, $|h_A|$ and $|h_B|$ are the channel gains, $B$ is the total transmission bandwidth, and $\alpha$ is the fraction of bandwidth allocated for device B [59].
Therefore, under the assumption of perfect SIC cancellation, it can be shown that when the Signal-to-Noise Ratios (SNRs) of the two users are different, the rates of devices A and B calculated using NOMA in (3.1) and (3.2) respectively are much greater than the rates calculated using the OMA technique in (3.3) and (3.4) [59].

\[
R_A = \log(1 + \frac{P_A|h_A|^2}{N_0}), \tag{3.1}
\]

\[
R_B = \log(1 + \frac{P_B|h_B|^2}{P_A|h_B|^2 + N_0}), \tag{3.2}
\]

\[
R_A = (1 - \alpha) \log(1 + \frac{P_A|h_A|^2}{(1 - \alpha) N_0}), \tag{3.3}
\]

\[
R_B = \alpha \log(1 + \frac{P_B|h_B|^2}{\alpha N_0}). \tag{3.4}
\]

Figure 3.2. OMA vs. NOMA capacity regions [59].

A downlink scenario of NOMA is illustrated in Figure 3.3. In this example, device B is presumed to be the device with poor channel gain and device A is the device with the good channel gain. The key idea is that the access point (AP) first sorts “orders” the channels of the devices and exploits the power domain by allocating more power to device B and less power to device A. Then
after that, the Quadrature Phase Shift Keying (QPSK) constellation of device A is superposed on QPSK constellation of device B using superposition coding which was first proposed by Cover for broadcast channels [60].

Then the AP sends a linear superposition of device’s A and B data, where $\beta_A$ and $\beta_B$ are the fractions of the total power allocated to devices A and B respectively. At the receiving ends, device A receives the superposed signal $\mathbf{r}_A$. As shown in Figure 3.4, device A (with the better channel gain) uses the successive interference cancellation (SIC) receiver. Device A will first decode the signal of device B and then subtracts the device B message from the superposed received signal $\mathbf{r}_A$ to decode its own message $\mathbf{x}_A$. On the other hand, device B can decode its message $\mathbf{x}_B$ by only treating device A’s signal $\mathbf{x}_A$ as a noise.

Figure 3.3. Superposition coding for quadrature phase shift keying (QPSK) constellation points of devices A and B [59].
3.3. Flexible Frame Structure for the SAN Protocol

One of the main challenges for enabling the practical implementation of the SAN protocol is the assignment of proper power levels for the IoT devices before transmitting the information; otherwise, the signals received from different IoT devices cannot be extracted successively from the composite received signal. This issue becomes more challenging in dynamic environments where the number of IoT devices with information ready to send is continuously changing.

In this section, this challenge is addressed via a flexible frame structure. Such a scheme provides great flexibility in adapting to changing network environments. This structure is in contrast to that of TDMA or FDMA, where arrival of a new user can completely change the overall frame structure such that the additional user must be assigned at least one slot within the frame. As illustrated in Figure 3.5, the proposed frame structure composed of 5 phases.

In the first phase, the IoT gateway transmits a beacon signal to announce its readiness to receive packets. Next, the IoT devices with packets ready to transmit send a request (with “dummy packets”) to aid the gateway in detecting the number of active IoT devices in the medium. The IoT
gateway detects the number of devices requesting transmission via a form of multiple hypotheses testing, as discussed in Section IV, and adjusts the degree of the SIC receiver for the optimum power levels. In practice, the SIC receiver has fixed range of optimum power levels (e.g. m=2,3).

Notice that it would lead to easier implementation if all IoT devices are registered to the gateway instead of relying on multi-hypothesis testing; However, this may significantly increase the length of control phase and thus decrease the payload or throughput considering the large number of IoT devices. If the detected number of active IoT devices is not in the range of optimal power levels, then the IoT gateway aborts the transmission and starts all over again by sending a beacon signal.

Otherwise, the IoT gateway broadcast the degree of SIC to the transmitters and then each one of the active IoT device randomly picks one of the optimum power levels. If the choices are distinct the SIC receiver can decode the self-identifying signals (device ID + payload) and then the gateway sends an ACK. If the active IoT devices could not select distinct power levels, the SIC detection will fail and a reselection process is repeated. If after a few attempts, there is no successful transmission, the users receive a NACK and enter a random back-off mode. It should be clear, that the proposed protocol will be most efficient when there is a small number of active devices, so that the probability of randomly choosing distinct optimum power levels remains high.
Figure 3.5. The proposed SAN protocol frame structure.

3.4. Detecting the Number of Active IoT Devices

This section presents a form of multiple hypothesis testing used to detect the number of active devices. The detection of active devices starts in the second phase of the SAN frame as presented in Fig.2. After receiving the beacon, all the IoT devices send at the same power level using SA a training sequence (known to the IoT gateway) of length $L$. The superposed received signal at the IoT gateway from $N$ active transmitting IoT devices is given by

$$
y = \mathbf{Hs} + \mathbf{w},
$$

where $\mathbf{H} = [h_1, h_2, ..., h_N] \in \mathbb{R}^{1 \times N}$, $h_n$ is the channel gain between the $n^{th}$ IoT device and the IoT gateway, $\mathbf{s} \in \mathbb{R}^{N \times L}$ is the transmit sequence (e.g. BPSK) from $N$ IoT active devices and $\mathbf{w} \in \mathbb{R}^{1 \times L}$ is the additive white Gaussian noise with zero mean and variance $\sigma^2$. In this simplified model, it is assumed that all the channel gains are known at the IoT gateway, so that the received powers can be controlled and set equal. The multiple hypothesis test is used to detect the number of $N$ active IoT devices from the total $M$ IoT devices. The following procedure is used to sequentially detect the number of active devices.
\( \mathcal{H}_0 \): Received signal contains only noise

\[
y = w, \quad (3.2)
\]

\( \mathcal{H}_1 \): Received signal contains data from one IoT device

\[
y = h_1 s_1 + w, \quad (3.3)
\]

\( \vdots \)

\( \mathcal{H}_N \): Received signal contains data from N IoT device

\[
y = Hs + w. \quad (3.4)
\]

We assume \( h_n = 1, \forall n \in \{1, 2, ..., N\} \). Following the Neyman-Pearson (NP) test, we can write the Likelihood Ratio (LR) testing \([\mathcal{H}_N \text{ Vs. } \mathcal{H}_{N-1}\) as

\[
\frac{p(y; \sum_{n=1}^{N} s_n, \mathcal{H}_N)}{p(y; \sum_{n=1}^{N-1} s_n; \mathcal{H}_{N-1})} = \exp \left[ -\frac{1}{2\sigma^2} (y - \sum_{n=1}^{N} h_n s_n)^T (y - \sum_{n=1}^{N} h_n s_n) \right] \geq_{\mathcal{H}_N} \leq_{\mathcal{H}_{N-1}}, \quad (3.5)
\]

where \( y, N = 1, ..., M \) and \( s_n \in \mathbb{R}^{1 \times L} \) is the transmitted sequence from \( n^{th} \) IoT device. By taking the logarithm, (3.5) is simplified to

\[
T(y) = \frac{1}{L} \sum_{l=0}^{L-1} y_{H_{N-1}} \leq_{\mathcal{H}_N} \frac{2\sigma^2}{\mathcal{H}_{N-1}} \ln \frac{\left( \sum_{n=1}^{N-1} h_n s_n \right)^T \left( \sum_{n=1}^{N-1} h_n s_n \right) + \left( \sum_{n=1}^{N} h_n s_n \right)^T \left( \sum_{n=1}^{N} h_n s_n \right)}{-2 \left( \sum_{n=1}^{N-1} h_n s_n + \sum_{n=1}^{N} h_n s_n \right)} = \gamma'. \quad (3.6)
\]

The NP detector, or the test statistic, in (3.6) compares the sample mean of the received signal to the threshold \( \gamma' \) to decide on a hypothesis \( \mathcal{H}_N \) or \( \mathcal{H}_{N-1} \). The NP test terminates if the number of detecting devices exceeds the number of SIC receiver optimum power levels, which are 3 levels in this chapter. To compute the threshold \( \gamma' \) in (4) for a desired probability of false alarm PFA, which occurs when deciding \( \mathcal{H}_N \) if the test in (4) is greater than the threshold \( \gamma' \), so that PFA can be written as
\[ P_{FA} = P(T(y) > y'; \mathcal{H}_N), \quad (3.7) \]

Since the test in (3.6) under both hypothesis is a Gaussian distribution, that
\[ T(y) \sim \mathcal{N}(\sum_{l=0}^{L-1} \sum_{n=1}^{N-1} s_n, \sigma^2) \] under \( \mathcal{H}_{N-1} \) and \( T(y) \sim \mathcal{N}(\sum_{l=0}^{L-1} \sum_{n=1}^{N} s_n, \sigma^2) \) under \( \mathcal{H}_N \) we rewrite (5) as
\[ P_{FA} = Q\left( \frac{y' - \sum_{l=0}^{L-1} \sum_{n=1}^{N} s_n}{\sqrt{\sigma^2/L}} \right). \quad (3.8) \]

Thus, the threshold \( y' \) is given by
\[ y' = Q^{-1}(P_{FA}) \sqrt{\sigma^2/L} + \sum_{l=0}^{L-1} \sum_{n=1}^{N} s_n. \quad (3.9) \]

Following the same steps, the probability of detecting the number of active devices is
\[ P_D = Q\left( \frac{y' - \sum_{l=0}^{L-1} \sum_{n=1}^{N} s_n}{\sqrt{\sigma^2/L}} \right). \quad (3.10) \]

From (3.9), (3.10) we can write the \( P_D \) as a function of energy to noise ratio as
\[ P_D = Q\left( Q^{-1}(P_{FA}) + \frac{\sum_{l=0}^{L-1} \sum_{n=1}^{N} s_n - \sum_{l=0}^{L-1} \sum_{n=1}^{N} s_n}{\sqrt{\sigma^2/L}} \right). \quad (3.11) \]

The probability of correct detection of the number of active users as a function of the SNR for \( P_{FA} = 0.1 \) is shown in Figure 3.2. Observe that the detection performance increases monotonically and smoothly with increasing SNR.
3.5. Simulation Results

In this section simulation results are presented to evaluate the throughput performance of the SAN protocol. Throughout the simulation, we assume there is one IoT gateway with a single antenna and a total of $M$ IoT devices. A binomial distribution is considered to model the random number of active IoT devices $N$, each with probability of transmission $p_T$.

$$P_T(N; p_T, M) = \binom{M}{N} p_T^N (1 - p_T)^{M-N}. \tag{3.12}$$

Figure 3.7 shows the throughput of the SAN protocol for different values of probability of transmission $p_T$, total number of IoT devices $M = 50$ and $k = 3$ attempts for the random selection of distinct optimum power levels. Throughput is the number successful transmissions for each probability of transmission. As expected, we observe that the throughput decreases with increasing probability of transmission. More importantly, we can see that the throughput of the SAN protocol is always higher than that of the SA protocol. In particular, when the probability of transmission is 0.07, the throughput of SAN with 3 power levels becomes almost 3.6 times higher than that of
(conventional) SA. This demonstrates that NOMA with a SIC receiver can help improve the throughput of SA, especially when the IoT devices are active.

![Graph showing throughput vs probability of transmission.](image)

Figure 3.7. Throughput of slotted ALOHA-NOMA (SAN) with 2 and 3 optimum power levels vs slotted ALOHA as a function of probability of transmission.

In order to see the impact of the number of optimum power levels on the throughput of SAN, we show the throughput of SAN for different power levels in Figure 3.8. With $M = 10$, $k = 3$ and probability of transmission of 0.25, the throughput of SAN increases with the increase in optimum power levels (SIC degree). For example, SAN with 3 power levels has a higher throughput than SAN with 2 power levels. However, the throughput improvement becomes negligible for optimum power levels greater than 5 (saturation in the throughput gain). Also, the simulation results shown in Figure 3.8, illustrates the impact of the number of attempts $k$, for picking distinct optimum power levels, on throughput for different optimal power levels. The more attempts allowed for picking the optimum power levels, the higher the throughput that can be achieved at the cost of increased delay.
Figure 3.8. Throughput as a function of optimum power levels for different number of attempts (2, 3 and 5) using SAN protocol.

Figure 3.9 shows the average channel access delay as a function of probability of transmission for the SAN protocol vs. the Slotted ALOHA protocol. It is shown that at a low probability of transmission, the Slotted-ALOHA-NOMA (SAN) protocol has a larger average channel access delay compared with conventional Slotted-ALOHA protocol. While, at a high probability of transmission both ALOHA NOMA\(^2\) and Slotted-ALOHA deliver the same average channel access delay.

\(^2\) ALOHA NOMA and Slotted ALOHA NOMA (SAN) will be used interchangeably in this chapter.
Figure 3.9. Average channel access delay as a function of probability of transmission for the SAN protocol and the well-known slotted ALOHA protocol.

3.6. Slotted ALOHA-NOMA with MIMO Beam-Forming (BF-SAN)

In this section, the simulation results evaluate the throughput and channel access delay performance of the Beamforming Slotted ALOHA-Non-Orthogonal Multiple Access (BF-SAN\textsuperscript{3}) protocol. Throughout the simulation, we assume there is one IoT gateway with $N_r$ receiving antennas and a total of $M$ IoT devices each with $N_t$ transmitting antennas.

Figure 3.10 shows the throughput of BF-SAN for different values of probability of transmission $P_t$, when $M = 60, N_t = 2, N_r = 2, m = 3$ optimum power levels and $k = 3$ attempts for the random selection of distinct optimum power levels. In the simulation, throughput is defined as the average number of packets that are successfully decoded for a given probability of transmission. As expected, we can observe that the throughput decreases as the probability of

\textsuperscript{3} BF-SA-NOMA and BF-SAN will be used interchangeably in this chapter.
transmission increases. Also, we can see that the throughput of the BF- SAN protocol is always higher than that of SAN and conventional Slotted ALOHA protocols. In particular, when the probability of transmission is 0.15, the throughput of BF- SAN is almost 60 times higher than that of conventional Slotted ALOHA and 2 times higher (i.e., the beamforming gain) than SAN protocol.

![Figure 3.10. Throughput as a function of probability of transmission for BF-SAN, SAN and conventional slotted ALOHA protocols.](image)

This demonstrates that BF-SAN provides a dramatic improvement in throughput in comparison with conventional Slotted ALOHA and SAN protocols due to the additional “virtual” sub-channels in the spatial and power domains. More importantly, it is worth mention that achieving a throughput gain similar to that of the proposed BF- SAN using SAN requires an excessive increase in SIC power levels which is impractical when using power-limited IoT devices. In Figure 3.11 and Figure 3.12 we show the beamforming gains in terms of the channel access delay and throughput of BF-SAN for different values of probability of transmission when
$k = 2, 3$ attempts, $N_r = 2$, $N_t = 2$, $M = 60$, and $m = 3$ optimum power levels. In simulation, the channel access delay is composed of three components, namely: round trip delay, delay due to the attempts for random selection of distinct power levels, and the back-off delay (that occurs at the event of the number of active IoT devices is greater than the degree of SIC receiver or when the IoT devices fail in selecting distinct optimum power levels in $k$ attempts). Figure 3.11 and Figure 3.12 illustrates the beamforming gains, where it is shown that BF-SAN when $k = 2$ can still achieve a higher throughput than ALOHA-NOMA when $k = 3$ with a lower channel access delay resulting from reducing the probability of collision and the average back-off delay by creating more virtual sub-channels via MIMO beamforming.

![Figure 3.11. Throughput of BF-SA-NOMA VS. ALOHA-NOMA for different values of probability of transmission when $M = 60$, $m = 3$, $k = 2$ and 3 attempts, $N_t = 2$ and $N_r = 2$.](image-url)
Figure 3.12. Average delay of BF-SA-NOMA VS. ALOHA-NOMA for different values of probability of transmission when $M = 60$, $m = 3$, $k = 2$ and 3 attempts, $N_r = 2$ and $N_t = 2$.

Also, the simulation results in Figure 3.11, shows the impact of the number of attempts $k$, for picking distinct optimum power levels, on throughput for different optimal power levels. The more attempts allowed for picking the optimum power levels, the higher the throughput that can be achieved at the cost of increased delay.
CHAPTER 4: OPTIMIZATION OF 5G VIRTUAL-CELL BASED COORDINATED MULTIPOINT NETWORKS USING DEEP MACHINE LEARNING

4.1. Introduction

As discussed in earlier chapters, the concept of virtual-cell networks has emerged as new paradigm for serving wireless users [61]. In a virtual-cell (VC) architecture, a group of base-stations (or transmission points [TPs]) is created and is updated as the user moves within the network. The users that are within the domain of a virtual-cell can be associated with multiple base-stations (BSs) and their data transmitted (or received) will be coherently processed by multiple BSs. The association of users with multiple BSs is dynamically managed and updated at a centralized unit (Centralized Radio Access Network (C-RAN)). This approach removes the notion of cell borders and creates virtual cells that move according to the user mobility in the network.

The main contribution of this chapter is to propose a novel algorithm for automated managing of 5G virtual-cell networks [26], [27]. This algorithm is based on machine learning (ML) [62] and proactively predicts the optimal triggering conditions that can be used to dynamically enable/disable Coordinated Multipoint (CoMP) transmission [9] (or to form/dissolve VCs), instead of relying on passive approaches, as illustrated in Figure 4.1. The proposed algorithm uses

---

1 This chapter is an enhanced version of
Recurrent Neural Networks (RNNs), and in particular it uses Gated Recurrent Units (GRUs) [63]. Assuming that the sequences of Received Signal Strength (RSS) values of different mobile nodes in the network are available for training the GRU, the simulation results have shown that the future RSS values can be rapidly predicted with a very high accuracy based on the user mobility as shown in Section 3.4 of this dissertation. The predicted RSS values are used for making proactive decision on enabling/disabling the virtual-cell mode as the user moves in the network [26], [27]. One major challenge in order to make a reliable prediction is to collect a sufficient number of sequences in the target areas, so-called “big data”.

Figure 4.1. Adaptively enabling / disabling coordinated multipoint (CoMP) transmission according to the trajectory of a user in the network. (a) Mobile associating with the nearest base station (BS). (b) Forming a virtual cell (VC) using an enabling CoMP cooperation set. (c) Dissolving a VC.
4.2. CoMP Management Based on Recurrent Neural Network

A. System Model

A road network of dimensions 6 km × 6 km with a regularly spaced intersections of 1 km apart is generated, as shown in Figure 4.2. Within this grid, Eight BSs are placed in arbitrary locations. A mobile node is randomly generated on a road with an initial speed drawn from a uniform distribution between 8-12 km/h for pedestrians and between 55-65 km/h for vehicles. It is assumed that a mobile node continues to maintain a constant speed until leaving the road network. At any intersection a mobile node is assigned a probability of 0.5 for moving straight ahead, or an equal probability of 0.25 for turning either right or left. When a node is within the road network, it measures the RSS values from the closest three BSs. This method was chosen to demonstrate the capability of RNNs to learn sequences in order to predict RSS values to proactively enable/disable CoMP transmission. The path loss model adopted in the system model, is suggested by 3GPP [64], with an additional term to account for large- scale shadow fading. Here, σ is normally distributed with mean zero and variance 9 dB and \( d \) is measured in kilometers.

\[
PL(d) = 128.1 + 37.6 \log(d) + \sigma, \quad (4.1)
\]
Figure 4.2. A road network with eight base stations (BSs) (crosses) and a mobile node (circle). The grid lines are the roads. Each mobile node measures the received signal strength (RSS) from the three closest BSs [49].

B. Dataset

The data used is in the form of eight-element feature vector is created, where the positions correspond to BSs with their RSS values. The RSS values are recorded every 500 ms from the closest three BSs to the mobile node. The rest of the vector elements are set to zero. As the mobile node moves along the road network and the set of closest three BSs changes, the RSS values (feature vectors) are recorded and stored consecutively in a queue data structure. Table 4.1 presents a sample of the RSS data recorded for nine nodes (in rows) in the road network from the eight base stations (in columns).

Since the recorded RSS values from the distant past do not have a large predictive value regarding the direction the mobile node is heading to, the oldest elements are dropped from the queue when its size exceeds 150. With this method mobile nodes were simulated until a total of 100,000 sequences were collected.
Table 4.1: Sample of RSS values for different nodes

<table>
<thead>
<tr>
<th>BS1</th>
<th>BS2</th>
<th>BS3</th>
<th>BS4</th>
<th>BS5</th>
<th>BS6</th>
<th>BS7</th>
<th>BS8</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>-83.5</td>
<td>-107.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>0.0</td>
<td>-82.1</td>
<td>-60.7</td>
<td>0.0</td>
<td>-98.4</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>0.0</td>
<td>0.0</td>
<td>-95.9</td>
<td>-91.6</td>
<td>-82.5</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>0.0</td>
<td>0.0</td>
<td>-79.5</td>
<td>-82.5</td>
<td>-101.8</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>0.0</td>
<td>-73.5</td>
<td>-117.0</td>
<td>-80.5</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>0.0</td>
<td>-85.1</td>
<td>-90.7</td>
<td>-60.3</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>0.0</td>
<td>0.0</td>
<td>-85.9</td>
<td>-81.6</td>
<td>-92.5</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>-98.5</td>
<td>-112.1</td>
<td>-105.8</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

C. Recurrent Neural Network (RNN) Based on Gated Recurrent Units (GRU)

The RNN is used to discover patterns in sequential information (or temporal data). Unlike traditional neural networks where all inputs (and outputs) are assumed to be independent of each other, RNNs have a memory which captures information about what happened in all previous time steps to help learn a large range of dependencies. The GRU was first introduced by Cho et al. [63] as a modification to the hidden layer (H) RNN, to solve the vanishing gradient problem. The RNN architecture tends to give priority for the most recent inputs and neglects the effects of inputs that are further away in the past, as illustrated in Figure 4.3. A GRU is made up of two gates, as shown in Figure 4.4. The first is the update gate, which controls how much of the current cell content should be updated with the new candidate state. The second is the reset gate, which resets the memory of the cell if it is closed i.e. the unit acts as if the next processed input was the first in the sequence. The state equations of the GRU are [63].

reset gate: \( r[t] = \sigma(W_r h[t-1] + R_r x[t] + b_r) \), \( (4.2) \)

current state: \( h'[t] = h[t-1] \odot r[t] \), \( (4.3) \)

candidate state: \( z[t] = g(W_z h'[t-1] + R_z x[t] + b_z) \), \( (4.4) \)

update gate: \( u[t] = \sigma(W_u h[t-1] + R_u x[t] + b_u) \), \( (4.5) \)
new state: \( \mathbf{h}[t] = (1 - \mathbf{u}[t]) \odot \mathbf{h}[t - 1] + \mathbf{u}[t] \odot \mathbf{z}[t], \tag{4.6} \)

where, \( g(\cdot) \) is non-linear function usually implemented by a hyperbolic tangent, \( \sigma \) is the logistic sigmoid\(^2\), \( \mathbf{W}_r, \mathbf{W}_z, \mathbf{W}_u \) are rectangular weight matrices, that are applied to the input \( \mathbf{x}[t] \) (RSS values from all eight BSs), \( \mathbf{R}_r, \mathbf{R}_z, \mathbf{R}_u \) are square matrices that define the weights of the recurrent connections, \( \mathbf{b}_r, \mathbf{b}_z, \mathbf{b}_u \) are the bias vectors, and \( \odot \) is the Hadamard product.

After constructing the RNN-GRU model, the prediction error is evaluated using the Mean Square Error (MSE) loss function. The loss function, which measures the difference between the true and the predicted RSS values, is defined as follows.

\[
\text{MSE} = \mathcal{L}(\mathbf{Y}_i, \hat{\mathbf{Y}}_i) = \frac{1}{N} \sum_{i=1}^{N} \sum_{t=1}^{T} (Y_{i,t} - \hat{Y}_{i,t})^2 = \frac{1}{N} \sum_{i=1}^{N} \sum_{t=1}^{T} (Y_{i,t} - \mathbf{H}(\mathbf{X}_{i,t}))^2 , \tag{4.7}
\]

where \( N \) is the number of training examples, \( \mathbf{X}_{i,t} \) is the vector of observed RSS values from each of the 8 BSs, \( \mathbf{Y}_{i,t} \) is the vector of true RSS values and \( \hat{\mathbf{Y}}_{i,t} \) denotes the predicted RSS values (the output of the GRU-RNN model). The purpose of training the GRU-RNN model is to minimize the loss function by choosing a proper weighting matrix \( \mathbf{W} \). Thus, the optimization problem can be formulated as follows.

\[
\min_{\mathbf{W}} \frac{1}{N} \sum_{i=1}^{N} \sum_{t=1}^{T} (Y_{i,t} - \mathbf{H}(\mathbf{X}_{i,t}))^2 = \min_{\mathbf{W}} \frac{1}{N} \sum_{i=1}^{N} \sum_{t=1}^{T} (\mathbf{Y}_{i,t} - \mathbf{W}^{T}\mathbf{X}_{i,t})^2 , \tag{4.8}
\]

\(^2\) The logistic sigmoid is defined as \( \sigma = \frac{1}{1+e^{-x}} \)
Figure 4.3. (a) RNN structure. (b) RNN structure unfolded in time.

Figure 4.4. Gated recurrent unit (GRU) architecture. Dark gray circles with a solid line are the variables whose content is exchanged with the input and output of the network. Dark gray circles with a dashed line represent the internal state variables, whose content is exchanged within the cells of the hidden layer. White circles with +, 1 and represent linear operations.

4.3. Simulation Results

In the simulation, a Gated Recurrent Unit (GRU) mechanism with 512 units is used to predict the RSS values for making a proactive decision on whether to form/dissolve VCs via enabling/disabling CoMP transmission (predict the optimal triggering conditions for
enabling/disabling CoMP and forming VCs). A total of 70,000 sequences were used for training the GRU model and 30,000 were used for testing it. With 15 training epochs and 75 steps per epoch, the testing error gradually converged. The performance of the GRU model in predicting the RSS values of two different UE from the closet three BSs is shown in Figure 4.5. It was shown that the GRU model can output predictions on the RSS values that are very close to the true values most of the time. Based on these predictions, the network can proactively make a decision on enabling the CoMP transmission for the goal of providing a uniform user experience.

Figure 4.6 shows the Cumulative Distribution Function (CDF) of the number of enabled virtual cells when the GRU-RNN predictive model is applied. Note that the virtual-cell mode is enabled 14 times during the whole duration of time that nodes spend within the network with a probability approximately of 0.95, instead of relying on static virtual cell. Figure 4.7 depicts the gradual decrease of the loss function represented in Equation (4.7) as a function of the number of epochs. It shows that with 75 training epochs, the testing error based on MSE gradually converges.
Figure 4.5. True versus predicted received signal strength (RSS) values for two different UEs. (a) True and predicted RSS values measured from closest three BSs (3, 4 and 5). (b) True and predicted RSS values measured from closest three BSs (4, 5 and 6).
Figure 4.6. The CDF of enabling virtual cells using the GRU-RNN predictive model.

Figure 4.7. Convergence of the training model calculated based on MSE loss function.
4.4. Performance Analysis and Discussion

In this section, the performance of ML-based virtual-cell systems are evaluated in terms of the precision, recall, and accuracy, which are defined later. As illustrated in Section 4.3, for any node in the network, the future RSS values from each of the 8 BSs are predicted using the GRU-RNN model. The predicted RSS values are then used to make cooperative decisions among the BSs for enabling or disabling the virtual cell mode. The process of making proactive decisions about enabling or disabling VC mode becomes a binary classification problem and hence the precision and recall performance measures can be used for evaluating our model.

A. Precision

Precision is used to measure the exactness of a classifier. It is calculated by dividing the number of True Positives (TPs) by the sum of the True Positives (TPs) and False Positives (FPs). In our system, TP and FP represent respectively the number of true predictions about enabling the VC mode and the number of false predictions about enabling the VC mode, respectively, for each node in the network. Low precision of the classifier indicates that the classifier enables the VC mode when it’s not actually needed, which results in wasting network resources.

\[
\text{Total precision} = \sum_{k=1}^{M} \frac{TP}{TP + FP} \% = 86.68 \%
\]  
(4.9)

B. Recall

Recall is used to measure the competence of a classifier. It is calculated by dividing the number of True Positives (TPs) by the sum of True Positives (TPs) and the False Negatives (FNs). In our system the number of False Negatives (FNs) represents the number of false predictions about disabling VC mode. The low recall of the classifier indicates that the classifier disables the VC mode when it is actually needed, resulting in degradation of the quality of signal received by a node.
Total recall = \( \sum_{k=1}^{M} \frac{TP}{TP + FN} \% = 89.71 \% \)  \hspace{1cm} (4.10)

C. Accuracy

This performance measure is defined as the ratio of decisions made by the correct predictions of enabling /disabling the VC mode to the total number of decisions.

\[
\text{Accuracy} = \sum_{k=1}^{M} \frac{TP + TN}{TP + FP + FN + TN} \% = 92.80 \% \hspace{1cm} (4.11)
\]

The proposed machine learning algorithm addresses the problem of automated proactive mobility management in 5G virtual cell networks. This is a promising approach to achieving the goal of providing a uniform user quality of experience everywhere in the network via dynamic formation of VCs. Moreover, considering the very low-latency requirements of next-generation 5G networks, the proposed algorithm assists the mobility management function to rapidly create VC formation. As shown in Equation (4.11), an accuracy of 92.90 was achieved indicating that the proposed algorithm can detect patterns in arbitrary-length of RSS sequences to make accurate prediction on whether to enable or disable the virtual cell mode and also which BS the user should be associated with in the future.
CHAPTER 5: TOWARDS LOW LATENCY IN 5G HETNETS: A BAYESIAN CELL SELECTION / USER ASSOCIATION APPROACH}

5.1. Introduction

Among 5G’s headline features are expanding the cellular ecosystem to support an immense number of connected devices and creating a platform that accommodates a wide range of emerging services of different traffic types and Quality of Service (QoS) metrics. One of the key 5G performance metrics is ultra-low latency (~1ms) to enable new delay-sensitive use cases. Some network architectural changes have been proposed to facilitate the 5G ultra-low latency objective [19], [20], [65]. With these paradigm shifts in system architecture, it is of cardinal importance to rethink the cell selection / user association process to achieve substantial improvement in system performance over conventional maximum Signal-to-Interference plus Noise Ratio (Max-SINR) [50] and Cell Range Expansion (CRE) algorithms [51]. In this chapter we present a novel cell selection / user association approach that is directed towards 5G Heterogenous Networks (HetNet) to achieve the ambitious 5G low latency objective [28].

5.2. System Model

We consider a simplified two-tier Heterogeneous Network (HetNet) consisting of one High Power Node (HPN) overlaid by several Fog-Low Power Nodes (F-LPNs) with caching and computation capabilities as shown in Figure 5.1. The set of all radio Access Nodes (ANs) in the

---

1 This chapter is an enhanced version of
HetNet is defined as $N = \{n_0, n_1, n_2, \ldots, n_N\}$, where $n_0$ represents the HPN and the subset $L = \{n_1, n_2, n_3, \ldots, n_N\}$ denotes the F-LPNs. The F-LPNs are randomly distributed in the service area with spatial density of $\lambda$ (F-LPNs/km2). The set of all User Equipment (UE) under the coverage area of the two-tier HetNet is denoted by $U = \{u_0, u_1, u_2, \ldots, u_K\}$. Each UE $k \in U$ requests a service class defined as the tuple $\varphi_k = (\eta_k, \tau_k)$, where $\eta_k$, $\tau_k$ are the UE $k$ required data rate and latency respectively. Hence, the UE’s traffic in the proposed system model, can be mainly classified as delay-sensitive (DS) or delay-tolerant (DT) according to their latency requirement $\tau$.

The signal to interference plus noise ratio (SINR) at UE $k$, associated with access node $n \in N$ whose transmitting in downlink on the resource block $r$, is expressed as

$$
\gamma_{nk}^r = \frac{h_{nk}P_{nk}^r}{\sum_{n \in N_r/n} h_{nk}P_{nk}^r + WN_0^r},
$$

where $P_{nk}^r$ is the received power at UE $k$ from access node $n$ on resource block $r$, $R$ is the total number of resource blocks (RBs), $W$ is the bandwidth of each $R_B$, $N_0$ is the thermal noise spectral power, $\sum n \in N_r/n$ are the access nodes which are using the resource block $r$ and causing interference on the UE $k$ associated with the access node $n$, and $h_{nk}$ is the channel between access node $n$ and the UE $k$. The channel model incorporates the effects of small-scale fading and large-scale fading (the latter includes path loss and shadowing). The achievable downlink data rate at the UE $k$ from the AN (access node) $n$ on single resource block $r$ is given by

$$
R_{nk}^r = W \log_2(1 + \gamma_{nk}^r).
$$

(5.2)
Figure 5.1. Two-tier 5G HetNet consisting of a radio access controller (RAC), high-power node (HPN) and fog-low power nodes (F-LPNs).

Hence, the total achieved data rate from all the access nodes associated with UE $k$ can be written as

$$\mathbb{R}_k = \sum_{n \in \mathcal{N}} \alpha_k^n \sum_{r \in \mathcal{R}} \psi_{nk}^r R_{nk}^r,$$

(5.3)

where $\alpha_k^n \in \{0,1\}$ and $\psi_{nk}^r \in \{0,1\}$ are the user association and the resource allocation control variables respectively.

To meet the data rate requirement, at any time instant $t$, we assume that the UE $k$ can be associated with $N$ access nodes $n \in \mathcal{N}$, which is referred to as Coordinated Multipoint (CoMP) transmission [66].

$$\sum_{n \in \mathcal{N}} \alpha_k^n \leq M,$$

(5.4)

Thus, the AN $n$ serves the UE $k$ if the minimum data rate requirement in the service class tuple is GBR (Guaranteed Bit Rate), otherwise the UE will be severed by multiple ANs (CoMP transmission). Hence, we can write
The access nodes assign the needed resource blocks to UE $k$ to satisfy Equation (5.5). The number of resource blocks is given by

$$\Gamma = \sum_{n \in N} \sum_{r \in R} \psi_{nk}^r R_{nk}^r \geq \eta_k,$$

where $[\cdot]$ is the ceiling function. Similarly, to meet the primary delay requirement $\tau_k$ of UE $k$

$$\frac{1}{\Delta_{nk}} \leq \tau_k,$$

where $\Delta_{nk}$ is the inverse of the average Round-Trip delay-Time$^2$ (RTT) from access node $n$ to UE $k$. As mentioned previously, the F-LPNs are fitted with caching and computation capabilities that are brought close to the network edge. Therefore, the average RTT of the F-LPN is assumed to be much smaller than the average RTT of the HPN $\Delta_{nk}^{FLPN} > \Delta_{nk}^{HPN}$ [67].

5.3. Problem Formulation and Proposed Bayesian Game Approach

A. Problem Formulation

In this subsection, the problem of proper cell selection / user association is formulated. The utility functions are defined to represent the network’s resource utilization and the UE’s degree of satisfaction with the Quality of Service (QoS) including latency. The aim is to maximize the UE’s utility function for improved QoS satisfaction and the network’s resource utilization with respect to its preferences. Therefore, the UE and the network utility functions, are respectively written as

$$U_k^{UE}(\alpha, \psi) = \sum_{n \in N} \alpha_k^n \sum_{r \in R} \psi_{nk}^r R_{nk}^r - \theta \Delta_{nk},$$

---

$^2$The time required by a processor to serve a request is not included. It is assumed that with the advancement in computing powers and by aggregating multiple CPUs in a central unit, the processing latency can be neglected compared to the other latency components.
\[ U_{n}^{Net} = \omega_n(\theta_k) \left[ \frac{\Gamma_n}{\sum_{n \in N} \alpha_k^N n} \right], \tag{5.9} \]

where \( n \) is the total number of available resources for access node \( n \) and \( \omega_n(\theta_k) \) is representing the AN \( n \) preferences for UE \( k \) traffic type \( \theta_k \). The objective of the UE is to maximize (5.8) as

\[ \max_{\alpha, \psi} : U_k^{UE}(\alpha, \psi), \tag{5.10} \]

subject to: (5.4), (5.5), (5.6), (5.7).

Similarly, the objective of the network is to maximize (5.9) as

\[ \max_{\omega} : U_{n}^{Net}, \tag{5.11} \]

subject to: \( \alpha_k^N, \psi_k^R, \eta_k < C, \)

where \( C \) is the cap on the data rate allowed for UE \( k \) by access node \( n \). We modeled the cell selection/ user association problem as a Bayesian Game\(^3\) [68]. The motivation for selecting a Bayesian game to model for this problem is as follows. As previously mentioned, some of the proposed solutions such as the densification in deploying small cells and bringing the content closer to the user, are designed to accommodate the expected increase in data rates and to lower the end-to-end latency for certain type of applications. However, achieving the optimum system performance may not be feasible by just applying those solutions due to the \( a \) priori unawareness of access nodes to the UE’s type of traffic. Furthermore, assuming that the radio access nodes know the exact type of the user’s traffic is unrealistic since this take considerable time and makes achieving low latency quite unlikely. Hence, selecting a Bayesian game to model this problem,

\[^3\] The proposed cell selection / user association Bayesian Game algorithm is implemented in a central unit hypervisor referred to as the Radio Access Controller (RAC) which manages multiple access nodes.
where perfect knowledge about UE’s traffic types is not available at the access node, is well justified.

B. Proposed Bayesian Game Approach

The cell selection/user association Bayesian game is defined in the strategic form as $G = (\mathcal{P}, \theta, \mathcal{A}, \mathbb{P}, U)$, where:

- **Players ($\mathcal{P}$):** Set of two players $\mathcal{P} = \{u_k, n_n\}$: the UE $k$ and the AN $n$, respectively.
- **Types ($\theta$):** Set of possible types for a UE $k$ according to its traffic $\theta = \{\theta_{k,1} = DS, \theta_{k,2} = DT\}$, where $\theta_{k,j}$ is the type $j$ for the UE $k$.
- **Actions ($\mathcal{A}$):** The space of all possible combinational actions $\mathcal{A} = \mathcal{A}_1 \times \mathcal{A}_2$. Where, the UE’s action space $\mathcal{A}_1 = \{H = \text{UE selects } HPN, L = \text{UE selects } LPN\}$ and the AN’s action space $\mathcal{A}_2 = \{S = \text{serve}, C = \text{CoMP}\}$.
- **Prior Probabilities ($\mathbb{P}$):** The probability $\mathbb{P}(\theta_j)$ over the types of users.
- **Utility functions ($U$):** The UE and the network utility functions as defined in (5.8) and (5.9), respectively.

The player $i$’s strategy is mapping $s_i: \theta_i \rightarrow \mathcal{A}_i$, which represents the player $i$’s action for each possible type. In our Bayesian model, we started with a prior probability distribution over the UE types which is assumed to be common knowledge at the ANs. However, a given AN only knows its type and strategy, and does not know the strategies selected by the UE or its actual type.

The expected utility of a player $i$ under strategy profile $s$ is

$$E[U_i] = \sum_{\theta_j \in \theta} U_i(s(\theta_j), \theta_j) \mathbb{P}(\theta_j),$$

Hence, the expected utility (payoff) of UE $k$ is

58
\[
\mathbb{E}[U_k^{UE}] = p(\theta_j) \left[ \sum_{n \in N} \alpha_k^n \sum_{r \in R} \psi_{nk}^r R_{nk}^r - \theta \Delta_{nk} \right] 
+ \left(1 - p(\theta_j)\right) \left[ \sum_{n \in N} \alpha_k^n \sum_{r \in R} \psi_{nk}^r R_{nk}^r - \theta \Delta_{nk} \right].
\]

Similarly, the expected utility (payoff) of network is
\[
\mathbb{E}[U_n^{Net}] = \left(\omega \left[ \frac{\Gamma_n}{\sum_{n \in N} \alpha_k^n N_n} \right] \right) 
+ \left(1 - \omega \right) \left[ \frac{\Gamma_n}{\sum_{n \in N} \alpha_k^n N_n} \right].
\]

The UE \( k \) is of type \( \theta_{k,1} \), when \( U_k^{UE}(L,S) > U_k^{UE}(H,S) \) and \( U_k^{UE}(L,C) > U_k^{UE}(H,C) \). Consequently, the L strategy strictly dominates the H strategy. On the other hand, the UE is of type \( \theta_{k,2} \), when \( U_k^{UE}(H,S) > U_k^{UE}(L,S) \) and \( U_k^{UE}(H,C) > U_k^{UE}(L,C) \). Then, the H strategy strictly dominates the L strategy. Starting with an unbiased belief (\textit{a priori} probability) about the types of the UEs, then the expected utility (payoff) of the dominant strategies for both traffic types is assumed to be even (at the beginning the network is not biased in its choice for any type of users).

Hence, using Table 5.1, the posterior probability \( P \) can be calculated as in Equation (5.15).

\[
P = \frac{\omega_1 \left[ \frac{\Gamma_n}{\sum_{n \in N \setminus n_0} \alpha_k^n N_n} \right] - \omega_2 \left[ \frac{\Gamma_{n_0}}{\alpha_k^n N_{n_0}} \right]}{\omega_1 \left[ \frac{\Gamma_{N \setminus n_0}}{\alpha_k^n N_{n_0}} \right] - \omega_2 \left[ \frac{\Gamma_{n_0}}{\alpha_k^n N_{n_0}} \right] - \sum_{n \in N \setminus n_0} \frac{\Gamma_n}{\alpha_k^n N_n} + \sum_{n \in N} \frac{\Gamma_n}{\alpha_k^n N_n}}
\]
Table 5.1. Payoff matrix.

<table>
<thead>
<tr>
<th></th>
<th>Delay Sensitive (DS)</th>
<th>Delay Tolerant (DT)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( p )</td>
<td>( 1-p )</td>
</tr>
<tr>
<td><strong>UE Connect to</strong></td>
<td><strong>Network</strong></td>
<td><strong>Network</strong></td>
</tr>
<tr>
<td>( S )</td>
<td>( (U_k^{UE}(H,S), U_n^{Net}) )</td>
<td>( (U_k^{UE}(H,C), U_n^{Net}) )</td>
</tr>
<tr>
<td>( C )</td>
<td>( (U_k^{UE}(L,C), U_n^{Net}) )</td>
<td>( (U_k^{UE}(L,S), U_n^{Net}) )</td>
</tr>
</tbody>
</table>

### 5.4. Simulation Results and Analysis

In this section, we present the simulation setup and discuss the performance results of the proposed Bayesian cell selection / user association algorithm. The performance of the proposed Bayesian cell selection approach is evaluated in terms of the probability of proper association and the achieved latency with respect to conventional CRE and Max-SINR based cell selection / user association algorithms used in LTE-Advanced [50]. In the simulation, proper association is defined as the average number of delay-sensitive and delay-tolerant traffic associated with LPNS and HPN, respectively.

The simulation parameters are presented in Table 5.2. In the simulation setup, we assume that there is one HPN and F-LPNs that are randomly deployed over the service region. The UEs are randomly distributed following a homogeneous PPP and each UE could be one of the two types DS or DT. The DT and DS applications traffic demands are modeled as a uniform random variable on \([5,10]\) Mbps and \([0.1,4]\) Mbps, respectively. The path loss is based on TS 36.942 model [69], for an urban environment as
\[ PL = 40(1 - 4 \times 10^{-3}hb)\log_{10} d - 18\log_{10}(hb) + 21\log_{10}(f) \quad (5.16) \]

+ 80 dB,

where \(d\) is the separation between the UE and the AN, \(hb\) the height of the AN antenna and the \(f\) is the carrier frequency. The shadow fading is lognormal distributed, and the fast fading is based on the Winner II model [70]. In simulation, the average RTT for the F-LPN (1/\(\Delta F\)-LPN) is modeled as a uniform random variable on [0.5,1.5] ms and for HPN (1/\(\Delta HPN\)) as the sum of three uniform random variables, namely: radio access [0.5,1.5] ms, the core network [1,2] ms and [5,10] ms for the Internet.

### Table 5.2. Simulation parameters.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Service area</td>
<td>400 m × 400 m</td>
</tr>
<tr>
<td>Number of ANs (\mathcal{N} = \mathcal{N}_0 \cup \mathcal{L})</td>
<td>10 = 1+9</td>
</tr>
<tr>
<td>Number of UEs per AN</td>
<td>40</td>
</tr>
<tr>
<td>Total transmit power of ANs (\mathcal{N})</td>
<td>{46.02,40} dBm</td>
</tr>
<tr>
<td>Transmit antenna height of ANs (hb)</td>
<td>20 m</td>
</tr>
<tr>
<td>Carrier frequency (f)</td>
<td>2.14 GHz</td>
</tr>
<tr>
<td>Bandwidth</td>
<td>20 MHz</td>
</tr>
</tbody>
</table>

After performing simulations, we used the CCDF of achieved latencies of the delay-sensitive traffic to evaluate the Bayesian game in comparison with the Max-SINR approach. As shown in Figure 5.2, the minimum achieved latency using the SINR approach is \(\approx 6.8\) ms due to the improper association that happens when delay-sensitive traffic is associated with HPN using the max-SINR criterion, while significantly the Bayesian approach achieves latencies < 1 ms.
In addition to latency, we study the proper association of different schemes. Where the proper association refers to when the delay-sensitive traffic is associated with Fog-Low Power Nodes (F-LPNs) and the delay-tolerant traffic is associated with High Power Node (HPN). Figure 5.3 illustrates the cumulative distribution functions (CDFs) of proper associations for the proposed Bayesian game compared with the CRE and Max-SINR approaches. It is noted from Figure 5.3 that the minimum percentage of proper associations for the delay-sensitive traffic of the Bayesian approach is greater than or equal to 82%, which outperforms the proper association of the conventional Max- SINR and the CRE approaches. Similarly, using the CDF of the delay-tolerant traffic as illustrated in Figure 5.4, we observed that the CRE and Max-SINR approaches are inferior to the Bayesian game approach. The Bayesian game approach attains a proper association greater than or equal to 80% for delay-tolerant traffic.
In this chapter, a novel method based on Bayesian games for cell selection/ user association in 5G heterogeneous networks was presented. Such a methodology can be quite important in achieving the 5G low latency objective of 1ms. As illustrated in Section 5.3, simulation results demonstrate that the proposed Bayesian game algorithm provides a significant improvement in
terms of the probability of proper association (as a function of traffic type) and the achieved latency compared with the Maximum Signal-to-Interference-plus-Noise Ratio (Max-SINR) criterion and Cell Range Expansion (CRE) approaches, that do not use delay as a metric. Simulation results show that Bayesian game approach attains the 5G low end-to-end latency target with a probability exceeding 80%.
CHAPTER 6: CONCLUSIONS

The main contributions in this dissertation can be summarized as below:

6.1. **Slotted ALOHA Non-Orthogonal Multiple Access (SAN) Protocol for IoT Networks**

This work is directed towards a novel Medium Access Control (MAC) layer protocol that is easy to implement, energy efficient, scalable and compatible with the low complexity requirements of IoT devices. The synergistic combination of the Slotted ALOHA (SA) protocol with NOMA and a Successive Interference Cancellation (SIC) receiver was demonstrated to significantly improve the throughput performance with respect to the SA protocol.

This methodology enables the use of the SAN protocol by providing the IoT gateway a means to determine the number of active IoT devices in the medium. Knowing the number of active IoT devices is essential in order to optimize the SIC power levels and the ability to distinguish between signals from different IoT devices transmitting at the same time and on the same frequency. In the case of massive number of active IoT devices the power levels become high, whereas the IoT devices usually have a limited power capability. Therefore, with this technology the number of active devices that can achieve a successful transmission using SIC is limited to keep the power levels modest. It was shown that with correct detection, there is a substantial throughput gain of 5.5 dB relative to SA for 0.07 probability of transmission and up to 3 active transmitters.

Next, as a further enhancement to SAN protocol the merit of Multiple-Input Multiple-Output Beamforming (MIMO-BF) was studied as a mechanism for reducing collisions, It was
shown that the addition of beamforming to the SAN protocol, dubbed BF-SAN, demonstrates a significant improvement in throughput for a large number of IoT devices, compared with conventional Slotted ALOHA (SA) and SAN protocols, through the exploitation of power and spatial domains without excessively increasing the SIC power levels which is impractical when using power-limited IoT devices.

Also, BF-SAN addresses the SAN channel access delay problem by reducing the probability of collision and consequently lowering the average back-off delay when beamforming is employed. Simulation results have shown that BF-SAN can achieve higher throughput than the SAN protocol with fewer attempts and consequently a lower channel access delay. Such a random-access protocol can be quite important for massive M2M communication in IoT networks.

6.2. **Recurrent Neural Networks for Proactive Mobility Management in 5G Virtual-Cell Networks**

A deep machine learning (ML) algorithm for proactive mobility management in 5G wireless networks was proposed and evaluated in a particular scenario. Using ML to proactively predicting the triggering conditions to enable/disable Coordinated Multipoint (CoMP) transmission and Virtual-Cell (VC) selection was investigated. A Recurrent Neural Network (RNN) based on a Gated Recurrent Unit (GRU) mechanism was applied to eight element feature vectors of the Received Signal Strength (RSS) values from all the eight BSs in the network to make an accurate prediction about the RSS values in the future. Simulation results have shown that the proposed GRU-RNN model achieves an accuracy of 92.80 % to predict the triggering conditions for enabling and disabling the Virtual-Cell (VC) mode as required based on the mobility of users. As mentioned earlier, using the Virtual-Cell (VC) mode can help attaining a uniform user experience for mobile users by improving the throughput of cell-edge users and minimizing the
number of hard handovers. However, to enable the formation of VCs, for instance in downlink (DL), the User-Equipment (UE) data must always be available to the BSs in the CoMP cooperating set, hence, this leads to a substantial cooperation cost due to the exchange of data and the control information among the cooperating BSs. On the other hand, waiting for CoMP transmission to be enabled can lead to a degradation in throughput since the time required to form VCs is severely time constrained. Simulation results have shown that the virtual-cell mode was enabled 14 times during the duration of time that nodes spend within the network, with a probability approximately of 0.95, instead of relying on static virtual cells.

The proposed algorithm is a promising approach to achieving the goal of providing a uniform user experience everywhere in the network via dynamic formation of VCs. Moreover, considering the very low-latency requirements of next-generation 5G networks, the proposed algorithm helps the mobility management function facilitate VC formations.

6.3. **Bayesian Game Approach for Cell Selection / User Association in 5G HetNet**

As discussed in CHAPTER 1 and CHAPTER 5, the three solutions proposed for 5G to meet the low latency and high data rate requirements of 5G, namely: Dense LPNs, F-RANs and CoMP, are not sufficient in themselves to meet the diverse requirements for a wide range of emerging applications and in particular the low-latency target. For this reason, it is essential to rethink the traditional cell selection/user association procedures, by considering both the traffic type and the access node capabilities. A novel method for cell selection/user association for 5G heterogeneous networks was proposed using a Bayesian game model. The utility functions of the user equipment and the network are defined based on the achievable data rate, the average Round-Trip-Time (RTT), and the access node’s traffic preferences. Simulation results demonstrate that the proposed Bayesian game algorithm provides a significant improvement in terms of the
probability of proper association (as a function of traffic type) and the achieved latency compared with Max-SINR criterion and Cell Range Expansion (CRE) approach. It was shown that the Bayesian game approach attains a proper association greater than or equal to 80% of the time for delay-tolerant traffic. Such a methodology can be quite useful for achieving the 5G low latency objective.
REFERENCES


[31] “Institute of Electrical and Electronics Engineers (IEEE): The Internet of Things (IoT),” Special Report, Mar. 2014.


[34] 3GPP, “Study on RAN improvements for machine-type communications,” TS 37.868 V11.0, October 2016.


[58] 3GPP TD RP-150496: “Study on Downlink Multiuser Superposition Transmission”.


APPENDIX A: COPYRIGHT PERMISSIONS

The permission below is for the use of material in Chapter 3.

**Title:** Enabling slotted Aloha-NOMA for massive M2M communication in IoT networks

**Conference Proceedings:** 2018 IEEE 19th Wireless and Microwave Technology Conference (WAMICON)

**Author:** Mohamed Elkourdi

**Publisher:** IEEE

**Date:** April 2018

Copyright © 2018, IEEE

### Thesis / Dissertation Reuse

The IEEE does not require individuals working on a thesis to obtain a formal reuse license, however, you may print out this statement to be used as a permission grant:

Requirements to be followed when using any portion (e.g., figure, graph, table, or textual material) of an IEEE copyrighted paper in a thesis:

1) In the case of textual material (e.g., using short quotes or referring to the work within these papers) users must give full credit to the original source (author, paper, publication) followed by the IEEE copyright line © 2011 IEEE.
2) In the case of illustrations or tabular material, we require that the copyright line © [Year of original publication] IEEE appear prominently with each reprinted figure and/or table.
3) If a substantial portion of the original paper is to be used, and if you are not the senior author, also obtain the senior author's approval.

Requirements to be followed when using an entire IEEE copyrighted paper in a thesis:

1) The following IEEE copyright/ credit notice should be placed prominently in the references: © [year of original publication] IEEE. Reprinted, with permission, from [author names, paper title, IEEE publication title, and month/year of publication]
2) Only the accepted version of an IEEE copyrighted paper can be used when posting the paper or your thesis online.
3) In placing the thesis on the author’s university website, please display the following message in a prominent place on the website: In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of [university/educational entity's name goes here]'s products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to http://www.ieee.org/publications_standards/publications/rights/rights_link.html to learn how to obtain a License from RightsLink.

If applicable, University Microfilms and/or ProQuest Library, or the Archives of Canada may supply single copies of the dissertation.
Title: Slotted Aloha-NOMA with MIMO Beamforming for Massive M2M Communication in IoT Networks

Conference Proceedings: IEEE 88th Vehicular Technology Conference (VTC2018-Fall)

Author: Mohamed Elkourdi

Publisher: IEEE

Date: August 2018

Copyright © 2018, IEEE

Thesis / Dissertation Reuse

The IEEE does not require individuals working on a thesis to obtain a formal reuse license, however, you may print out this statement to be used as a permission grant:

Requirements to be followed when using any portion (e.g., figure, graph, table, or textual material) of an IEEE copyrighted paper in a thesis:

1) In the case of textual material (e.g., using short quotes or referring to the work within these papers) users must give full credit to the original source (author, paper, publication) followed by the IEEE copyright line © 2011 IEEE.

2) In the case of illustrations or tabular material, we require that the copyright line © [Year of original publication] IEEE appear prominently with each reprinted figure and/or table.

3) If a substantial portion of the original paper is to be used, and if you are not the senior author, also obtain the senior author's approval.

Requirements to be followed when using an entire IEEE copyrighted paper in a thesis:

1) The following IEEE copyright/credit notice should be placed prominently in the references: © [year of original publication] IEEE. Reprinted, with permission, from [author names, paper title, IEEE publication title, and month/year of publication]

2) Only the accepted version of an IEEE copyrighted paper can be used when posting the paper or your thesis on-line.

3) In placing the thesis on the author's university website, please display the following message in a prominent place on the website: In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of [university/educational entity's name goes here]'s products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to http://www.ieee.org/publications_standards/publications/rights/rights_link.html to learn how to obtain a License from RightsLink.

If applicable, University Microfilms and/or ProQuest Library, or the Archives of Canada may supply single copies of the dissertation.

BACK CLOSE WINDOW

Copyright © 2019 Copyright Clearance Center, Inc. All Rights Reserved. Privacy statement. Terms and Conditions.
Comments? We would like to hear from you. E-mail us at customercare@copyright.com
The permission below is for the use of material in Chapter 4.

Mohamed Elkoudi
Copyright Permission
To: wire Mobil

Hi,

I am sending this email to request a formal copyright permission to use my paper titled “Optimization of 5G Virtual Cell Based Coordinated Multipoint Networks Using Deep Machine Learning” in my PhD dissertation.

Please try sending me the copyright permission ASAP so I can append it in my PhD dissertation.

Thanks.

wire Mobil
Re: Copyright Permission
To: Mohamed Elkoudi

Hi,

Since you are the author of the article, you are free to use it in your PhD dissertation. No worries.

Cheers!

**************************************************************************************************************
Editorial Secretary
International Journal of Wireless & Mobile Networks (IJWMN)- ERA Indexed
http://acmce.org/journal/ijwmn.html
The permission below is for the use of material in Chapter 5.

**Thesis / Dissertation Reuse**

The IEEE does not require individuals working on a thesis to obtain a formal reuse license, however, you may print out this statement to be used as a permission grant:

Requirements to be followed when using any portion (e.g., figure, graph, table, or textual material) of an IEEE copyrighted paper in a thesis:

1) In the case of textual material (e.g., using short quotes or referring to the work within these papers) users must give full credit to the original source (author, paper, publication) followed by the IEEE copyright line © 2011 IEEE.
2) In the case of illustrations or tabular material, we require that the copyright line © [Year of original publication] IEEE appear prominently with each reprinted figure and/or table.
3) If a substantial portion of the original paper is to be used, and if you are not the senior author, also obtain the senior author's approval.

Requirements to be followed when using an entire IEEE copyrighted paper in a thesis:

1) The following IEEE copyright/credit notice should be placed prominently in the references: © [year of original publication] IEEE. Reprinted, with permission, from [author names, paper title, IEEE publication title, and month/year of publication]
2) Only the accepted version of an IEEE copyrighted paper can be used when posting the paper or your thesis on-line.
3) In placing the thesis on the author’s university website, please display the following message in a prominent place on the website: In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of [university/educational entity's name goes here]'s products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to http://www.ieee.org/publications_standards/publications/rights/rights_link.html to learn how to obtain a License from RightsLink.

If applicable, University Microfilms and/or ProQuest Library, or the Archives of Canada may supply single copies of the dissertation.