

2022-09

# Improved handover decision scheme for 5g mm-wave communication: optimum base station selection using machine learning approach.

Mollel, Michael

NM-AIST

---

<https://dspace.nm-aist.ac.tz/handle/20.500.12479/1653>

*Provided with love from The Nelson Mandela African Institution of Science and Technology*

**IMPROVED HANDOVER DECISION SCHEME FOR 5G MM-WAVE  
COMMUNICATION: OPTIMUM BASE STATION SELECTION USING  
MACHINE LEARNING APPROACH.**

**Michael S. Mollel**

**A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor  
of Philosophy in Information and Communication Science and Engineering of the Nelson  
Mandela African Institution of Science and Technology,**

**Arusha, Tanzania.**

**September, 2022**

## ABSTRACT

The rapid growth in mobile and wireless devices has led to an exponential demand for data traffic and exacerbated the burden on conventional wireless networks. Fifth generation (5G) and beyond networks are expected to not only accommodate this growth in data demand but also provide additional services beyond the capability of existing wireless networks, while maintaining a high quality-of-experience (QoE) for users. The need for several orders of magnitude increase in system capacity has necessitated the use of millimetre wave (mm-wave) frequencies as well as the proliferation of low-power small cells overlaying the existing macro-cell layer. These approaches offer a potential increase in throughput in magnitudes of several gigabits per second and a reduction in transmission latency, but they also present new challenges. For example, mm-wave frequencies have higher propagation losses and a limited coverage area, thereby escalating mobility challenges such as more frequent handovers (HOs). In addition, the advent of low-power small cells with smaller footprints also causes signal fluctuations across the network, resulting in repeated HOs (ping-pong) from one small cell (SC) to another.

Therefore, efficient HO management is very critical in future cellular networks since frequent HOs pose multiple threats to the quality-of-service (QoS), such as a reduction in the system throughput as well as service interruptions, which results in a poor QoE for the user. However, HO management is a significant challenge in 5G networks due to the use of mm-wave frequencies which have much smaller footprints. To address these challenges, this work investigates the HO performance of 5G mm-wave networks and proposes a novel method for achieving seamless user mobility in dense networks. The proposed model is based on a double deep reinforcement learning (DDRL) algorithm. To test the performance of the model, a comparative study was made between the proposed approach and benchmark solutions, including a benchmark developed as part of this thesis. The evaluation metrics considered include system throughput, execution time, ping-pong, and the scalability of the solutions. The results reveal that the developed DDRL-based solution vastly outperforms not only conventional methods but also other machine-learning-based benchmark techniques.

The main contribution of this thesis is to provide an intelligent framework for mobility management in the connected state (i.e HO management) in 5G. Though primarily developed for mm-wave links between UEs and BSs in ultra-dense heterogeneous networks (UDHNs), the proposed framework can also be applied to sub-6 GHz frequencies.

## DECLARATION

I, **Michael S. Mollel** do hereby declare to the Senate of Nelson Mandela African Institution of Science and Technology that this thesis is my original work and to the best of my knowledge has not been submitted or presented to any other institution for similar or different award.

Michael S. Mollel




10.05.2021

**Candidate Name and Signature**

**Date**

The above declaration is confirmed

Dr. Michael Kisangiri



10.05.2021

**Name and Signature of Supervisor 1**

**Date**

Dr. Shubi Kaijage

10.05.2021

**Name and Signature of Supervisor 2**

**Date**

## **COPYRIGHT**

This thesis is copyright material protected under the Berne Convention, the Copyright Act of 1999 and other international and national enactments, in that behalf, on intellectual property. It may not be reproduced by any means, in full or in part, except for short extracts in fair dealings; for research or private study, critical scholarly review or discourse with an acknowledgment, without a written permission of the Deputy Vice Chancellor for Academic, Research and Innovation, on behalf of both the author and the Nelson Mandela African Institution of Science and Technology.

## CERTIFICATION

The undersigned certify that they have read and hereby recommend for acceptance by the Nelson Mandela African Institution of Science and Technology a research report entitled: ***Improved Handover Decision Scheme for 5G Mm-wave Communication: Optimum Base Station Selection using Machine learning Approach.***, in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Information and Communication Science and Engineering of the Nelson Mandela African Institution of Science and Technology.

Dr Michael Kisangiri



---

**Name and Signature of Supervisor 1**

10.05.2021

---

**Date**

Dr. Shubi Kaijage

---

**Name and Signature of Supervisor 2**

10.05.2021

---

**Date**

## ACKNOWLEDGMENTS

I acknowledge with delight the assistance received from many people and appreciate all those who assisted me during my studies at the Nelson Mandela African Institution of Science and Technology (NM-AIST). To begin with, I must express my sincere gratitude to Dr. Michael Kisangiri, Dr. Shubi Kaijage and Dr. Qammer H. Abbasi. They gave me the liberty to manage this research, enabled me to recognize research problems, offered valuable guidance and instilled enthusiasm for solving problems into me. They have also enabled me to develop the mindset, and acquire the relevant skills and experience needed to carry out quality engineering research, which I did not have prior to the commencement of my PhD studies. I also wish to thank Dr. Ahmed Zoha, Dr. Sajjad Hussain, and Professor Muhammad Ali Imran for accepting me into the Communication, Sensing and Imaging (CSI) group during my tenure at the University of Glasgow. They enabled me to know more about my research problem and future direction.

My thanks also go out to colleagues and friends in the CSI group and at the NM-AIST. I wish to thank Metin Ozturk, Attai Abubakar, Kenechi Omeke, Aysenur Turkmen, Wasiwasi Mgonzo and Oluwakayode Onireti for their valuable suggestions, discussions and contributions to my articles. They also influenced my thinking and gave me motivation when I needed it most.

I would like to express my profound gratitude to African Development Bank (AfDB) for sponsoring my studies and the Prevention and Combating of Corruption Bureau (PCCB) for granting me the study leave to accomplish this study. My sincere gratitude goes to Mr Julius Lenguyana, Ms Victoria Ndossi, Mr Japhet Laizer, Mr Humphrey Robert and the NM-AIST project team for their spontaneous outpouring of support and cooperation.

Special thanks to my beautiful wife Amelia for her love, patience, and support throughout the period of my research until the completion of my thesis. To my parents Samwel, and Elizabeth, who made ensure that I had an excellent education, I am extremely grateful.

Finally, my uttermost gratitude goes to almighty God, who sustained my life, gave me understanding, vitality, and strength while on my path to the completion of this thesis.

Since it is not possible to mention everybody who has contributed to this work, allow me to express my heartfelt appreciation to everybody who stretched my mind or hand towards this study's accomplishment.

## **DEDICATION**

This thesis work is dedicated to my lovely wife, Amelia, who has been a constant source of support and encouragement during graduate school and life challenges. I am deeply grateful for having you in my life. This work is also dedicated to my children. You have made me stronger, better and more fulfilled than I could have ever imagined. Lastly, this work is dedicated to my parents and siblings, who have always loved me unconditionally and whose good examples have taught me to work hard for the things I aspire to achieve.

## TABLE OF CONTENTS

ABSTRACT . . . . .	i
DECLARATION . . . . .	ii
COPYRIGHT . . . . .	iii
CERTIFICATION . . . . .	iv
ACKNOWLEDGEMENTS . . . . .	v
DEDICATION . . . . .	vi
TABLE OF CONTENTS . . . . .	vii
LIST OF TABLES . . . . .	xi
LIST OF FIGURES . . . . .	xii
LIST OF APPENDICES . . . . .	xiv
LIST OF ABBREVIATIONS AND SYMBOLS . . . . .	xvii
CHAPTER ONE . . . . .	1
INTRODUCTION . . . . .	1
1.1 Background of the Problem . . . . .	1
1.1.1 Channel Characteristics of 5G Wireless Systems . . . . .	4
1.1.2 Heterogeneous Networks . . . . .	7
1.1.3 Internet of Things . . . . .	9
1.1.4 Device-to-Device Communication . . . . .	11
1.1.5 Vehicular Communications . . . . .	12
1.1.6 High Speed Train Communication . . . . .	13
1.1.7 Beyond 5G System . . . . .	14

1.2	Problem Statement . . . . .	14
1.3	Rationale of the study . . . . .	15
1.4	Research Objectives . . . . .	16
1.4.1	General Objective . . . . .	16
1.4.2	Specific Objectives . . . . .	16
1.5	Significance of the Study . . . . .	16
1.6	Delineation of this Study . . . . .	17
CHAPTER TWO . . . . .		19
LITERATURE REVIEW . . . . .		19
2.1	5G System Architecture . . . . .	19
2.2	Mobility Management in 5G . . . . .	22
2.2.1	Radio Resource Control State Machine . . . . .	23
2.2.2	Idle and Inactive State Mobility . . . . .	27
2.2.3	Connected State Mobility . . . . .	29
2.3	Handover Management in 5G and Beyond . . . . .	30
2.3.1	Types of Handover . . . . .	30
2.3.2	Handover Requirements and Key Performance Indicators . . . . .	34
2.3.3	Handover and Radio Resource Management . . . . .	34
2.3.4	Dual Connectivity . . . . .	35
2.3.5	Handover Management in NR . . . . .	35
2.3.6	Mobility and Handover Management in B5G . . . . .	39
2.4	Machine Learning for Handover management . . . . .	41
2.4.1	An Overview of Machine Learning Algorithms . . . . .	43
2.4.2	Machine Learning based Handover Optimization . . . . .	47
CHAPTER THREE . . . . .		57

MATERIAL AND METHODS . . . . .	57
3.1 System Models Implementation . . . . .	57
3.1.1 5G Heterogeneous Network . . . . .	57
3.1.2 Channel Model . . . . .	58
3.1.3 Beamforming Model . . . . .	59
3.1.4 SINR Model . . . . .	59
3.1.5 Radio Link Failure Model . . . . .	59
3.1.6 User and Traffic Models . . . . .	60
3.2 Handover as Combinatorial Optimization Problems . . . . .	61
3.2.1 Conditions for Initiating Handover based on 3GPP . . . . .	62
3.2.2 Handover Cost . . . . .	63
3.2.3 Trajectory and Service-aware Handover . . . . .	64
3.3 Benchmarks Solutions . . . . .	65
3.3.1 Selection of Target HO BS in Sparse 5G Networks based on Clustering of UDNs . . . . .	65
3.4 Proposed Solutions . . . . .	68
3.4.1 Proposed Solution for Handover Management using RL . . . . .	69
3.4.2 DRL-based optimal BS selection . . . . .	74
3.4.3 DDRL-based optimal BS selection . . . . .	77
CHAPTER FOUR . . . . .	80
RESULTS AND DISCUSSION . . . . .	80
4.1 Performance Evaluation . . . . .	80
4.1.1 Simulation Setup . . . . .	80
4.1.2 Performance Metrics . . . . .	81
4.1.3 Comparative Analysis . . . . .	82

CHAPTER FIVE . . . . .	92
SUMMARY AND CONCLUSION . . . . .	92
5.1 Summary . . . . .	92
5.1.1 Intelligent Handover Scheme . . . . .	92
5.1.2 Benefits of the Developed Intelligent Handover Scheme . . . . .	92
5.2 Conclusion . . . . .	93
REFERENCES . . . . .	95
APPENDICES . . . . .	121
RESEARCH OUTPUTS . . . . .	128

## LIST OF TABLES

Table 1:	Requirements of 5G and B5G —key performance indicators. . . . .	2
Table 2:	Enabler to enhance capacity gain for cellular networks—5G and B5G ap- proaches. . . . .	4
Table 3:	Cell types in wireless networks in term of coverage and capacity . . . . .	9
Table 4:	Summary of the RRC State and mobility handling in 5G. . . . .	24
Table 5:	Types of Machine Learning Algorithm. . . . .	45
Table 6:	Summary of the State-of-the-art ML-based HO Optimization in 5G mm- wave Communication Systems. . . . .	56
Table 7:	Radio link failure parameters. . . . .	60
Table 8:	Traffic model parameters. . . . .	61
Table 9:	Parameters for designing and developing DQN model . . . . .	76
Table 10:	Environment simulation parameters . . . . .	81

## LIST OF FIGURES

Figure 1:	An illustration of heterogeneous networks. . . . .	8
Figure 2:	Sensors and IoT use case. . . . .	10
Figure 3:	The 3GPP-5G architecture with reference points (3GPP, 2019). . . . .	20
Figure 4:	Overall architecture of 5G system showing network elements and inter-faces (3GPP, 2020). . . . .	21
Figure 5:	UE state machine and state transitions in 5G (3GPP, 2018b). . . . .	23
Figure 6:	RAN Areas and Tracking Areas. . . . .	28
Figure 7:	An illustration depicting intra-frequency HO in scenario 1 and inter-frequency HO in scenario 2. . . . .	31
Figure 8:	UE undergo HO from once cell to another with both cells use the same RAT (intra-RAT). . . . .	32
Figure 9:	Inter-RAT HO scenarios in distributed and centralized RAN architectures. .	33
Figure 10:	Dual connectivity with HO scenarios in future communication networks. .	36
Figure 11:	UE performs intra-gNB HO which involves the change of cells in the same gNB. . . . .	37
Figure 12:	UE performs inter-gNB HO, which involves the change of gNBs with same UPF and AMF for scenario 1 and change of UPF for scenario 2. . . . .	38
Figure 13:	UE performs inter-gNB HO with AMF change, involving the change of gNBs while UPF is maintained in scenario 1 and change of UPF in scenario 2.	39
Figure 14:	HO procedure in 5G-NR involving no change of AMF and UPF (3GPP, 2020). . . . .	40
Figure 15:	The system model of mm-wave UDN. . . . .	57
Figure 16:	HO problem in the overlapping BSs coverage area . . . . .	65
Figure 17:	Overview of generic RL algorithm . . . . .	69
Figure 18:	DRL-based framework comprising environment, states, actions, and rewards.	74
Figure 19:	The structure of the proposed DDRL - with double Q- networks . . . . .	77
Figure 20:	Number of HOs and $\gamma_{ave}$ as functions of $t_d$ , for $\gamma_{th} = 20$ dB and $\lambda = 10$ BSKm <sup>-2</sup>	83
Figure 21:	Number of HOs as a function of $\gamma_{th}$ for $t_d = 0.7, 1, 2, 3$ sec and $\lambda = 10$ BSKm <sup>-2</sup>	84
Figure 22:	Number of HOs as a function of $\gamma_{th}$ for $t_d = 0.7, 1, 2, 3$ sec and $\lambda = 50$ BSKm <sup>-2</sup>	85
Figure 23:	Relationship between the number of HOs and UE velocity . . . . .	86

Figure 24:	HO performance showing the relationship between average system throughput and UE velocity . . . . .	87
Figure 25:	Average running time as a function of number of mm-wave BSs . . . . .	88
Figure 26:	Ping Pong rate as a function of UE velocity . . . . .	89
Figure 27:	Frequency of HOs as a function of mm-wave BS density . . . . .	90
Figure 28:	Average system throughput as a function of mm-wave BS density . . . . .	91

## LIST OF APPENDICES

Appendix 1: Python codes for System Model . . . . .	121
Appendix 2: Python codes for Agent $Q$ -Neural Network . . . . .	123
Appendix 3: Python codes for Main function . . . . .	124

## LIST OF ABBREVIATIONS AND SYMBOLS

### SYMBOLS

$[\dots]$	iverson brackets, $[P] \doteq 1$ if P is true, else 0
$\alpha$	learning rate parameter
$\arg \max_x f(x)$	a value of $x$ at which $f(x)$ takes its maximal value
$\beta_c$	handover cost
$\gamma$	signal to Noise ratio
$\mathbb{1}_{\text{HO}}$	indicator function ( $\mathbb{1}_{\text{HO}} \doteq 1$ if the HO occur is true, else 0)
$\nabla \hat{Q}(s, a, \theta)$	column vector of partial derivatives of $\hat{Q}(s, a, \theta)$ with respect to $\theta$
$\zeta$	discount factor for weighting future rewards

### ABBREVIATIONS

$k$ NN	$k$ -nearest neighbour
3GPP	3 <sup>rd</sup> Generation Partnership Project
4G	Fourth Generation
5G	Fifth Generation
6G	Sixth Generation
A2C	asynchronous actor-critic
AMF	Access Mobility Function
ANN	artificial neural networks
AS	Access stratum
B5G	Beyond 5G
BS	Base Station
CMAB	contextual Multi-Armed Bandit
CN	Core network
CNN	convolution neural networks
CSI	channel state information
D2D	Device-to-device
DBN	deep belief networks
DDPG	deep deterministic policy gradient
DDRL	Double Deep Reinforcement Learning
DRL	Deep Reinforcement Learning

E-UTRA	Evolved Universal Mobile Telecommunications System Terrestrial Radio Access
gNB	Next-Generation NodeB
HetNet	Heterogeneous Networks
HO	Handover
HST	High speed train
IoT	Internet of Things
IP	internet protocol
ISS	Intra-Cluster Sum of Squares
KPI	Key Performance Indicator
LIDAR	Light Detection and Ranging
LOS	Line of Sight
LTE	Long Term Evolution
MAB	Multi-Armed Bandit
MARL	Multi-Agent Reinforcement Learning
MIMO	Multiple-Input Multiple-Output
ML	Machine learning
mm-wave	Millimetre Wave
MME	Mobility management entity
mMTC	Massive Machine-Type Communications
MR	Measurement report
NF	Network Function
NG	Next-Generation
NG-C	Next-Generation control-plane part
NG-RAN	NextGen Radio Access Network
NG-U	Next-Generation user-plane part
NR	New Radio
NSSF	Network Slice Selection Function
PCF	Policy control function
PPP	Poisson Point Process
QoE	Quality-of-Experience
QoS	Quality-of-Service

RAI	RAN area identifier
RAN	Radio Access Network
RAT	radio access technology
RBH	Rate-based HO policy
ReLU	Rectified Linear Unit
RIAI	RRC_Inactive assistant information
RL	Reinforcement Learning
RLF	Radio Link Failure
RNA	RAN-based notification area
RNAU	RAN-based Notification Area Update
RNN	recurrent neural networks
RRC	Radio resource control
RSRP	Reference Signal Received Power
S-BS	serving BS
SC	Small Cell
SGW	Serving Gateway
SHP	Smart HO Policy
SMF	Session Management Function
SNR	signal to noise ratio
SVM	support vector machine
T-BS	target BS
TAI	Tracking area identifier
THz	TeraHertz
TTT	Time-to-trigger
UAVs	unmanned aerial vehicles
UDHN	Ultra-Dense Heterogeneous Network
UDM	Unified data management
UDN	Ultra-dense network
UE	User equipment
UPF	User Plane Function
URLLC	Ultra-Reliable Low-Latency Communications
V2X	Vehicular-to-everything

## CHAPTER ONE

### INTRODUCTION

#### 1.1 Background of the Problem

Wireless communication networks have been witnessing an unprecedented demand in terms of bandwidth and number of connections in this so-called *information age*—in particular the age of big data<sup>1</sup> where data is regarded as *new oil* (Zhang & Letaief, 2020). It is reported in the Ericsson Mobility Report that the mobile network traffic soared by 56% in the first quarter of 2020 (Ericsson, 2020), indicating the imminent issue that needs to be addressed. There are strong evidence for the correlation between such growth in the global data traffic and the proliferation of emerging applications, including tactile-internet, virtual reality, high-definition video streaming. For example, we learnt from the same report (Ericsson, 2020) that video streaming alone constitutes more than half of the mobile data traffic, and there is a tendency towards higher resolutions—putting the issue at an alarming level in terms of data demand. This, in turn, poses serious challenges to legacy networks and paves the way for the fifth generation of cellular networks (5G), which offers a thousandfold increase in capacity (Agiwal et al., 2016; Busari et al., 2018; Tayyab et al., 2019). As such, enhanced mobile broadband has been included in 5G New Radio (5G-NR) as one of the scenarios—along with ultra-reliable low-latency communications (URLLC) and massive machine-type communications (mMTC)—in order to support the aforementioned bandwidth-hungry applications (3GPP, 2017a). Table 1 summarizes expectation and key performance indicators (KPI) for 5G and beyond (B5G) systems.

On the other hand, Internet of things (IoT) devices have already pervaded our daily life, as they can be seen in numerous domains, including agriculture (Lashari et al., 2018), healthcare (Islam et al., 2015), smart living (Pal et al., 2019; Meng et al., 2018; Mocrii et al., 2018), smart home, smart industry, and smart city (Zanella et al., 2014), to name a few. In the case of smart city, for example, city waste, building health monitoring, traffic, etc., are managed smartly using IoT technology by deploying the IoT devices to the required places accordingly (Zanella et al., 2014). A good example of this can be found in the publication by the Mayor of London on the

---

<sup>1</sup>Steve Lohr, *The Age of Big Data*, The New York Times, 11 Feb. 2012. Available online at <https://www.nytimes.com/2012/02/12/sunday-review/big-datas-impact-in-the-world.html>. Accessed on 25 Oct. 2020.

Table 1: Requirements of 5G and B5G —key performance indicators.

Performance Indicator	5G	B5G
Downlink Peak Rate	20 Gbps	1 Tbps
Uplink (UL) Peak Rate	10 Gbps	1 Tbps
Traffic Capacity	10 Mbps/m <sup>2</sup>	10 Gbps/m <sup>3</sup>
Latency	1 msec	0.1 msec
Energy Efficiency	Not set	1 pJ/bit

road map for smart city agenda<sup>2</sup> with the slogan “*Smarter London Together*”, which dictates a heavy use of IoT technology in London to make the City more efficient and to boost the standard of living of its residents. IoT technology owes this popularity to the promises in terms of making our everyday life as well as industrial processes more manageable and efficient with continuous monitoring and quick response (Xu et al., 2014; Lee et al., 2017; Akpakwu et al., 2018). The alarming point here is that IoT devices are becoming more pervasive each year and are projected to gain more dramatic prevalence in the near future, albeit a slight deceleration due to COVID-19 pandemic (Ericsson, 2020).

The challenges of 5G and B5G cellular communication networks, therefore, are primarily twofold: a) the bandwidth demand due to more advanced smartphones with more computational capabilities, and the rise in data demanding applications, such as online gaming, augmented reality, etc. (Öztürk, 2020); b) the number of cellular connections that is exponentially growing mainly due to IoT technology. Various solutions have already been proposed in order to combat these issues: network densification and millimetre wave (mm-wave) communications are among the most important candidates for network capacity enhancement (Shafi et al., 2017). Network densification is a phenomenon, whereby the base station (BS) density in a given environment is increased in order to provide more radio access network (RAN) capacity. This concept mainly uses the idea of frequency reuse, which states that the frequency spectrum of one BS can be reused by other BSs as well only if they avoid interfering with each other. This avoidance is provided by lowering the transmit power in order to reduce the footprints of BSs, so that the overlapping regions are minimized—the less footprint of BSs results in more BSs

<sup>2</sup>The road map can be found at the following link. Accessed on 22/11/2020. Available online at [https://www.london.gov.uk/sites/default/files/smarter\\_london\\_together\\_v1.66\\_-\\_published.pdf](https://www.london.gov.uk/sites/default/files/smarter_london_together_v1.66_-_published.pdf).

deployment opportunity, which then leads to more RAN capacity. mm-wave, on the other hand, offers a great enhancement in the RAN capacity of cellular networks by exploiting the abundant bandwidth available in the mm-wave frequency spectrum. Moreover, as antenna sizes reduce with increasing carrier frequency, the use of mm-wave communication enables Multiple-Input Multiple-Output (MIMO) technology, which in turn enhances the reliability and capacity of the network (Jameel et al., 2017). In other words, the capacity enhancement supplied by mm-wave communications are mainly due to two factors: a) increased bandwidth made available, and b) MIMO technology (Rappaport et al., 2017; Jameel et al., 2017).

Even though these are sensible and effective methods of enhancing the capacity of cellular networks, a serious side effect immediately emerges: mobility management (Öztürk, 2020). The common ground for network densification and mm-wave communication concepts is that both lead to more frequent handovers (HOs), which is defined as the user equipment's (UE's) change of channel, resource, or cell <sup>3</sup> association while keeping an ongoing call or session. The underlying reasoning behind this consequence is mainly due the reduction of the footprint of BSs. First, in the case of network densification, the footprint is deliberately reduced with the use of small cells (SCs) in order to facilitate more BS deployments through frequency reuse. Second, concerning the mm-wave communications, the footprint of BSs reduces due to the higher propagation losses incurred at mm-wave frequencies (more dependency on line of sight (LOS)). Furthermore, the increased amount of bandwidth also shortens the range of mm-wave signals (Björnson et al., 2017).

As such, the frequency of HOs grows due to the smaller footprints of BSs: mobile UEs would need to perform more HOs, given that there are now more BSs in a certain environment. Given that the average throughput of a user is a function of the number of HOs with an inverse proportionality (Arshad et al., 2016a), this issue has severe consequences in terms of communication quality—degrades the quality-of-service (QoS). Besides, as service interruptions are experienced during HOs, the user satisfaction rates are also affected negatively, undermining the great promises of 5G networks. These adverse effects are mainly caused by two reasons: a) the number of HO experienced during a call or data transfer session; and b) the HO cost incurred for each HO experienced. In this regards, the research activities on HO management have predominantly focused on these two aspects, such that minimizing the number of HOs and/or the cost incurred

---

<sup>3</sup>Cell and BS are interchangeably used throughout this work unless stated otherwise.

per HO.

Although the figures in terms of the growing number of IoT devices and BSs along with increasing demand for data-oriented applications have been discussed negatively so far, there are some positive impacts as well. The volume of data being generated by cellular networks is also growing considerably, making it a gold mine for network operators to exploit in such a way that more efficient management can be facilitated (Sun et al., 2019b; Morocho Cayamcela & Lim, 2018; Zhang et al., 2019; Öztürk, 2020; Mollel et al., 021b). In other words, although growing network sizes results in more complexity, the immense data volume generation becomes a key to alleviate such complexity: this so-called challenge brings its own opportunity and solution. In that regard, machine learning (ML) techniques have gained significant attention in the field of wireless communications, since such amount of data can be very well utilized for training ML models, which could help the networks gain experience and take proactive and more informed actions.

With other introduced use-cases, the 5G aims to provide high system capacity as discussed early, and Table. 2 summarizes approaches used to increase network capacity to accommodate UEs and achieve the 5G requirement. Among different techniques, only those associated with HO are discussed in the following subsection.

Table 2: Enabler to enhance capacity gain for cellular networks—5G and B5G approaches.

Approach	Capacity Gain	Reference
Frequency Division	5	(Balakrishnan, 2015)
Higher-order Modulation	5	(Balakrishnan, 2015)
IoT and Device-to-Device Communication	$\geq 17$	(Mustafa et al., 2015)
Access to higher frequency band	$\geq 25$	(Mollel et al., 020b)
Frequency reuse using more SCs	1600	(Huq et al., 2019a)

### 1.1.1 Channel Characteristics of 5G Wireless Systems

Table. 2 shows the increase of system capacity from using the high-frequency band and other approaches. The new spectrum introduced in 5G is mm-wave, while for B5G is anticipated to use the TeraHertz (THz) band (Chataut & Akl, 2020; Huq et al., 2019a). Nevertheless, the

lower band — sub-1 and 6 GigaHertz (GHz) — continue to exist in the 5G and are expected to exist in B5G for specific purposes. Compared to the sub-6 GHz band, the mm-wave band's advantages include more available bandwidth and small antennas in devices. Antenna size is inversely proportional to frequency; therefore, mm-wave antennas for UE and BS are small and can be placed in small devices. However, the mm-wave band has some drawbacks that necessitate the use of sub-6 GHz frequencies in 5G. In this subsection, we present the rationale for the co-existence of multiband frequencies in 5G and the characteristics and applications of different spectrum bands from sub-1 GHz to mm-wave. We also emphasize the adverse effect of each spectrum on HO in each subsection.

#### **(i) Sub-1 GHz and Sub-6 GHz in 5G**

In its early phases of implementation, 5G's main spectrum options were around 3.5 GHz and 4.5 GHz for sub-6 GHz with time division duplexing technology. For the 3.5/4.5 GHz band, 5G aims to use existing BSs to help in the roll-out and implementation (Schumacher et al., 2019). The 3.5 GHz band provides comparatively less coverage than the 2 GHz band used in legacy networks, and this is because radio propagation decreases as frequency increases. However, introducing MIMO beam-forming antennas at 3.5 GHz and higher spectrum reduces propagation losses, thereby significantly increasing coverage for 3.5/4.5 GHz. The effect of propagation loss can be reduced by designing MIMO beam-forming antennas with a good receiver sensitivity, and high antenna gains.

The sub-1 GHz bands are also used through frequency division duplex in 5G, especially for deep indoor penetration (Nokia, 2017). With its broader coverage, low data rate IoT connectivity and other critical communication like remote control or automotive applications can be introduced. Therefore, extensive coverage becomes imperative for these new use cases which can be served by the sub-1 GHz band. This frequency band is assigned at 700 MHz and potentially 900 MHz, if communications service providers reduce the spectrum allocation for legacy radios in order to make room for 5G (Nokia, 2017; Chen et al., 2018a; Yue et al., 2019). Generally, a sub-6 GHz band has less impact on HO than using a higher frequency band due to high coverage, making even heuristic solutions capable of solving the HO problem in the sub-6 GHz channel.

## **(ii) mm-wave in 5G**

The propagation of waves at mm-wave is more prone to adverse effects of obstacles which can be caused by movement of people, presence of trees, foliage (outdoor scenario), furniture and walls (indoor scenario). Since the mm-wave spectrum is severely affected by rain and other atmospheric conditions, previous studies suggested that it was impractical to use this frequency band for mobile communications. However, this has been proven to be wrong, as recent studies have shown that atmospheric absorption does not create a significant loss when used in picocells—coverage below 200 m from transmitter— (Attiah et al., 2020; Rangan et al., 2014; Mollel et al., 2019). These studies also show that even under very extreme rainfall, the rain attenuation would cause 1.4 dB and 2 dB loss at 28 GHz and 73 GHz, respectively. The impact of rain attenuation on mm-wave propagation, especially in urban picocell areas, will therefore become insignificant (Rangan et al., 2014). The short-range coverage of mm-wave has both advantages and disadvantages. Spatial reuse of frequency band, strong multi-path behaviour due to reflection are among the advantages of using mm-wave while one of the disadvantages of using this band is that many SCs are required to provide coverage due to the high propagation loss of mm-wave.

mm-wave is an inherently directional wave which means that there is a need for the transmitter and receiver to focus the beam towards each other, this is commonly known as beam steering. The main advantage of beam steering is to achieve high gain by focusing the transmitter and receiver towards each other. The beam steering is completed through a beam training/tracking process. Beam training is a process of finding the desired beam to connect the UEs in order to reduce initial access delay. The beam training protocols (i.e., algorithms which carry out beam tracking) developed for both the BS and UEs have to be run very frequently and fast due to outage events and UE's mobility. Another critical parameter to consider is sensitivity to blockage. mm-wave has a higher frequency, making the size of its wavelength small compared to many physical objects, and thus the low ability of mm-wave to diffract through large objects makes it sensitive to blockage. For example, at the 60 GHz band, it is observed that there is a 20-35 dB increase in the path-loss if an obstacle (e.g., humans or furniture) is introduced between the mm-wave link (Rangan et al., 2014). Various studies have shown that human activities—including UE rotation, human movement, etc.— contribute to the channel blockage. Due to unavoidable activities such as signal blockage resulting from holding the UEs in the hand, the

network reliability would become a challenging issue in NextGen mm-wave systems (Rangan et al., 2014).

While it has been demonstrated that using mm-wave frequencies such as 28 GHz and 38 GHz is possible even in complex urban environments, many challenges such as low throughput and high signaling overheads associated with HO still needs to be addressed to realize the full potentials of the mm-wave band (Rangan et al., 2014; Arshad et al., 2016b).

### **(iii) Co-existence of Sub-1 GHz, Sub-6 GHz and mm-wave**

Given the rigid transmission efficiency standard for certain use cases such as vehicular networks, the use of mm-wave poses some significant difficulties in implementing reliable but high data rate communication. Critical IoT applications, including remote healthcare systems (for clinical remote monitoring and assisted living), traffic and industrial control (drone/robot/vehicle), and tactile Internet, etc., require higher availability, higher reliability, safety, and lower latency to ensure end-user experience as failure to satisfy these requirements would result in severe consequences, such as vehicle collision, and accident (Akpakwu et al., 2018).

A control plane and user plane (UP) decoupled network is designed to circumvent these challenges by using the sub-6 GHz for the control plane and mm-wave frequencies for the UP (Yan et al., 2019). This guarantees that signaling from the control plane reaches the UE with high reliability by using the sub-6 GHz spectrum. On the other hand, the use of mm-wave frequencies for UP provides unprecedented data speeds, due to the vast bandwidth availability at the mm-wave spectrum. Therefore, the main purpose of sub-6 GHz and sub-1 GHz bands is to provide uninterrupted access to the control plane or to provide coverage for areas where mm-wave cannot offer adequate coverage.

## **1.1.2 Heterogeneous Networks**

Spatial reuse of more cell sites provides more orders of magnitude in terms of capacity gain; the technique is among the various approaches showcased in Table. 2. This result in smaller low-powered cellular layers that proliferated on top of the macrocell layer. These small, energy-efficient cells include microcells, picocells, femtocells that can relay, among others, and Fig. 1 shows a scenario with macrocells overlaid with different small cell types. A heterogeneous network (HetNet) is such a network that consists of many overlapping cellular layers, each with

its own set of specific characteristics. HetNet supports aggressive spectrum spatial reuse and enables dynamic traffic offloading from macrocell for several purposes, including network load balancing, capacity boost or coverage extension. A summary of the types of cells in terms of coverage and capacity is presented in Table 3. A macrocell is a BS used in cellular networks

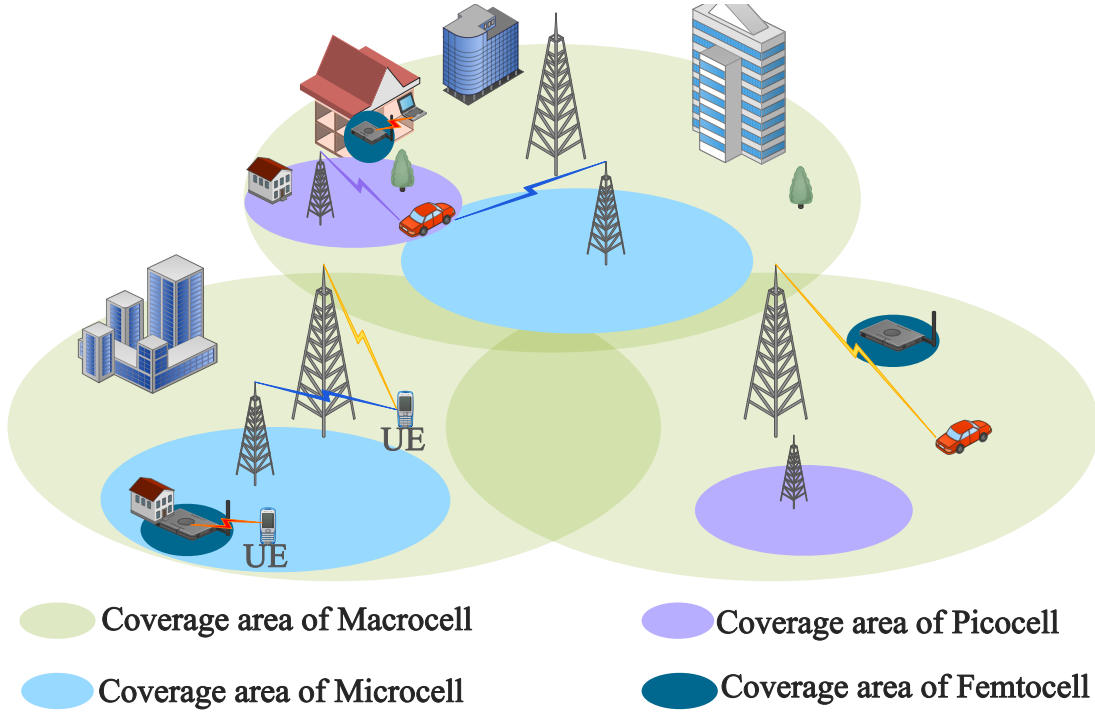


Figure 1: An illustration of heterogeneous networks.

with the function of providing radio coverage to a large area of mobile network access for long time. Currently, the macrocell overlaps several SCs, and it has high output power, usually in the range of tens of watts and can provide coverage to a large area. However, the macrocell suffers from interference caused by the use of sub-6 GHz, which can travel far by nature. While the macrocell transmits radio waves over a long distance, if not managed properly, signal interference with other cells is very likely, which in turn could result in the degradation of network performance (Ali et al., 2017). Nevertheless, macrocell has low spectral efficiency or area spectral efficiency, typically measured in (bit/s/Hz) per unit area, which results in less bandwidth and low data rate per UE. The data rate is the function of bandwidth, and SCs allows frequency reuse due to limited range hence more bandwidth and data rate per UE. Therefore, to increase the data rate, the idea of reducing BS footprint for macrocells was introduced (Hamed & Rao, 2018), led to more SCs.

As the BS footprint becomes smaller with smaller BSs, the use of mm-wave become more feasi-

Table 3: Cell types in wireless networks in term of coverage and capacity

Cell type	Coverage range(meter)	Capacity
Femtocell	10-20	A few UEs
Picocell	200	20 - 40 UEs
Microcell	2000	> 100 UEs
Macrocell	$(3-3.5) \times 10^4$	Many UEs

ble. The mm-wave frequency suffers from high penetration loss, enabling mm-wave frequency in an indoor environment for femtocell. However, the drawback is the high number of HO for the mm-wave link and the high chances of sub-optimally selecting BS as the serving BS (S-BS) in ultra-dense network (UDN) if HO is required.

### 1.1.3 Internet of Things

In this modern era, various applications used by billions of people are daily made available via the internet, thereby making the Internet an essential tool to interconnect these applications, among which services like video streaming, file sharing, electronic commerce, etc. are increasingly taking place online. The types of interconnected devices includes smart phones and IoT devices such as sensors, wearables, etc. These IoT devices are able to communicate with each other to share information with little or no human involvement. Fig 2 illustrates some common IoT uses cases. As the number of IoT devices keep increasing, the traffic generated by these devices also increases, hence, the underlying protocols that support IoT should be reconsidered to support the massive interconnection of both new and conventional devices (Yilmaz et al., 2016; Srinidhi et al., 2019). Conventional devices need to be made smarter by incorporating advanced technologies such as ubiquitous computing, artificial intelligence, embedded devices, different communication standards and technologies, various application services, and different Internet standards. However, the problem is that these devices used in IoT are memory-limited and energy-limited, so information should be routed efficiently, and the proper channel between source and sink should be carefully chosen (Srinidhi et al., 2019). IoT has different use cases such as smart cities, smart home, vehicular sensors, health monitoring, and sport & leisure scenarios. Several of these use cases are discussed in the following subsection, while focusing on the differences in application domains requirements.

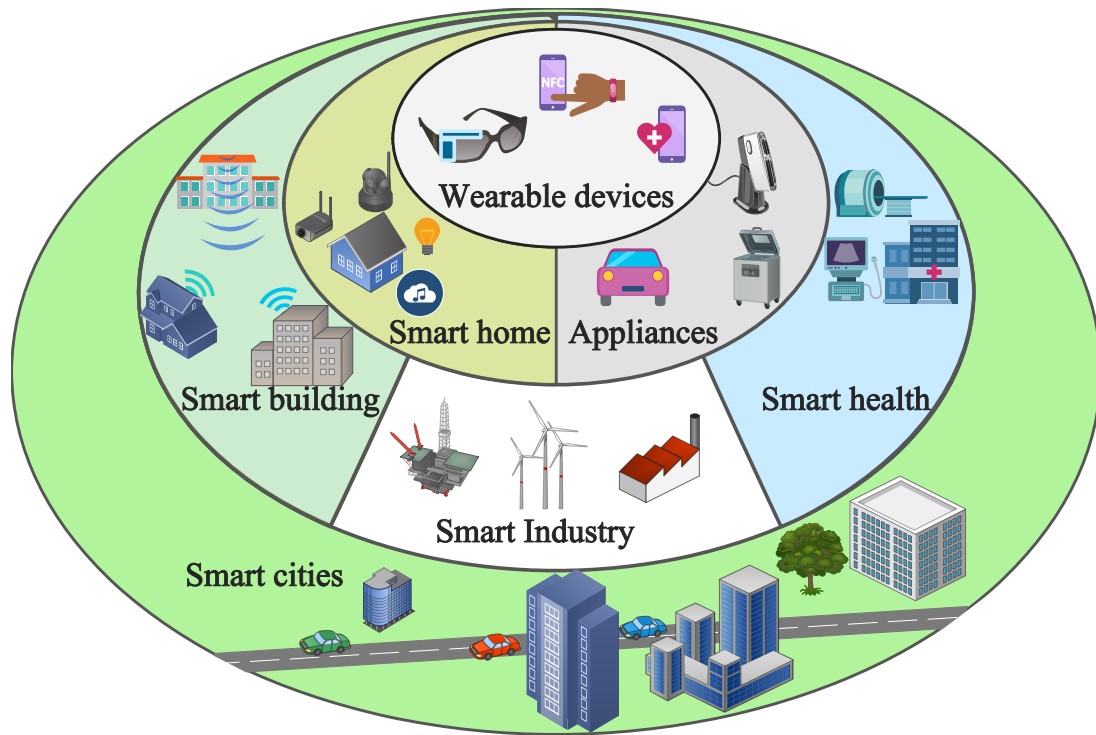


Figure 2: Sensors and IoT use case.

#### (i) Smart Cities

This involves the use of smart technologies to provide relevant information and automated services that would improve the standard of living of the people in a particular area. These smart technologies include: deployment of sensors for traffic monitoring and management, smart grid, waste management systems etc (Akpakwu et al., 2018). Some of these use cases are briefly discussed: a) *vehicular traffic monitoring* where sensors are deployed on the roads for the detection of traffic jams, polluted or damaged roads, as well dynamically proposing rerouting options for end-users who have GPS equipment in their vehicles and are able to receive such information, b) *Street lights* can be equipped with sensors for detecting cars or human movement, and can be dynamically turned on and off according to activity level around the area-of-interest. This can assist in energy and monetary saving for the city, whilst ensuring security by preventing unilluminated zones around people, and c) there could also be *environmental sensors* deployed in various locations to detect pollution, water level, or fire. In this case, the early detection of abnormal environmental situations can be used to alert the appropriate authorities in order to enable them take the necessary actions when any incident occurs.

## **(ii) Smart Home**

This use case is sometimes classified as a part of smart cities. However, it is mostly limited to user-oriented applications, particularly for home networks (Akpakwu et al., 2018). Different services that can be classified under the smart home use-case include: a) *Connected home appliances*, such as smart fridges which can automatically order for the restocking of the fridge with food items or beverages when it detects that it is running out of supplies by checking a pre-defined threshold for the amount of each item, and b) *home video monitoring* — homes can be equipped with small cameras that are mounted in different locations, and can be used to stream the video to the Internet for a remote monitoring. The sensors can also be used to send alarms upon the detection of unusual movement or abnormal behavior, smoke, carbon monoxide, etc. in the monitored area.

## **(iii) Healthcare/Telemedicine/Wearable**

This use case is becoming more popular as more devices such as watches and other wearable devices become increasingly available. Patients do not necessarily need to be monitored manually, but smart wearable devices track their health conditions for any abnormality. Such devices send an alarm message to a nearby hospital as soon as they detect anomalies with the patients being monitored.

All the use cases mentioned above face challenges that need to be addressed before IoT can become very efficient and able to integrate heterogeneous devices—device with different communication standards (protocols, technologies and hardware)—and applications envisaged for 5G (Uwaechia & Mahyuddin, 2020). These challenges include scalability, network management, security and privacy, interoperability and heterogeneity, network congestion and overload, and network mobility and coverage. To interconnect a massive number of devices and accommodate enormous traffic generated within 5G system, conventional sub-6 GHz is no longer sufficient, hence the need for the utilization of a new frequency band (mm-wave) (Yilmaz et al., 2016). This would lead enhanced QoS for IoT devices.

### **1.1.4 Device-to-Device Communication**

Device-to-device (D2D) communication involves the direct communication between two devices without passing through a BS. These devices could be smartphones, vehicles, etc.

This kind of communication usually occurs when both devices are in close proximity to each other (Astely et al., 2013). The introduction of D2D communication is necessary to cope with the rise in the number of devices as well as the increase in demand for high speed connections. It is one of the technologies that is being exploited in 5G and B5G networks as its use would lead to enhanced link reliability, spectral efficiency, system capacity, energy efficiency and reduced network delays. It is also useful for offloading traffic from the core of the network (Asadi et al., 2014).

The use of mm-wave in 5G would facilitate D2D communications as more direct links would be supported, thereby enhancing the capacity of the network. In addition, due to the directional nature of 5G antennas, it would be possible to support more simultaneous connections in mm-wave systems. Both local and global D2D connections are supported by mm-wave 5G networks. The latter occurs between two devices within the coverage of the same BS while the former occurs between devices under the coverage of different BSs (Gandotra & Jha, 2016). Despite the inherent advantages of D2D communications, due to UE mobility, and the fact that the UEs still need to connect to the BS in order to transmit control signals, the issue of HO needs to be carefully considered in order to prevent ping-pong effects which results in frequent HOs (Lai et al., 2020).

### **1.1.5 Vehicular Communications**

Vehicular-to-everything (V2X) is a special case of D2D communication. It is a technology that provides communication between vehicles and surrounding devices, including hand-held devices, moving/stationary cars, and all other external IoT appliances. V2X is categorized into two main components: vehicle-to-vehicle and vehicle-to-infrastructure. The former allows communication between two or more vehicles. On the other hand, the latter deals with communication between cars and other devices in its external environment, such as traffic/street lights. A particular case of vehicle-to-infrastructure called vehicular ad-hoc network or network on wheels, is used to provide communication between vehicular nodes (Abbasi & Shahid Khan, 2018).

The most common and popular communication protocol that supports vehicular networks is the dedicated short-range communication, and it can support all V2X architecture. Dedicated short-range communication uses 75 MHz bandwidth at 5.9 GHz band, and was expected to provide

the data rate up to 27 Mb/s and a transmission range up to 1000 m (Abbasi & Shahid Khan, 2018). However, in practice, the data-rate reached only about 2–6 Mb/s. To increase the data rate and extend the coverage, the coexistence of dedicated short-range communication and long term evolution (LTE) has been suggested as a solution for V2X application (Ghafoor et al., 2020).

As a result of the high mobility of vehicles, one of the major challenges that vehicular networks suffer from is HO. This occurs because in the course of the vehicles movement from one location to another, they often move out of the coverage area of one network also known as road side unit to another thereby leading to frequent change of connection from one road side unit to another. This issue would become more pronounced with the use of mm-wave in 5G as the coverage area of the road side units would become smaller (Ahmed & Alzahrani, 2019). Hence, HO management must be carefully considered for fast moving vehicles in 5G mm-wave communication networks in order to ensure seamless HO.

#### **1.1.6 High Speed Train Communication**

High speed train (HST) communications is one of the verticals that would be supported by 5G networks. The availability of large spectrum in the mm-wave frequency band would make the provision of enhanced mobile broadband services possible for passengers in high speed trains. HST communications is needed to improve passenger safety and also enhance their on-board experience. The services provided by HST communications includes closed-circuit television surveillance, uninterrupted Internet, high definition video streaming, e-ticketing, communication between trains, etc. (Chen et al., 2018a).

HST communication networks basically encompasses two kinds of communications, namely: critical and non-critical communications. The former is the communication between the HST and its associated infrastructures and is necessary to control the speed and ensure the safety, reliability and smooth functioning of the HST. The latter is required to provide services to the passengers on-board such as high quality video, and other data services (Briso-Rodríguez et al., 2017). Even though mm-wave has great potentials for application in HST communications, due to the high mobility of trains, HST communication is often prone to frequent HOs and fast fading channels, that potentially undermines its availability. As a result, some new technologies such as hybrid beamforming, beam management, network slicing, and distributed antenna

system have been introduced in mm-wave communications to enhance its application in high mobility applications such as HST communications (Yue et al., 2019).

### **1.1.7 Beyond 5G System**

5G is game-changer as it can provide data rates up to tens of gigabits per second, which is far beyond what is provided by legacy networks (Liu & Jiang, 2016). However, with the introduction of new use cases and applications such as virtual and augmented reality, remote surgery and holographic projection, 5G would not be able to meet the projected explosion in wireless data demands. As a result, research into higher frequency (beyond mm-wave) has risen, and THz frequency has become the B5G researcher's focus as the new spectrum for B5G systems. Only frequency bands in the THz range can provide the large amount of bandwidth that is needed to support the terabit-per-second data rates in order to support huge traffic types such as uncompressed videos that is envisioned in B5G networks (Mumtaz et al., 2017; Huq et al., 2019b).

The use of THz band in sixth generation (6G) is required to provide the reliable communication that is required to support various critical applications, accommodate high data rates per area, and support massive amounts of connected UEs. The THz frequency band has quite similar characteristics to that of mm-wave. However, because it has a higher frequency compared to mm-wave, this means that it would be prone to all the challenges facing mm-wave along side additional challenges because it is a higher spectrum compared to mm-wave. Therefore, there is a need for more advanced error control mechanisms, mobility management techniques, as well as other new features to enable the utilization of this frequency band in mobile cellular networks.

## **1.2 Problem Statement**

A significant challenge in wireless 5G systems is achieving smooth and robust mobility, particularly in HetNets with BSs and UEs using mm-wave links. A comprehensive mobility performance analysis for HetNets shows that the rate of HO failure, ping pong effects and other unnecessary HO is much higher in mm-wave than in sub-6 GHz links (Tayyab et al., 2019). A proper HO management scheme plays a vital role in the successful realization of the potential gains of using mm-wave in HetNets because poorly managed HOs leads to more HO signalling

which results in significant increase in latency during data transmission. In addition it is also associated with high cost of recovery from HO failures. The HO process generally involves the time delay from HO initiation to completion, resulting in high signalling overhead, latency, and poor service quality.

5G aims at achieving ultra-low latency (possible  $< 1$  ms) (Mollet et al., 2021b) among several other objectives. Thus, the unavailability of services during path switching must be reduced to a minimum level. Since HO failures and unnecessary HO would be more prominent in HetNet mm-wave networks, more robust frameworks for selecting target BSs (T-BS) that can minimize HO and prolong UE to BS connection are required to mitigate the impact of HO in the system.

In this study, ML algorithms were employed to develop models that improve the performance of HO management, achieve seamless mobility, and high UE throughput. In particular, reinforcement learning (RL) was used to model the trade-off between selecting the BS with the best instantaneous metric such as signal to noise ratio (SNR) and the one that can sustain UE connectivity for an extended period without initiating a new HO. The proposed model considered various features found in the wireless network environment, such as static and dynamic blockages, devices with different service demands, etc.

### **1.3 Rationale of the study**

Strategies for improving HO decision in ultra-dense heterogeneous networks have been proposed in various studies such as (Mezzavilla et al., 2016; Arshad et al., 2016b; Arshad et al., 2016; Mahira & Subhedar, 2017). However, these studies ignored some vital parameters that are required for HO optimization in 5G mm-wave communication systems. Such parameters include the speed of mobile UEs, UEs distribution, the presence of blockages, etc. Various techniques have been proposed for HO management from heuristics Zang et al. (2017); Demarchou et al. (2018) to ML-based ((Wang et al., 2018; Alkhateeb et al., 2018; Sun et al., 2018)) approaches. Heuristic methods lack intelligence while deciding the choice of the best beam in a cell or selecting the optimal T-BS among the available potential T-BSs while ML-based methods employ intelligence in deciding the optimal BS which results in more accurate selection of a T-BS. However, most of the techniques proposed in the literature are either impractical or have inefficient implementation in UDN-HetNets

Therefore, this study aims to develop an intelligent approach based on deep RL (DRL) for selecting the optimal BSs that would maximize the UE-BS connection duration to reduce HO cost while guaranteeing user QoS. The developed solutions consider information regarding user trajectory, network topology and network parameters to improve the accuracy of the model and make it practically applicable to wireless environments. Such information includes: a) HO cost along the UE trajectory, b) blocking probability due to environmental factors, c) a user's QoS, and d) probability of UE staying in a BS footprint (dwell time). HO is a critical limiting factor that should be carefully considered when planning network densification in 5G. Therefore, the optimal model for minimizing the HO rate is a crucial consideration for prior and post HO optimization schemes.

## **1.4 Research Objectives**

### **1.4.1 General Objective**

To study the performance of various HO decision schemes for 5G and B5G networks that use mm-wave links and propose a novel approach that uses ML to achieve seamless mobility for users in an UDN environment.

### **1.4.2 Specific Objectives**

- (i) To perform a comprehensive review of existing HO decision methods that are used in 5G and B5G mm-wave communication systems.
- (ii) To set up experiments and carry out signal measurements in a simulated 5G mm-wave environment for mobile users.
- (iii) To develop a model that can intelligently optimize HO rate by training agents to maximize throughput by maintaining connection to a BS for longer durations.

## **1.5 Significance of the Study**

The results of this research are crucial for the development of an intelligent 5G wireless system, as they focus on increasing the system throughput by improving the accuracy of HO decision in 5G mm-wave communication. Other essential outcomes are described as follows:

- (i) This work paves the way for developing advanced intelligent HO decision-making so-

lutions to use in 5G and B5G systems operating at mm-wave and THz frequencies.

- (ii) To ensure the low latency is achieved by reducing the effect of HO delay.
- (iii) To reduce the UEs power consumption by reducing the frequency with which it scans and sends measurement reports to neighboring cells during the HO operation.
- (iv) To improve the overall throughput of the mm-wave communications system by reducing the number of HOs and optimizing the BS selection during HO.
- (v) This study is critical for infrastructure-to-infrastructure connectivity, such as vehicle-to-vehicle and device-to-device communication, for fast HO decisions for high-speed devices.

## **1.6 Delineation of this Study**

This study uses advanced machine learning algorithms to effectively make informed choices about selecting the next serving cell for a UE to HO to from a list of potential target cells. In this research, various experiments were conducted using RL algorithms as an optimization tool and the model developed proved efficient in terms of ensuring high QoS, less ping pong effects and reducing unnecessary HOs. The developed model works efficiently under all assumptions postulated in this thesis. The approaches considered are described as follows:

- (i) The BSs are distributed in a square dimensional area following a Poisson Point Process (PPP). Random waypoint user mobility model was used to simulate a more realistic UE movement with a velocity set at a given mean and standard deviation (since the velocity of the UE is not always constant).
- (ii) It was assumed that UEs take a specific route (trajectory-aware). Moreover, a trajectory-aware HO optimization approach was developed such that instead of using the exact user location, i.e. geo-coordinates of the user's location (since it is difficult to obtain), we correlate the user location to SNR values received from all BSs at a particular point. The mapping of exact locations to SNR values considers various kinds of obstacles that are found in a typical network environment, such as buildings, trees, vehicles and human beings. This evaluation was done using wireless Insite software™ (WI).

- (iii) Finally, an offline learning framework was developed to simulate the environment for data collection. Then, a Double Deep Reinforcement Learning (DDRL) algorithm was used to train an agent with the collected data as the data set. This was followed by testing of the developed model with a new data set generated by altering some of the points along the UE trajectory. This was done to test the robustness and generalization ability of the developed model to select the optimal BS that maximizes the user-BS connection. By choosing the optimal BS, this process simultaneously reduces the number of HOs per trajectory and guarantees high user throughput.

## CHAPTER TWO

### LITERATURE REVIEW

#### 2.1 5G System Architecture

The Next generation (NextGen) architecture is based on network function (NF) instead of a network entity that is obtained in LTE, according to 3rd Generation Partnership Project (3GPP) specification for LTE and new 5G systems (3GPP, 2020; Alsaeedy & Chong, 2018; 3GPP, 2017b). In LTE's core network (CN) also known as evolved packet core (EPC), the appropriate network protocols and interfaces are defined among the entities for each network entity (e.g. serving gateway (SGW) and the mobility management entity (MME)). In contrast, network protocols and interfaces in 5G CN (5G-CN) are specified for each NF. The NF is the processing functionality in 5G networks, and it can be implemented in three ways (Alsaeedy & Chong, 2018): a) as a network element on dedicated hardware, b) as a software instance running on dedicated hardware, or c) as a virtualized function built on an appropriate platform, such as a cloud infrastructure. The advantage of NF over NE is that it dramatically decreases latency. This is achieved by carefully controlling the UE mobility (e.g. tracking and paging procedures) scheme and separating the user plane from the control plane to ensure that each plane's resources are independently scaled and that more NF can be deployed in a distributed manner (3GPP, 2019). Figure. 3 shows the 5G system architecture along with NFs and reference points. A reference point shows the interaction between the services in two NFs (e.g. N4 is the reference point that connects UPF and SMF).

The architecture includes the user plane carrying UE traffic and control plane carrying UE signaling and control functions. The NF in the UP consists of user plane function (UPF) acting as a gateway for the UE traffic passing through RAN to external networks such as the Internet. It is responsible for packet routing and forwarding, packet inspection, QoS handling, packet filtering, and traffic measurement.

Several components of NFs run in the control plane. Some of the components are: access mobility function (AMF), session management function (SMF), network slice selection function (NSSF), unified data management (UDM), policy control function (PCF), authentication server function.

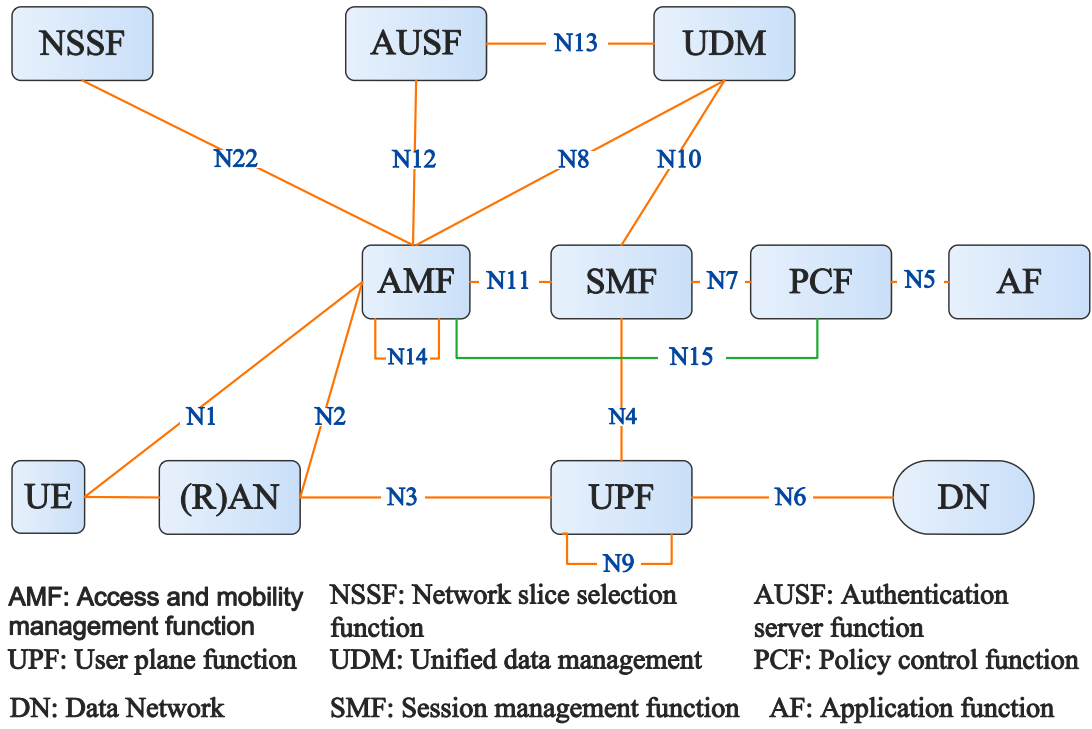


Figure 3: The 3GPP-5G architecture with reference points (3GPP, 2019).

AMF is responsible for control signaling between the CN and UE and provides additional functionality such as access authentication, authorization, mobility management control, location services, and integrity protection algorithms. SMF is primarily responsible for handling the interaction between the decoupled planes (user and control plane) and other functions such as internet protocol (IP) address allocation to UEs and policy enforcement. NSSF selects the network slice instance based on information provided during UE attachment. The UE's information provided include bandwidth, capacity and latency requirements for the application to be run on a network slice. Network slicing is a way of partitioning network resources to distinguish between the services provided to different UEs. UDM is responsible for credential authentication and access authorization while PCF takes care of the unified policy framework that governs network behaviour, and the policy rules for control plane function(s) that enforce them. Furthermore, authentication server function is used to facilitate 5G security processes.

In addition, there are other components of the NF that are found in the control plane but have not been depicted in Fig. 3. They include unstructured data storage function, unified data repository, network exposure function, and NR repository function. However, all depicted NFs can interact with the unstructured data storage function, unified data repository, network exposure function, and NR repository function. For further information on this functions, the reader is referred

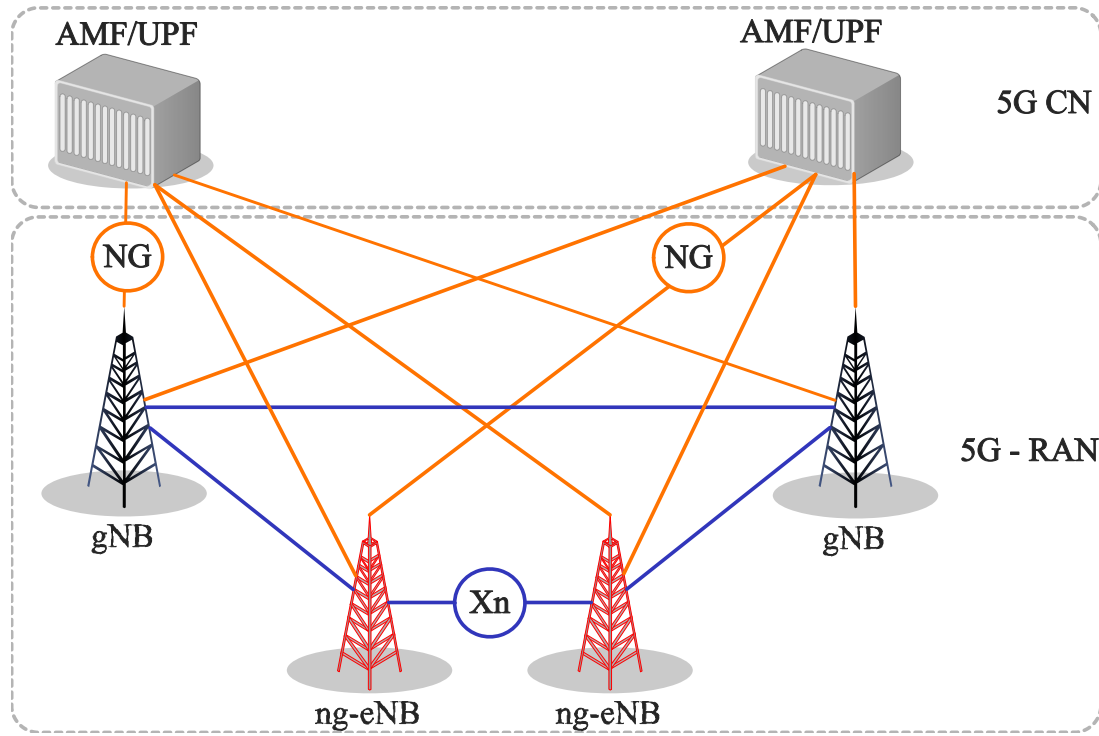


Figure 4: Overall architecture of 5G system showing network elements and interfaces (3GPP, 2020).

to (3GPP, 2019).

Overall, 5G architecture is divided into two parts, as shown in Fig. 4. The first part is the CN whose components have just been discussed while the second part is NextGen Radio Access Network (NG-RAN). The NextGen NodeB (gNB) serves as the access point for the 5G network, transmitting control plane and user plane traffic originating from N1, N2, N3 reference interfaces as shown in Fig 3. The purpose of the ng-eNB is to provide Evolved Universal Mobile Telecommunications System (UMTS) Terrestrial Radio Access (E-UTRA) user plane and control plane protocol terminations for UEs. In addition, 5G technology also supports LTE via ng-eNB. It allows existing 4G radio networks to coexist with the gNB. For example, if both LTE and 5G radio coverage are available, a 5G UE may use either LTE and 5G radio resources. Therefore, when there is no 5G coverage, LTE serves the 5G UE using the ng-eNB. The connection interface between gNB and ng-eNB is known as an Xn interface, and NG interface is the connection interface between gNB/ng-eNB and CN more specifically to the UPF the NG user-plane part (NG-U) and to the AMF the NG control-plane part (NG-C). The last interface that needs to be mentioned is the radio frequency interface, which is the circuit between the UE and the active gNB or ng-eNB which is also known as Uu interface. This interface supports a

broad spectrum from low to high frequencies (Ren et al., 2018).

5G is the first technology to introduce the use of high frequencies —above 6 GHz— intended for use in terrestrial mobile access networks along with the use of the sub-6 GHz spectrum. This is a real breakthrough that was made possible through different technological advances, including massive MIMO combined with beamforming technologies, advancements in processing capability of chipsets, and overall RF front-end/antenna subsystem integration/innovation for BS equipment and UE <sup>4</sup>(Chataut & Akl, 2020).

## 2.2 Mobility Management in 5G

Mobility management in 5G is quite different from that of legacy networks (2G-4G) and in this section, we present the concepts behind the radio access mobility in 5G cellular network. We also briefly explain the mobility state procedures in 5G system that makes it more efficient than legacy systems.

### **Definition 1** (Access stratum)

Access stratum (AS) is the set of protocols in 5G that contains the functionality associated with the UE's access to the RAN and the control of active connections between a UE and the RAN.

### **Definition 2** (Non-access stratum)

Non-access stratum (NAS) is the set of protocols in 5G that handles functionality operating between UE and CN.

### **Definition 3** (RRC context)

The radio resource control (RRC) context are the parameters necessary for establishing/maintaining communication between the UE and the CN.

### **Definition 4** (Cell selection)

Cell selection is the process of choosing a suitable cell<sup>5</sup> for the UE to camp on. This process is performed as soon as the UE is switched on (3GPP, 2018c).

### **Definition 5** (Cell re-selection)

Cell re-selection is the process of choosing a suitable cell after the UE camps on a cell and stays in the idle or inactive state.

---

<sup>4</sup>*Learn more about 5G mmWave Networks*, Published on 20 Oct. 2020. Available online at <https://www.5gmmwave.com/tag/mmwave/>. Accessed on 25 Nov. 2020.

<sup>5</sup>A cell with the measured cell attributes satisfy the cell selection criteria (3GPP, 2018e)

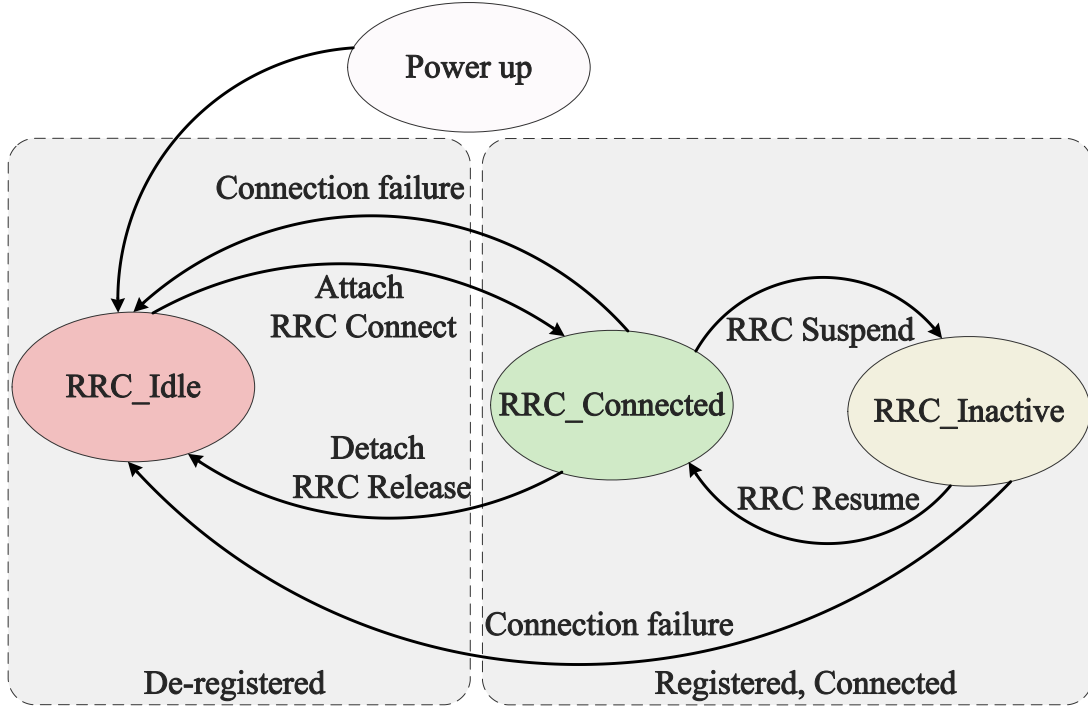


Figure 5: UE state machine and state transitions in 5G (3GPP, 2018b).

### 2.2.1 Radio Resource Control State Machine

The radio resource control (RRC) protocol is in the IP-level (Layer 3 /Network Layer) and is the protocol between UE and NG-RAN as specified by 3GPP (3GPP, 2018b). The RRC protocol's essential functions include: a) broadcast of system information, b) Control of the RRC connection —this procedure includes paging, establishment, modification and release of the radio bearer. It also involves establishing an RRC context, c) measurement configuration and reporting, and other functions specified by 3GPP that can be summarized in (3GPP, 2018b; Dahlman et al., 2018b). The RRC's operation is guided by a state machine that defines specific states where a UE may be present. The different states in this state machine have different amounts of radio resources that can be utilized by the UE once it enters into a particular state. Since different amounts of resources are available in different states, the state machine impacts the QoS that the user experiences and the energy consumption of the UE (Ahmadi, 2019b). In addition, RRC states provide a clear distinction between HO and cell (re-) selection. The UE can be in one of the three RRC states, namely: RRC\_Idle, RRC\_Connected, and RRC Inactive state. Figure. 5 depicts the UE state machine and state transitions in 5G while Table 4 summarizes the RRC protocols and functions in each RRC state.

Table 4: Summary of the RRC State and mobility handling in 5G.

RRC protocol	RRC State		
	Idle	Inactive	Connected
Network selection/registration	✓		
Cell re-selection	✓	✓	
5GC based Paging	✓		
NG-RAN based Paging		✓	
5G-RAN manages the UE RNA		✓	
5G-RAN knows UE serving cell			✓
Keep 5GC/5G-RAN connection for UE		✓	✓
UE AS context stored in 5G-RAN and UE		✓	✓

#### (i) **RRC\_Idle**

In RRC\_Idle state, the UE is not registered to a particular cell; hence, it does not have an AS context or receive any network information. This means that no specific link is established for communication between the UE and CN, and the UE does not belong to any specific cell. From the CN perspective, the UE is in the CN\_Idle state<sup>6</sup>, and the UE is in (a kind of) sleep mode and wakes up periodically (according to a configured discontinuous reception (DRX cycle)) to listen for paging messages from the network through the downlink channel. During this period, no data transfer takes place and the UE enters into sleep mode regularly to reduce battery consumption. The network can reach the UEs in the RRC\_Idle state by sending paging messages to notify them of changes in system information, warning messages such as earthquake and tsunami warning service, and commercial mobile alert system which are send as short messages. In this state, the UE manages mobility based on the network configurations via cell re-selections. It also performs the neighbouring cell measurements needed for cell re-selection in order to determine which cell it is to connect (explained in Section 2.2.2).

<sup>6</sup>UE is said to be in CN\_Idle state from CN perspective when no connection is established between UE and the CN (Dahlman et al., 2018b; Kim et al., 2017).

In legacy network (such as LTE), whenever the UE needs to connect to the network (e.g., when some uplink data becomes available for transmission), it must transit from idle to the connected state, and then return back to the idle state if there is no data to send. This process involves RRC signaling and latency. Moreover, today's UE needs to connect to the network very often, thereby leading to regular idle-connected-idle state transitions. These frequent transitions lead to increased network signalings and latency. To lower the network signaling overhead and reduce the latency involved during the transition to RRC\_Connected state, the RRC\_Inactive was introduced. In 5G, the network initiates the RRC release procedure to transit a UE from the connected to the idle state. In addition, as UE moves from the idle to the connected state, both the UE and the network establish the RRC context.

## **(ii) RRC\_Inactive**

5G-NR introduced RRC\_Inactive state from lessons learned during the development of LTE. The findings revealed that the transition of wireless devices from idle state to connected state is the most frequent high-layer signaling event in existing LTE networks, occurring about 500 – 1,000 times a day<sup>7</sup>. This transition involves a significant amount of signaling overhead between the UE and the network, as well as between network nodes, which can lead to increased latency and power consumption in the UE. The solution is to switch to RRC\_Inactive state which will result in a significant reduction in both latency and UE battery consumption. When the UE is in inactive state, its behaviour is similar to that in idle mode in term of power-saving. However, unlike the idle state, in the inactive state, RRC context is kept in both UE and gNB, and the UE is in CN\_Connected state<sup>8</sup> from the CN perspective, meaning that its connection to the CN is kept intact. Different from RRC\_Idle state, the primary purposed of RRC\_Inactive state is to reduce the network signaling load and latency involved during RRC\_Idle to RRC\_Connected state transition. In RRC\_Inactive state, the network signaling becomes faster since the AS context is stored in both the UE and gNB. While 5G CN connection is still retained - (UE remains in CN\_Connected state), the UE in RRC\_Inactive state is in sleep mode and wakes up repeatedly- according to configured DRX cycle (which in this case is controlled by the

---

<sup>7</sup>*Meeting 5G latency requirements with inactive state*, Published on 19 June. 2019. Available online at <https://www.ericsson.com/en/reports-and-papers/ericsson-technology-review/articles/meeting-5g-latency-requirements-with-inactive-state>. Accessed on 25 Nov. 2020.

<sup>8</sup>CN\_Connected state is when the UE establishes connection to the CN (Dahlman et al., 2018b; Kim et al., 2017).

5G-RAN), and regularly monitors for paging messages from the network. The procedure for notifying the UEs about any change of system information or warning message is the same as that of the idle state.

In RRC\_Inactive state, UE remains in CN\_Connected state and moves freely within an area configured by NG-RAN without notifying the RAN-based notification area (RNA). When it becomes necessary to transmit data or signaling, the resume procedure is initiated to enable the UE to switch from inactive to connected state. In this state, the UE controls mobility based on network configuration through cell re-selection with the same procedure as in the idle state.

### (iii) **RRC\_Connected**

In the RRC\_Connected state, the RRC context and all parameters needed to establish communication between the UE and the RAN are known to all entities. This means that in RRC\_Connected state, the network configures all required parameters for communication between the network and the UE. In RRC\_Connected state, the UE is in CN\_Connected state from the CN point of view. The cell to which the UE belongs and the UE's identity is known. In addition, the cell radio-network temporary identifier (C-RNTI) used for signaling purposes between the UE and the CN is configured. The connected state is intended to transmit data to or from the UE, and to minimize excessive power consumption of the UE. DRX is optimized while maintaining user's QoE (Liang et al., 2018). With a configured DRX cycle, the UE only monitors downlink signaling when active, and then goes into sleep mode for the rest of the time with the receiver circuitry turned off. This process allows significant power consumption reduction, as the longer the DRX cycle, the lesser the power consumption. For exhaustive discussion on how DRX reduces excess power consumption, please refer to (Dahlman et al., 2018a; Da Silva et al., 2016). Also, the RRC context is established in gNB for the connected state, therefore, data transmission/reception can commence relatively fast, as no connection setup, and signaling is needed. In this state, the network manages mobility by the process known as HO, explained in the Section 2.3.5

As regards cell re-selection when leaving RRC\_Connected state, the UE attempts to camp on a suitable cell according to *redirectedCarrierInfo* when transitioning from RRC\_Connected state to RRC\_Idle or RRC\_Inactive state (3GPP, 2018e). In the connected state, if the network initiates the RRC release message or the UE and CN are no longer attached, the UE moves into an

idle state, on the other hand, if the network initiates the RRC suspend procedure, the UE would transit from connected to inactive state (3GPP, 2018e,b) (see Fig. 5).

One significant difference among the different states, as seen from the preceding discussions, is the mobility mechanisms involved. Efficient mobility management is an essential aspect of any mobile communication system. In the following subsections, we describe the different mobility mechanisms including idle- and inactive-state mobility.

### **2.2.2 Idle and Inactive State Mobility**

Most importantly, RRC states ensure that the mobile UE is accessible via network mobility mechanisms, mainly when the UE is in the idle or inactive states, during which it has limited connection to the network. The network, through paging, communicates with the UE occasionally, and also sends short broadcast message which carries information about changes in the system (3GPP, 2018e). The area over which a paging message is sent is an essential feature of the paging process. Also, in both states, the device can switch from one cell to another via cell re-selection. The UE scans for candidate cells for cell re-selection, and if the UE discovers a cell with received power sufficiently higher than its current one, it deems this the best cell and contacts the network through random access (3GPP, 2018e).

UE tracking needs to be intelligently carried out to avoid high overhead due to paging, and signaling at the network and cell level respectively. High overhead would occur if the network broadcasts the paging message from every cell in the network in order to track the target UE as paging transmissions would occur even in cells that do not contain the target UE. On the other hand, signaling would occur if only the cell which the UE is located broadcast the paging message. In this case, the UE needs to inform the network whenever it moves from one cell's coverage to another. This would lead to high overhead due to the signaling required to notify the network about the updated UE's location. Hence, the cell-group level tracking system was introduced in 5G-NR to tackle the challenge of high overhead due to signalling and paging. Figure. 6 illustrates how tracking of UE in the idle and inactive state is carried out in 5G-NR. In order to enable effective UE tracking, the cells are organized into cell groups, and the UEs are only monitored on the cell-group level, as shown in Fig. 6. The network only receives new UE location information when the UE moves into another cell group outside its previous cell-group. In case of paging the UE, the broadcasted paging message is sent to all cells within the specific

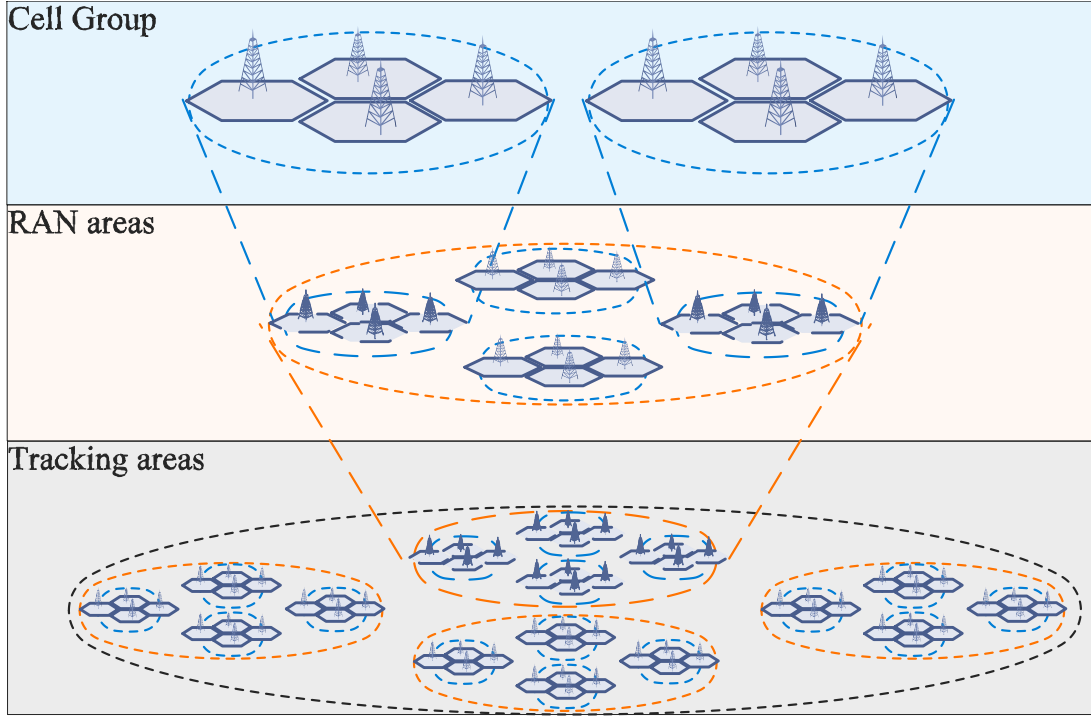


Figure 6: RAN Areas and Tracking Areas.

cell group — this is done to reduce the paging overhead. This is the primary tracking procedure in the NR for both states. However, there is a difference in the way that cells are grouped in both states as well as how paging is initiated.

For the idle state, cell groups are grouped into *RAN areas*, where a *RAN area identifier* (RAI) identifies each RAN area. The *RAN areas*, in turn, are grouped into an even larger group known as tracking areas, where a tracking area identifier (TAI) is used to identify tracking area. Thus, each cell belongs to one cell group which also belongs to one RAN area as well as a tracking area, and their respective identities are provided as part of cell system information.

Tracking areas are the basis for CN-based UE tracking, and the CN is responsible for managing and initiating paging. The CN assigns each UE to a *UE registration area*, which consists of a list of TAIs. When a UE enters a cell belonging to a tracking area outside its assigned registration area, it accesses the CN and performs a *non-access stratum registration update*. The CN records the UE's location and updates the UE's registration area, then it provides the UE with a new TAI list that includes the TAIs that the UE has been assigned. The UE is assigned a set of TAIs to avoid repeated *non-access stratum* registration updates in case the UE moves back and forth between two neighbouring tracking areas. If the UE moves back to the old TAI within the

updated UE registration area, no new update is needed.

In the inactive state, RAN Area becomes the basis for UE tracking, which is carried out in the 5G-RAN level. 5G-RAN is responsible for initiating the paging and managing RAN-based notification area. UEs are assigned a RNA comprising the following: a list of cell identities, a list of RAN areas, or a list of tracking areas. The RNA is assigned to a UE by its serving NG-RAN based on the UE's registration area and can cover a single or multiple cells (a subset of the tracking areas). As a result, the UE can move freely within the allocated RNA without contacting the NG-RAN. However, if it moves to an area outside its current RNA, it initiates RAN-based Notification Area Update (RNAU). Once the serving cell (ng-eNB or gNB) receives the RNAU request from the UE, it may send the UE to one of the following RRC states: RRC\_Inactive, RRC\_Connected, or RRC\_Idle. If UE remains in the inactive state, the serving NG-RAN may continue to send a periodic RNAU timer to the UE, which is used to notify the network that the UE is still active. The value of the RNAU time is assigned based on the RRC\_Inactive assistant information (RIAI) (Alsaeedy & Chong, 2018). In summary, two levels of paging can be applied for reaching the UE depending on its RRC state: CN-based paging for idle state and 5G-RAN-based paging for the inactive state (see Table 4).

### **2.2.3 Connected State Mobility**

The connection between UE and network is established in the connected state. Connected-state mobility aims to maintain connectivity without interruption or noticeable degradation as the UE moves within the network. To maintain the connection between UE and network in the connected state, the UE is continuously searching for new BSs to connect to. The BS search is based on current carrier frequency (intra-frequency measurements) and different carrier frequencies (inter-frequency measurements) from the UE perspective.

Cell search in the connected state results in HO if suitable condition are met while for idle and inactive state, it results in cell re-selection. When it becomes necessary to perform HO in the connected state, the UE is not responsible for the decision. Instead, the UE performs signal measurement of the serving cell and neighbouring cells and generates the measurement report (MR)—containing cell level measurement results such as reference signal received power (RSRP), signal-to-interference-plus-noise ratio, reference signal received quality, etc.—sent to the network. Based on this report, the network decides whether or not the UE is to HO

to a new cell. The above procedure is not applied to the very small SCs (e.g 5G femtocell) that are tightly synchronized to each other (3GPP, 2020; Ahmadi, 2019b).

## **2.3 Handover Management in 5G and Beyond**

This section describes the step-by-step procedure for HO in 5G-NR, introduces the various categories of HO, and also discusses HO requirements alongside its relationship with radio resources management.

### **2.3.1 Types of Handover**

There are two broad categories of HO, namely; intra-/inter- frequency and intra-/inter- radio access technology (RAT) HO.

#### **(i) Intra-/Inter-Frequency Handover**

Intra-frequency and inter-frequency HO are the HO types for which the carrier frequency is the subject of interest. If the UE is to move to the target cell with the same frequency as that in the serving cell, it is generally known as intra-frequency HO as seen in Fig. 7 Scenario 1. In contrast, inter-frequency HO occurs if the UE is to use a different carrier frequency in the target cell as shown in Scenario 2 in Fig. 7. Event A3 and A6 initiate intra-frequency HO. Both Event A3 and A6 are triggered when the neighbouring BSs RF condition is higher than that of the S-BS. Moreover, Event A6 is used for intra-frequency HO of the secondary frequency on which the UE camps. Event A4 and A5 are typically used for inter-frequency HO. Event A4 is triggered when the RF condition of one of the neighbouring BSs is higher than the threshold compared to that of the other BSs. On the other hand, Event A5 is triggered when the S-BS RF condition becomes lesser than the lower threshold and the RF condition of one of the neighbouring BS becomes higher than the upper threshold (where the threshold values are parameters that are optimized based on the network) (3GPP, 2018d,b).

As mentioned in Section 2.2.1.iii, HO occurs in the connected state and in that state, UE regularly sends the MR—containing cell level measurement results such as RSRP, signal-to-interference-plus-noise, reference signal received quality, etc.— of all neighbouring cells to the serving cell. Once the conditions for Event A3 or A4 are met, the serving cell communicate with the target cell and starts the HO procedure. Intra-frequency HO is the most commonly selected

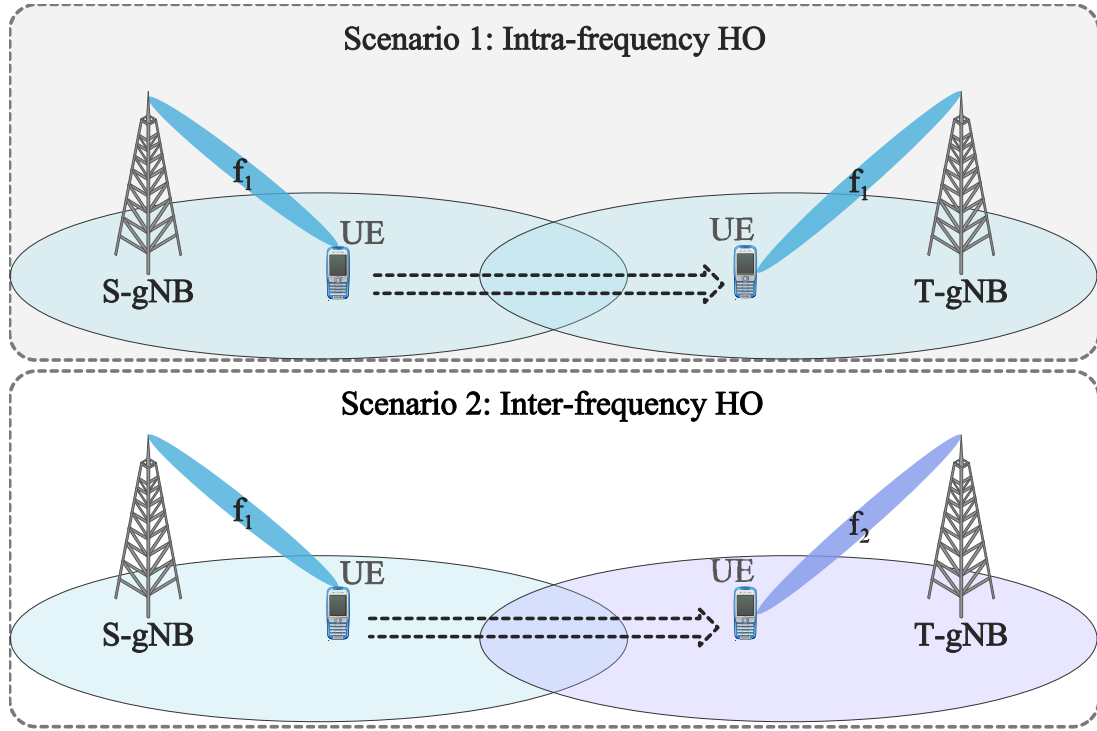


Figure 7: An illustration depicting intra-frequency HO in scenario 1 and inter-frequency HO in scenario 2.

choice of HO in mobile networks; however, there are certain instances where it is more preferable to choose inter-frequency HO. Such instances include: a) when the serving frequency does not have sufficient bandwidth to satisfy user service demands thereby resulting in poor QoS or call failure, b) when the requested service cannot be supported by current serving cell e.g. Voice over NR, and c) when the serving frequency is overloaded and load balancing becomes necessary to reduce congestion (Zaidi et al., 2020). The UE essentially carries out the measurements in the measurement gap at different frequencies for inter-frequency cases (3GPP, 2018d) and (Tayyab et al., 2019). The measurement gap is necessary because without it, the UE would not be able to measure the target carrier frequency while transmitting/receiving to/from the serving cell simultaneously. The measurement gap specifies the time interval when no downlink or uplink signal is transmitted. The need for measurement gap in 5G-NR depends on the capability of the UE, the active part of the UE bandwidth part and the current operating frequency (3GPP, 2018a).

5G-NR supports intra-frequency, inter-frequency, and inter-RAT measurement gaps. Unlike intra-frequency HO in LTE, intra-frequency measurements in NR may require a measurement gap for cases where the intra-frequency measurements are to be done outside the active band-

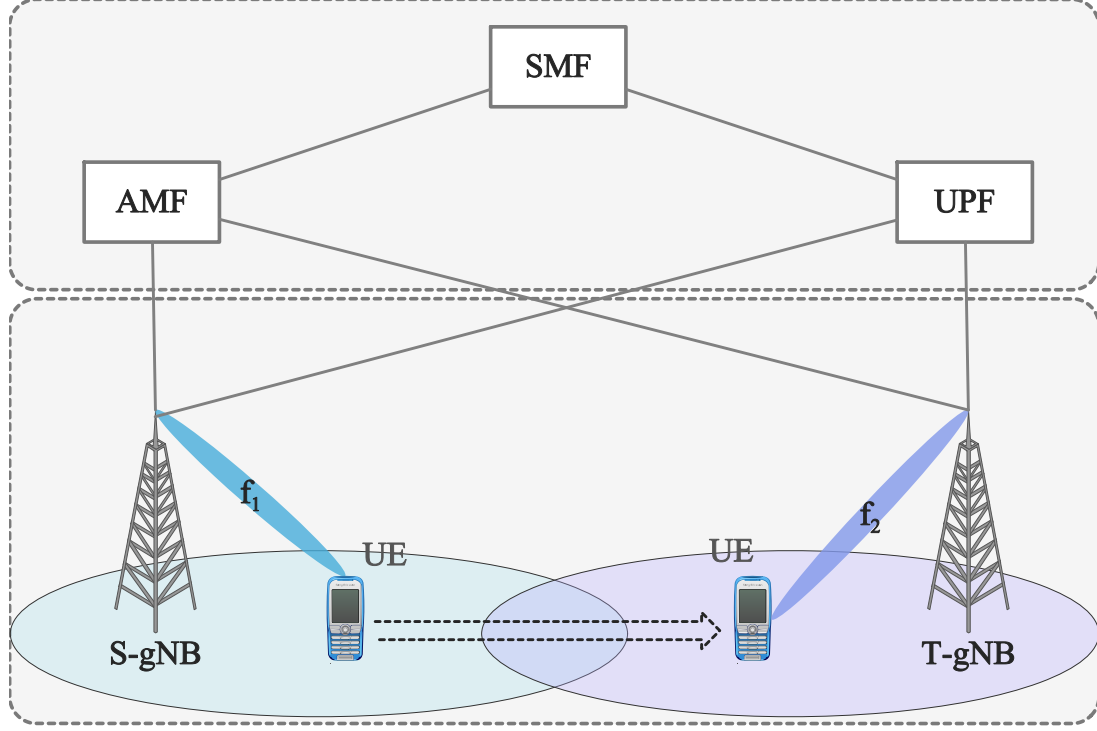


Figure 8: UE undergo HO from once cell to another with both cells use the same RAT (intra-RAT).

width part. That means that the measurement gap only applies to some cases of intra frequency HO where enhanced UE coverage is not guaranteed to be aligned with the serving gNB's centre frequency (3GPP, 2018d; Tayyab et al., 2019). However, the measurement gap is required for all cases of the inter-frequency HO as specified in 3GPP (3GPP, 2018d). Researches are concerned with fundamentals question on how the measuring gaps can be reduced, as large measuring gap results in lower throughput and higher UE energy consumption.

## (ii) Intra-/Inter-RAT Handover

In the case of intra-RAT HO, UE hands over from S-BS to the T-BS which both use the same RAT. Intra-RAT HO is commonly referred to as horizontal HO (Tayyab et al., 2019) as shown in Fig. 8. Intra-RAT HO can be either intra- or inter-frequency HO. Intra-RAT HO aims to preserve the connectivity of the UE with the existing network and the primary reason for this kind of HO can be attributed to load balancing or measurement trigger conditions. Once UE HO occurs, it prefers to camp on the cell which provides the strongest received signal.

In contrast to intra-RAT HO, inter-RAT (or vertical) HO occurs when the UE hands over to a T-BS which uses a different RAT from the S-BS. Unlike in intra-RAT HO where the cell with the

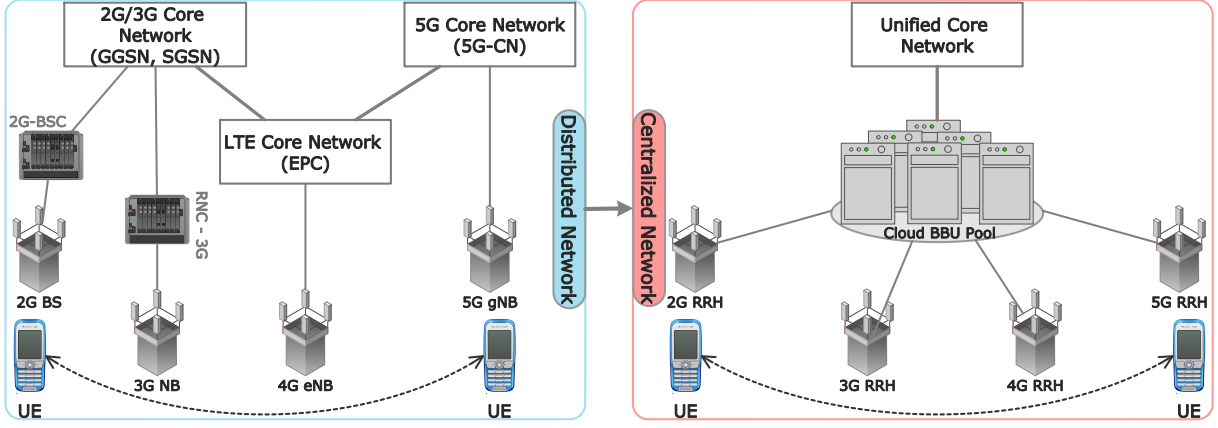


Figure 9: Inter-RAT HO scenarios in distributed and centralized RAN architectures.

highest received signal is selected, in inter-RAT HO, other factors such as user mobility, service type, as well as the network property and state are considered when selecting the target cell. It also involves the switching of the logical interface between the two RATs (Ahmadi, 2019a). The latency incurred during inter-RAT HO is still prohibitive for many application and services, thus, it poses a severe problem in the NexGen mobile systems (Ahmadi, 2019a). In order to improve the user experience, centralized architecture for inter-RAT HO, which integrates legacy and NR network protocol was proposed (Ahmadi, 2019a). Figure. 9 demonstrates how the UE performs inter-RAT HO. From the figure, it can be seen that both distributed, and centralized CN architecture for multi-RATs are possible. The advantage of using centralized architecture is that it can lead to a significant reduction in HO signaling and interruption time<sup>9</sup>. The centralized architecture comprises unified CN along with the baseband unit (BBU) and remote radio head (RRH) separated through a transport mechanism such as optical fiber. In a C-RAN architecture, the RRHs are connected to the BBU pool through high-bandwidth transport links known as fronthaul.

Other HO schemes such as intra/inter cell layer HO, which is the HO between two different cell layers (macro, micro and pico) and intra-/intra-operator HO, which is the HO that involves the same operator or different operators (roaming) are not considered in this thesis as they already fall into any of the two categories of HOs already discussed.

<sup>9</sup>Michael Wang, *5G, C-RAN, and the Required Technology Breakthrough*, Published on 21 Jun. 2018. Available online at <https://medium.com/@miccowang/5g-c-ran-and-the-required-technology-breakthrough-a1b2babf774>. Accessed on 25 Oct. 2020.

### 2.3.2 Handover Requirements and Key Performance Indicators

Since HO has adverse effect on the overall performance of wireless networks, different features and requirement are necessary to reduce the impact of HO. Also, various KPIs are used to measure how the network performs during a HO. The various HO requirements and KPIs are presented as follows:

- (i) **Seamless HO:** a seamless HO occurs when UE perceives continuation of connection during HO with little or no interruption during gNB switch. This guarantees the UE's active connection.
- (ii) **HO interruption time:** is a period where the UE is not permitted to send user plane packets to the BS.
- (iii) **HO failures rate:** For any given UE trajectory or unit time, the HO's failure rate is the number of HO failures—unsuccessful HOs—divided by the number of times the UE experienced the HOs.
- (iv) **Ping-pong** refers to situations in which the number of HOs over a specified period reaches a predetermined threshold ( $T_{pp}$ ), for example, more than one HO in every 10 sec. The ping-pong effect degrades the QoS by introducing defections, including throughput reduction, long HO delay, and high dropping probability due to these unnecessary HOs.
- (v) **Signaling overhead:** HO signaling overhead are the various data generated during the process of HO to facilitate the operation. However, the HO process interrupts the data flow and results in the reduction of the UE throughput (Zaidi et al., 2020).

There are other performance metrics that are essential to ensure optimal performance in wireless networks, particularly for HO optimization. Further details is presented by Tayyab et al. (2019).

### 2.3.3 Handover and Radio Resource Management

In 5G, the term radio resource includes both traditional (from the legacy system) and extended resource concept (Li et al., 2017). These legacy resources include energy consumption (cell and UE transmitting power), frequency (channel bandwidth, frequency of the carrier) and antenna configurations. In addition, the extended resource definition in 5G covers the hard resource

(number/type/configuration of antennas, the existence of nomadic/unplanned nodes, or mobile terminal relays) and soft resources (network node and UE software capabilities). It is also important to meet UE requirements such as QoS or QoE for all the UEs while properly managing resources. On the other hand, proper resource management can help networks fulfill HO KPIs, for example, by reducing the probability of HO failures while maintaining the QoE during and after HO (Van Quang et al., 2012; Long, 2018). To increase wireless system efficiency, it is necessary to address and take into account the fundamental issues related to HO and resource management such as admission control, bandwidth and power control.

#### **2.3.4 Dual Connectivity**

Dual connectivity means that the UEs can establish connection to two different cells at the same time (Shayea et al., 2020). Usually, in dual connectivity, UEs either connect to BSs of different sizes (macro cell and SC) or two different RATs simultaneously (e.g. 4G and 5G network), as illustrated in Fig. 10. Since the UE can be connected to two different RAT over different frequency bands simultaneously, the interruption time is reduced to zero. However, this would trigger an additional likelihood of HO where new HO cases are introduced relative to a single connection. These new HO scenarios (see Fig. 10) occur in two situations; when the UE switches the connection either from SC to macrocell or from SC to SC. With the introduction of mm-wave, the use of dual connectivity could lead to an increase in HO probability, thereby causing additional problems with mobility management, including an increase in signaling overhead, synchronization complexity between RATs for multi-RAT connectivity, simultaneous utilization of resources in multiple BSs, and reduction in battery lifespan. The increase in signaling overhead is due to flow control between the RATs (Shayea et al., 2020; Antonioli et al., 2018), and these issues could be addressed using intelligent approaches.

#### **2.3.5 Handover Management in NR**

NR physical layer uniquely differs from the legacy RAT with the following features: dual connectivity, high-frequency spectrum, forward compatibility, ultra-lean design, use of mm-wave and relay for devices (device-to-device). NR supports both multi-connectivity and single-connectivity selection depending on the configuration set. For both configurations, hard HO is used during path switching (Zaidi et al., 2020). In both licensed and unlicensed spectrum, NR operates between 600 MHz and 73 GHz. Forward compatibility means designing radio-

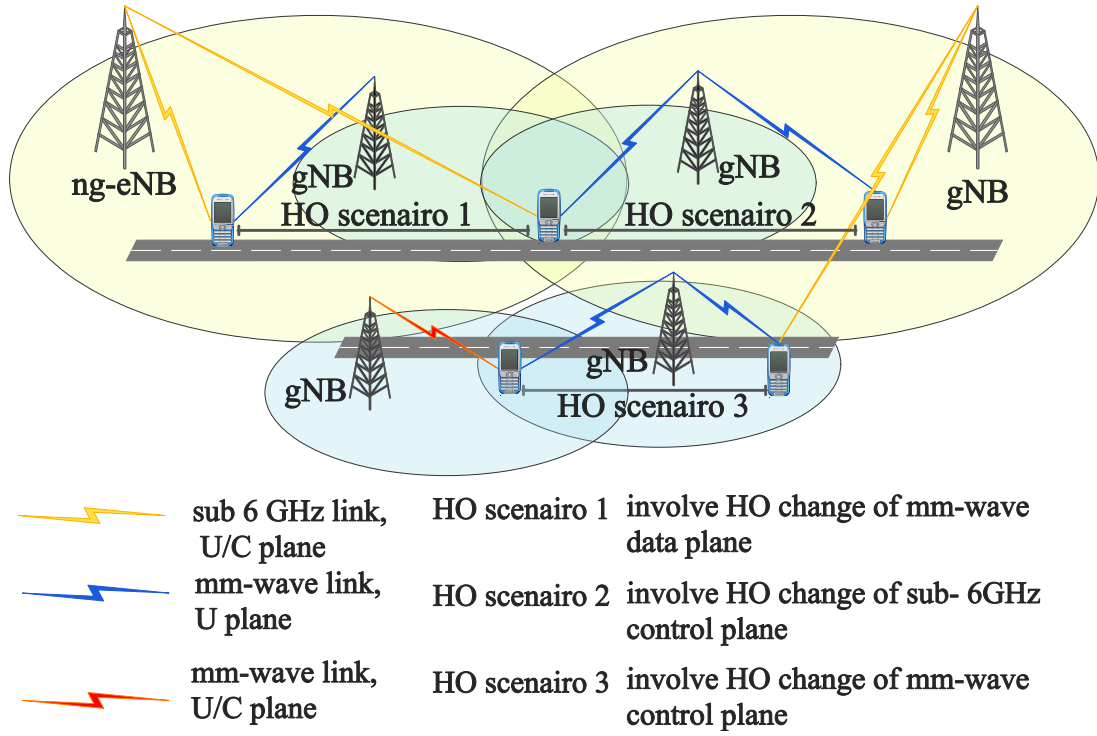


Figure 10: Dual connectivity with HO scenarios in future communication networks.

interface architecture that enables new service requirements and accommodate new technologies while supporting legacy network UEs. While the ultra-lean design principle aims to decrease the always-on transmissions (for example, signals for BS detection, broadcast of system information) to achieve high data rates with low energy consumption in the network. The main challenge for NR is the coverage due to the use of high frequency with high penetration loss that makes the cell footprint to become smaller. In this section, we describe the NR HO with a brief introduction of critical features and the entities involved in NR mobility. Also, a step by step HO procedure is provided for intra-AMF/UPF. The types of HO in NR are described as follows:

**1) Intra-gNB HO:** This occurs when both the source and target cells<sup>10</sup> belong to the same gNB, as shown in Fig. 11.

**2) Inter-gNB HO without AMF Change:** Inter-gNB HO generally occurs when serving and target cells are from different gNBs. There are two different types of HO within inter-gNB HO without AMF change, depending on whether the HO involves a change of UPF or not. However, the inter-gNB HO discussed here does not include a change of AMF in both cases, as shown in

<sup>10</sup>Cell here means the part of sector gNB that has specific beams and covers the specific environment.

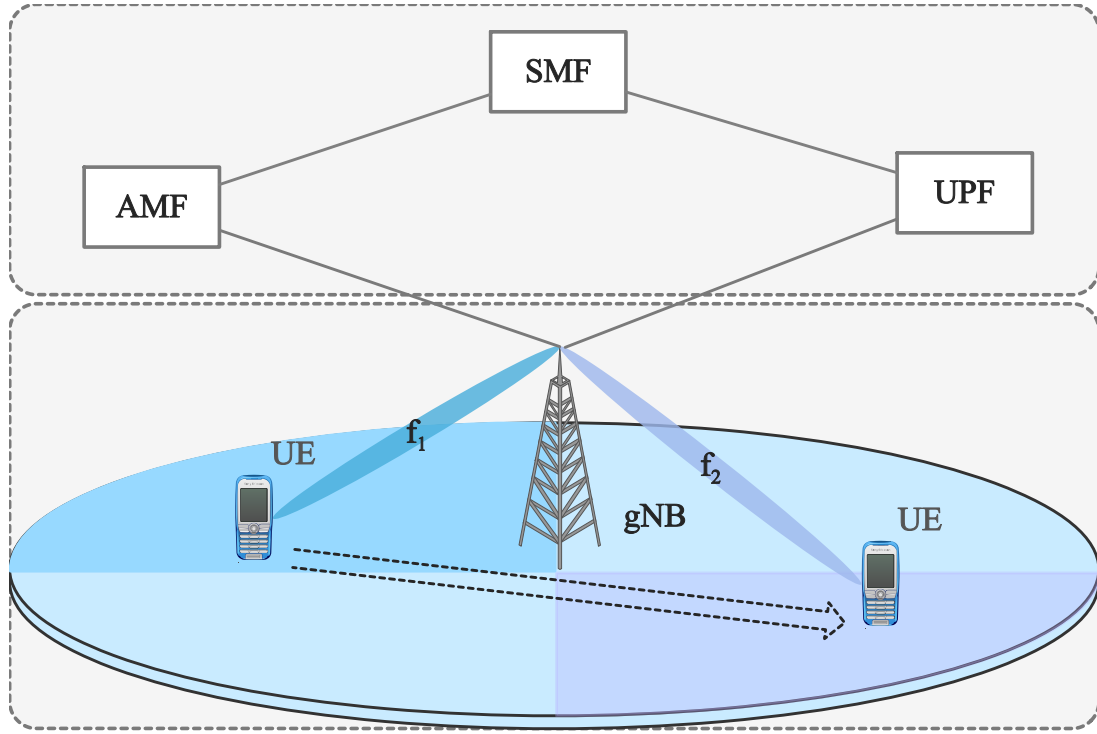


Figure 11: UE performs intra-gNB HO which involves the change of cells in the same gNB.

Fig. 12. Inter-gNB with intra-UPF HO is presented in Fig 12 scenario 1 , where the HO involves a cell change with the same UPF, while Fig. 12 scenario 2 presents inter-gNB with inter-UPF HO where the cell switch involves a change of UPF.

**3) Inter-gNB HO with AMF Change:** In this case, the HO requires a change of AMF from the source to the target AMF. However, the HO involves no change of SMF, and only the NG interface is used as depicted in Fig. 13. There are two cases of inter-gNB HO with AMF change; in the first case (Fig. 13 scenario 1), the same UPF is maintained while the second case (Fig. 13 scenario 2) involves a change of UPF during HO. The basic HO procedure in NR is shown in Fig. 14 (3GPP, 2020; Tayyab et al., 2019). It consists of three phases, namely: HO preparation (Steps 0-5), HO execution (Steps 6-8) and HO completion (Steps 9-12), which are described as follows:

- **Step 1:** the UE measuring procedure is configured according to access restriction and roaming information by the serving gNB (S-gNB), and the UE sends an MR to the target gNB (T-gNB).
- **Step 2:** the S-gNB determines to HO the UE, based on the MR and radio resource management information.

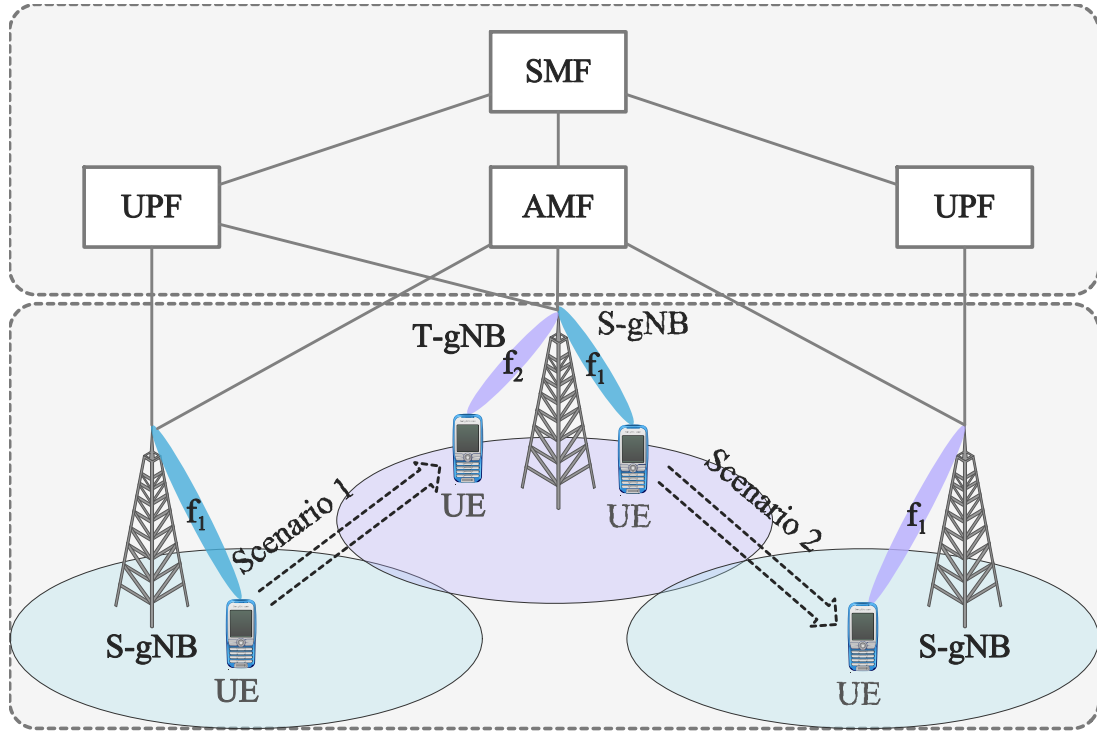


Figure 12: UE performs inter-gNB HO, which involves the change of gNBs with same UPF and AMF for scenario 1 and change of UPF for scenario 2.

- **Step 3:** the S-gNB sends a HO request message to the T-gNB (which includes the necessary information to prepare for HO to the T-gNB).
- **Step 4:** the T-gNB executes the admission control procedure if the T-gNB can grant the resources.
- **Step 5:** the T-gNB sends a HO request acknowledgement to the S-gNB. As soon as the S-gNB receives the HO request acknowledgement message, data forwarding may be initiated.
- **Step 6:** S-gNB sends a HO command to the UE.
- **Step 7:** S-gNB sends the Sequence number status transfer message to the T-gNB.
- **Step 8:** UE detaches from the S-gNB and synchronizes with the T-gNB.
- **Step 9:** the T-gNB informs the AMF that the UE has changed the cell, through the Path switch request message.
- **Step 10:** 5G-CN switches the downlink data path towards the T-gNB.

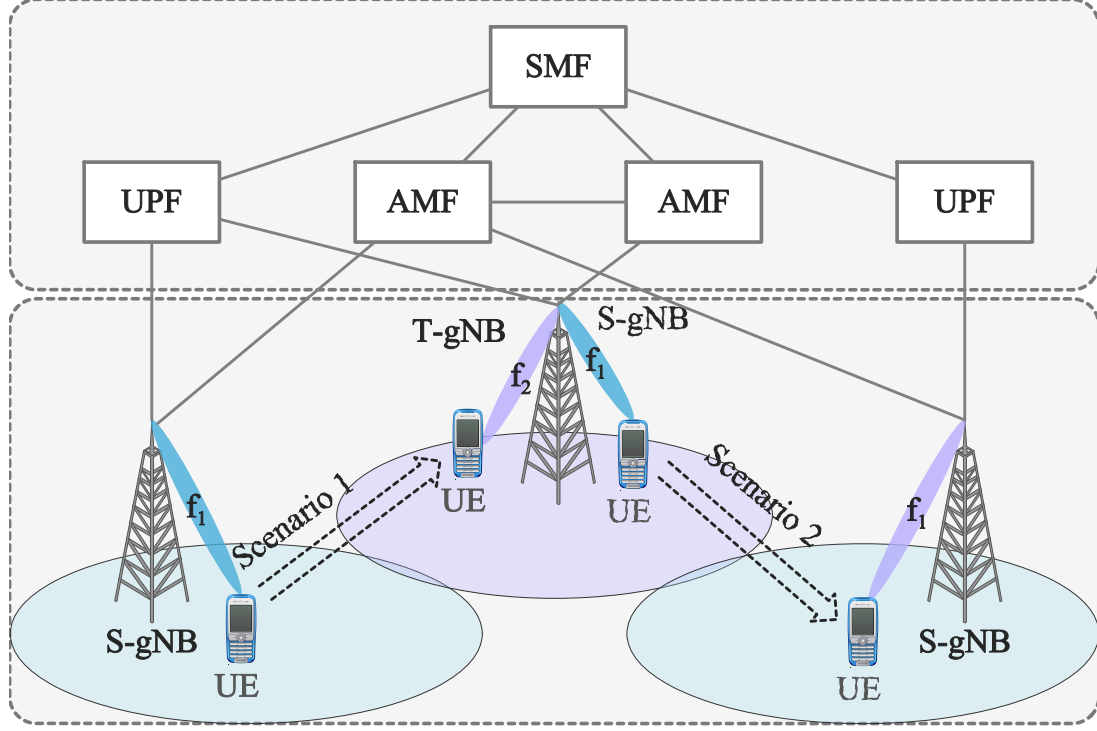


Figure 13: UE performs inter-gNB HO with AMF change, involving the change of gNBs while UPF is maintained in scenario 1 and change of UPF in scenario 2.

- **Step 11:** the AMF acknowledges the Path switch request.
- **Step 12:** the T-gNB informs the S-gNB that the HO was successful and triggers the release of resources by the S-gNB by sending a UE Context Release message. Finally, the S-gNB release the radio resources associated with the UE.

It is essential to point out that the above procedure is applied for HO between NR and NR technologies.

### 2.3.6 Mobility and Handover Management in B5G

Researchers have anticipated some use cases and applications that make B5G to be different from 5G. Some of these use cases include integrated unmanned aerial vehicles (UAVs) communications, high mobility of devices (above 500 kmph), holographic projection, etc (Yang et al., 2020). The high mobility of devices, UAVs, and other applications that use radio waves at the mm-wave and THz spectrum presents unprecedented wireless communication challenges in B5G. Among these challenges, mobility and HO management are anticipated to be the two most challenging issues in B5G networks since B5G networks would be highly dynamic, and

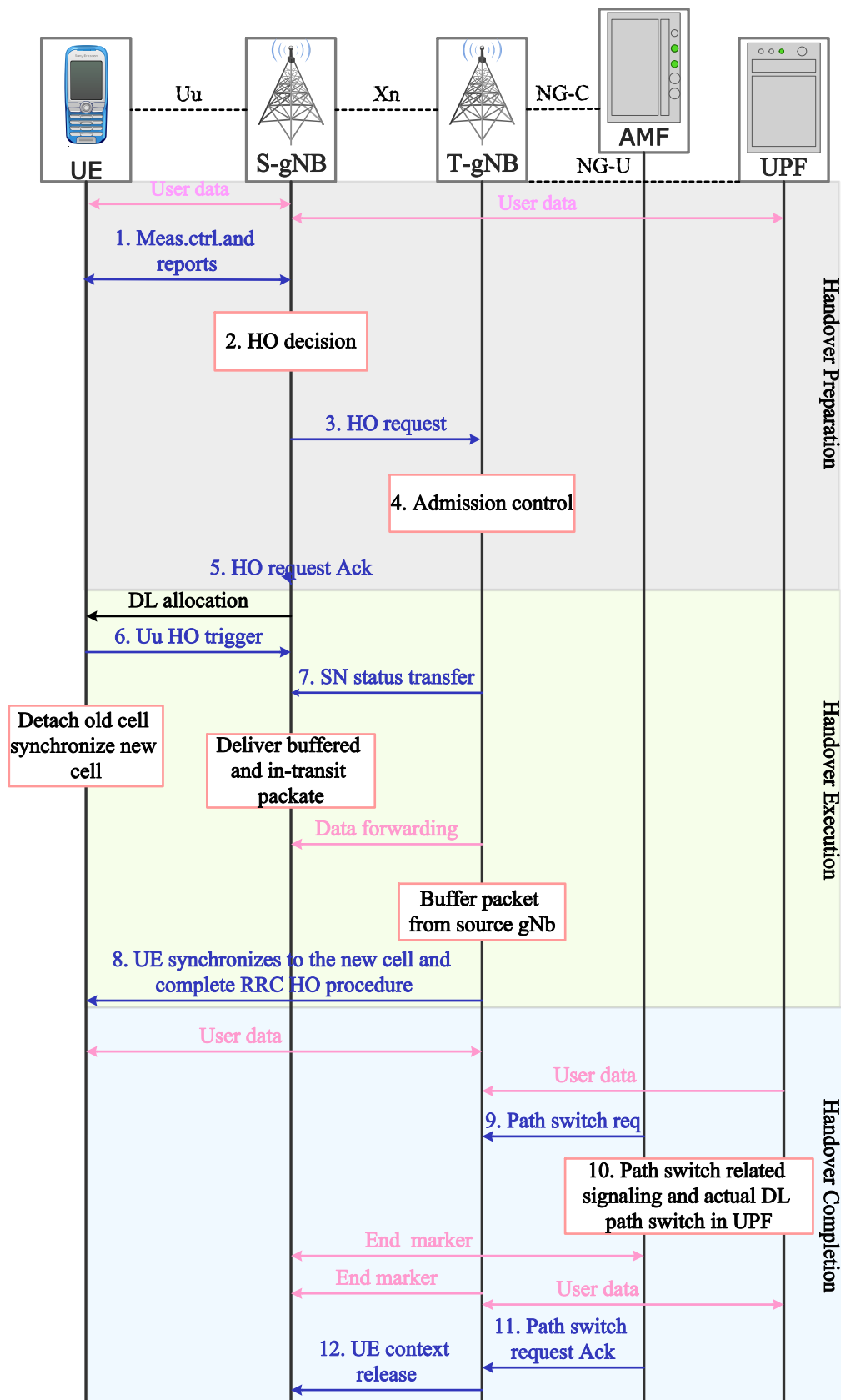


Figure 14: HO procedure in 5G-NR involving no change of AMF and UPF (3GPP, 2020).

multi-layered, which would lead to more frequent HO. High mobility of devices and UAVs results in uncertainties of their locations and keep in mind that high frequencies such as mm-wave and THz that would be used in B5G can be easily blocked by humans, buildings, etc.

Heuristic and traditional HO methods would not be able to react quickly. An alternative solution is to adopt artificial intelligence models in order to achieve mobility prediction and optimal HO strategy in order to guarantee communication connectivity. Even though the introduction of multi-connectivity is a very promising solution, the procedure still needs intelligent HO management strategies to optimize the cell (re-)selection process in order to reduce signaling, guarantee high data rate, high reliability, and low latency in the B5G (Yang et al., 2020). The HO procedure for the B5G might be similar to that of 5G, but there are no standards for B5G system yet.

## **2.4 Machine Learning for Handover management**

The use of mm-wave and higher frequencies in 5G and B5G networks is going to introduce new challenges and complexity to the HO management that would be difficult to handle by conventional methods. Firstly, these frequency ranges suffer from severe attenuation (e.g., larger penetration losses), which means their transmission distance will be small. As a result, more BSs need to be deployed to cover the same area that would have been covered by those utilizing the microwave frequencies (Xiao et al., 2017). This implies that the size and the complexity of the network is going to greatly increase and the users will be prone to more frequent HOs which would greatly affect their QoS, particularly for high mobility users and applications.

Secondly, due to the use of directional beams for transmission in mm-wave networks, the presence of obstacles on the path of the transmitted beam can partially or completely hinder the user from gaining access to the network or negatively impact the signal quality. As such, in mm-wave communication networks, the users are not only faced with the challenge of selecting the optimal BS but also the optimal beam to connect to per time in order to maximize their QoS. Hence, optimal beam selection has become another factor to consider in HO management process which would further add more complexity to the HO process because of the massive number of beams that the user has to select from during each HO instance (Kim et al., 2016; Niu et al., 2015).

Finally, there is also the need to provide some high mobility based essential services for emergency scenarios such as medical services to patients in ambulances en-route hospital through real-time consultations with the doctors that are situated in a remote hospital. Especially, in the pandemic situation that we find ourselves in now, this kind of services may be needed to sustain the lives of the patients in critical conditions before they get to the hospital to receive proper medical attention (Usman et al., 2019; Abubakar et al., 2020). Intelligent HO optimization would help predict the route of the ambulance, determine the optimal BSs to connect and also pre-allocate the resources that will be needed at the BSs. This will help prevent intermittent service interruptions and guarantee the QoS need to support the communication between the paramedics in the ambulance and the doctors at the remote location (Tabassum et al., 2019; Öztürk, 2020).

Therefore, effective HO optimization would enable the selection of the optimal BS and beam for user connection that will maximize user connection, reduction of excessive or unnecessary HO, the detection of obstacles and their avoidance. These are some of the issues that make HO optimization in mm-wave communications networks more challenging to handle compared to the previous generation of cellular networks. Moreover, since the HO process involves various network parameters that must be considered and optimized in real time in order to ensure seamless HO, this would be very challenging for most conventional methods to handle. The challenge with conventional methods of HO management is that they are computationally demanding to implement, particularly when the network dimension becomes very large. As such, before they can decide which T-BS to associate the user with, the user must have moved from that location. This would result in sub-optimal HO decision and degradation in user QoS. In addition, they cannot accurately capture certain details of the network such as the presence of different types and sizes of obstacles, as well as the dynamic traffic demand patterns that are typical of 5G and B5G networks, which are also important for making an optimal HO decision (Wang et al., 2019c; Tabassum et al., 2019; Bui et al., 2017).

However, ML techniques can assist in bringing intelligence and helping the network to self-optimize. ML techniques are able to learn various network characteristics from data generated from the network, in order to optimize various aspects of the network. They are able to capture hidden details and patterns in the network from the network data that cannot be represented by analytical models (Wang et al., 2019c). They are self-adaptive and as such can react to changes

in network environment and even predict future network or user demands before hand, thereby enabling the network to adequately prepare to handle such demands when they occur (Tabassum et al., 2019). They can be designed in a computationally efficient manner such that the training phase of the algorithm, which is often computationally demanding, can be carried out offline, and then the trained model deployed online to carry out real-time optimization after which the model can be updated periodically, as it experiences new data (Hussain et al., 2020).

In this section, we first present an overview of the major categories and types of ML algorithms used for HO optimization. Then, we delve into reviewing the state-of-the-art on ML-aided HO management. A top-level taxonomy is followed while reviewing the state-of-the-art, such that the ML-aided HO management methodologies are classified based on the source of the data they utilize. As such, two broad categories are considered: visual data based and wireless data based HO management techniques. The major objective of this novel taxonomy is to recognize the visual data aided HO management schemes—which has been long overlooked in the literature—by giving it a special place along with the traditional wireless data driven HO schemes. The visual aided wireless communications is an emerging research area in wireless communications where visual information (pictures/videos) captured from cameras, light detection and ranging (LIDAR), etc., are combined with wireless sensory data for wireless network optimization such as channel prediction, HO optimization, etc (Nishio et al., 2020; Tian et al., 2020). This is necessary because mm-wave communication networks possess unique challenges that would be difficult to handle using only wireless sensory data but with the assistance of visual data, some of these challenges can be handled properly. On the other hand, for the wireless data based HO management, the most recent works are extensively reviewed under two use cases: beam selection and BS selection.

#### **2.4.1 An Overview of Machine Learning Algorithms**

It has become very important to include AI/ML in the BS's and beam selection process during HO, in order to achieve the primary objective of providing a seamless HO and to ensure that the UE achieves maximum throughput during the entire duration of its connection to the network. The HO optimization problem is a decision-making problem, and intelligence is imperative to ensure that the optimal decision is taken at each HO instance in a more efficient and effective manner.

We begin by defining ML and discussing the various categories. According to Dey (2016), ML is a set of computation procedures that evolved from formidable techniques in the field of AI that allow the computer to self-learn, discover patterns, and generate models from historical data without being explicitly programmed. The objective of ML is to identify features of a given data set that are likely to influence an outcome of interest given the input, and then use those learned features to predict the result in a new situation not previously encountered (Klaine et al., 2017). A substantial collection of ML techniques (model and algorithms) has been created to solve various challenges in different domains. These algorithms can be classified according to how learning is performed. They have been broadly categorized into three major classes Sultan et al. (2018). Table 5 presents an overview of ML approaches based on their learning styles.

**Definition 6** (Labelled data set)

A labelled data set is a data set with clearly defined features (input) and target (output). The features are usually related to the target and enables the ML algorithm to identify the target or map the input to the output during the training phase. Labelled data sets are used for training supervised learning algorithms.

**Definition 7** (Unlabelled data set)

An unlabelled data set is a data set that does not have labels. That is, there is no clear description of the features or targets in the data set. This kind of data set is used for training unsupervised learning algorithms.

**Definition 8** (Model training)

Model training is the process of exposing an ML algorithm to the training data set (i.e., labelled or unlabelled data set) in order to enable it to learn the mapping between the features and the target. Thereafter, a model is obtained that can correctly predict the right target, even when it is feed with a new data set that it had not previously seen.

**(i) Supervised Learning**

As the name suggests, it is the learning technique which requires a labelled training set consisting of inputs features and output. The learning model tries to search for a function that maps the input to the desired output by minimizing both the bias and variance error of the predicted results. After that, new data set is then applied to the trained model in order to predict the output. Supervised learning is basically classified into regression—where the predicted output

Table 5: Types of Machine Learning Algorithm.

Learning category	Inference task	Learning algorithms
Supervised learning	Classification/Regression	SVM, ANN, extreme gradient boosting (XGBOOST), $k$ NN, Decision tree, Random forest
Unsupervised learning	Clustering/anomaly detection	$k$ -means Principal Component analysis Expectation maximization Independent component analysis
Reinforcement learning	Decision making	$Q$ -learning SARSA Policy gradient Proximal policy optimization Actor-critic
Deep learning	Advanced feature extraction/ data representation	Convolution neural networks Recurrent neural networks Deep belief networks Deep $Q$ -learning Deep deterministic policy gradient

is continuous—, and classification— where the predicted output is discrete or categorical. As regards the classification, two types exist; namely, binary classification (output consists of only two classes) and multi-classification (output comprises more than two classes). Examples of supervised learning algorithms include: artificial neural networks (ANN), support vector machine (SVM), extreme gradient boosting (XGBOOST),  $k$ -nearest neighbour ( $k$ NN), decision tree, random forest, etc, (Fourati et al., 2020).

## **(ii)      Unsupervised Learning**

Different from supervised learning, in unsupervised learning, the training data set is unlabelled. The learning model in this case, tries to find hidden patterns, structures, and correlations within the training data set. They are mainly employed for anomaly detection, pattern recognition, and the reduction of the dimension of a data set. Common examples of unsupervised learning algorithms are  $k$ -means clustering, principal component analysis, expectation-maximization, independent component analysis, etc (Wang et al., 2019a).

## **(iii)      Reinforcement Learning**

Unlike supervised and unsupervised learning that deal with continuous or discrete output prediction and identification of hidden pattern or structures in data, RL is concerned with making decisions in order to obtain an optimal policy in a given environment. It is a trial and error kind of learning whereby an agent interacts with the environments, takes action and gets feedback in terms of reward or penalty, depending on whether the action taken is good or bad for a given objective. The goal of the agent is to maximize the total expected reward (Arulkumaran et al., 2017). The outcome of RL is to learn the optimal policy that would enable the agent to make an optimal decision at any given state of the environment. RL algorithms can be value-based (e.g.  $Q$ -learning, SARSA), policy-based (e.g. policy gradient, proximal policy optimization (PPO) and asynchronous actor-critic (A2C))(Luong et al., 2019).

There is also another branch of ML known as deep learning. These set of algorithms make use of very sophisticated neural network architectures comprising multiple layers that have excellent generalization ability and are capable of more advanced computation, feature extraction, and data representation than classical ML algorithms. Examples of deep learning algorithms are convolution neural networks (CNN), recurrent neural networks (RNN), deep belief networks (DBN), etc. (Chen et al., 2019; Zappone et al., 2019). Deep learning, when combined with RL (deep RL (DRL)) generates a formidable solution to approximate the value function in order to solve the curse of dimensionality problem that is prevalent in conventional RL approaches. Examples of DRL algorithms are deep  $Q$ -learning, double deep  $Q$ -learning, deep deterministic policy gradient (DDPG), etc, (Arulkumaran et al., 2017).

### 2.4.2 Machine Learning based Handover Optimization

HO optimization is necessary when selecting the BS/beam that a user should connect to, in order to minimize frequent HO due to the small footprint of mm-wave BSs in 5G and THz wave BSs that are envisioned to be used in B5G. This is because frequent HOs increase the HO cost, thereby reducing the network throughput. Throughout this work, we will refer to HO as defined by Arshad et al. (2016b) which establish the term HO cost. With efficient HO optimization, the network is able to select the best T-BS that will provide a higher throughput for UE.

Before ML came into play, classical methods for BS selection were based on specific parameter measurement. These methods include selecting the T-BS based on distance, or the BS that provides a higher KPI such as RSRP, received signal strength indicator, and signal-to-noise ratio. In the measurement-based approaches, the channel state information (CSI) from the MR of all neighbouring BSs is measured, and the one with the best CSI is selected as the potential T-BS. These approaches are practical for sub-6 GHz frequencies; however, they are inefficient solutions in mm-wave and THz application band due to severe path loss and susceptibility to LOS blockage (Oguma et al., 2016).

ML techniques can play a significant role in HO optimization and BS station selection by reducing delays, computational overhead, and frequent HOs. They help predict the T-BS and also ensure that adequate resources are available at the T-BS before HO occurs in order to ensure a seamless HO. In this section, we consider ML-based HO management in 5G networks from the perspective of visual, and wireless data aided HO optimization. In Table 6, we present a summary of the state-of-the-art ML-based HO optimization in 5G mm-wave communication systems.

#### (i) Visual data aided handover optimization

Successive generations of cellular networks have mainly depended on wireless sensory information such as CSI, received power, etc., for network design and optimization. However, the use of mm-wave and THz frequencies in 5G and B5G networks would mean that BSs will have many antennas, communication will be through a large number of LOS beams, which would be subject to blockages of various types and would limit signal reception at the user end. In addition, much signaling overhead would be involved in the selection of the optimal beam for

user connection in mm-wave networks if only wireless sensory data are exploited for optimal beam selection considering the massive number of beams that would be involved (Nishio et al., 2020; Tian et al., 2020).

The vision assisted HO optimization has become necessary because of the complexity of the mm-wave networks, and it might not be possible to capture all the conditions of the environments like obstacles, buildings, etc., using wireless sensory data. As a result, detecting or predicting the presence of obstacles that would block the received beam and reduce the throughput at the user end is very difficult to achieve using only wireless sensory data. However, with vision assisted HO optimization, visual data (image/video) is combined with wireless sensory data to enable proactive obstacle prediction and optimized beam/BS selection that would help enhance user QoS (Alrabeiah et al., 2020b). In addition, with the advancement in computer vision, the training overhead that is normally associated with training ML models for optimal beam selection can be greatly reduced by utilizing the images of networks in developing deep learning algorithms for efficient HO operation (Alrabeiah et al., 2020a, 2019). In the following paragraphs, we review the research works that have been proposed on the use of visual data for HO optimization in mm-wave networks.

One application of visual data for HO optimization is the prediction of obstacles that might affect the magnitude of the received power or data rate at the user end. In this regard, the authors Koda et al. (2020b), proposed a cooperative sensing scheme for proactive HO in mm-wave networks using a combination of images captured from multiple cameras and received power. The idea is to map camera images with HO decision using DRL such that a proactive HO decision can be initiated before the received signal is blocked by an obstacle. The advantage of using multiple cameras is to cover areas that are inaccessible by other cameras so as to get a complete view of the network environment. The camera images also enable the prediction of obstacles that will affect the mm-wave links. The authors Koda et al. (2020a) developed a DRL framework using camera images for optimizing the HO timing by predicting the future data rate of mm-wave links and ensuring that proactive HO is performed before data rate degradation occurs due to presence of obstacles.

Another application of visual data for HO optimization is the prediction or selection of the optimal beam that the user should connect based on user mobility and the presence of obstacles in the mm-wave network environment. Following this research direction, the authors in Klautau

et al. (2019) demonstrated how data obtained from LIDAR sensors could be used to reduce the overhead associated with mm-wave beam selection and LOS detection, and proposed a decentralized architecture using deep CNN. Their work was extended by Dias et al. (2019) as they developing a deep learning-based centralized architecture for beam selection and the detection of the LOS in vehicular networks by combining location information and LIDAR data. In (Xu et al., 2020b), the authors proposed a novel beams selection scheme that is capable of predicting the optimal beam to connect to at any position in the cell using image-based 3D reconstruction and CNN. They argue that the proposed method takes images from ordinary cameras and is cheaper to implement compared to LIDAR-based approaches in (Klautau et al., 2019; Dias et al., 2019).

The work by Alrabeiah et al. (2020b) considered the problem of beam selection and blockage prediction using camera images, channel state, and deep learning for a single user communication in a mm-wave network. The beam selection problem was formulated as an image classification problem such that the UEs are mapped to a class of beams having a unique beam index, depending on their location in the image. However, to detect users that are blocked, the images are matched with channel information due to the difficulty of detecting obstacles in still images. The authors Alrabeiah et al. (2020a) first developed a realistic image data set for ML-based mm-wave network optimization that considers many BSs, users, different obstacles, and rich environmental dynamics. Then leveraging the image data set and information regarding previously selected beams, a ML based vision-aided beam tracking framework was proposed to predict the future beams of mobile users in a mm-wave communication system.

## **(ii) Wireless data aided handover optimization**

Non-vision assisted HO optimization, on the other hand, does not involve the uses visual sensory information such as images and videos for HO optimization. It uses the conventional wireless sensory information such as received signal level and channel state, and user location information to optimize the switching of user connection from one BS or beam to another. This is the general technique that is commonly used in wireless communications systems. In this session, we review the state-of-the-art in HO optimization in mm-wave communications networks from the perspective of the BS and beam selection by exploiting the CSI and user mobility information such as user location, trajectory, etc.

**a) Beam selection:** Due to the high path loss and sensitivity to blockage experienced by mm-waves, a large number of BSs comprising multiple directional antenna arrays have to be deployed. The use of multiple antenna arrays enables the formation of narrow signal beams with a high gain when the phase or amplitude of each antenna is adjusted. This approach, commonly known as beamforming (Roh et al., 2014), enables the formation of directional links between the BSs and UEs. However, because each BS comprises multiple beams, the challenge becomes selecting the optimal beam that will serve the UE in order to satisfy its QoS. In the following paragraphs, we review the most recent works on ML-based beam selection in mm-wave and THz communication systems.

The beam selection problem is sometimes modeled as a multi-classification problem, after which a supervised or deep learning algorithm is used to identify the beam class. In this regards, the authors Long et al. (2018) proposed a data-driven approach for analog beam selection in hybrid MIMO systems. The beam selection problem was first formulated as a multi-classification problem and then solved using SVM in order to obtain the optimal analog beam for each user. The performance evaluation shows that the proposed method has similar data rate to that of traditional methods but with lesser complexity. In (Antón-Haro & Mestre, 2019), the direction of arrival information was leveraged to developed a ML scheme for beams selection in mm-wave communications. The beam selection problem was expressed as a multi-class problem, and three supervised learning algorithms namely  $k$ NN, SVM, and ANN were used to solve the problem. The authors Ma et al. (2019), proposed a beam selection policy for THz systems based on ML approach with low complexity. The beam selection problem was first formulated as a multi-classification problem after which a random forest algorithm was used to determine the optimal beam class.

In (Yang et al., 2019) and (Yang et al., 2020), a ML framework for analog beam selection was proposed using SVM, which considers the transmit power of the SCs and channel information as inputs while the model training was performed using sequential minimal optimization in order to achieve high sum-rate at a lower computational complexity. A DNN model for beam selection where channel knowledge is not required was developed by Lin et al. (2019). The beam selection problem was modeled as an image reconstruction problem, after which the DNN was used for interpolation. The proposed model was first trained offline—to reduce the training overhead—before online implementation of the trained model was performed. A beam selection

framework for mm-wave vehicular networks using different ML-based classification models was proposed by Wang et al. (2018). The training data set comprised the vehicle location, type of receiver vehicle and its surrounding vehicles as well previously selected beams. It was observed that the random forest algorithm outperformed other classification algorithms in terms of accuracy and efficiency. A neural network framework for beam selection in THz communication networks was proposed by Li et al. (2019). The proposed model was trained using data samples obtained from the THz channel based on the multi-classification approach. The proposed model was able to determine the optimal beam for each user with low complexity compared to the conventional exhaustive search method.

Another category of mm-wave beam selection technique exploits the CSI of sub-6 GHz to minimize the search overhead involved in selecting the optimal beam as well as for initial beam establishment. In this regards, the authors Jagyasi & Coupechoux (2020) proposed a DNN based framework for selecting the optimal mm-wave SCs and beam in a HetNet involving mm-wave SCs and sub-6 GHz macrocells. They utilized the CSI from sub-6 GHz macrocells for both SCs and beam selection in order to minimize the latency resulting from using conventional exhaustive search approach for beam selection. The authors Sim et al. (2020) introduced a deep learning approach to mm-wave beam selection in 5G and B5G using sub-6 GHz CSI. They argue that using the sub-6 GHz CSI for the mm-wave beam selection would help reduce the search space required for establishing the initial beam. In (Alrabeiah & Alkhateeb, 2020), a deep learning framework was proposed for predicting mm-wave beam and blockage while using sub-6 GHz channel. They proved that under certain conditions, a mapping function exists, that can be used to predict the optimal beam and blockages in any environment. Then they went further to show that this mapping function can be learnt using a large enough neural network after which a DNN model was designed to perform both predictions. The work by Bian et al. (2020), suggested a deep learning approach for the prediction of the optimal mm-wave downlink beam. The developed DNN model takes as input a combination of features extracted from both the sub-6 GHz channel and mm-wave band in order to enhance prediction accuracy and achievable data rate.

RL techniques have also been applied to mm-wave beam selection in literature. In this regard, different (deep) RL algorithms such as multi-armed bandit (MAB),  $Q$ -learning, deep  $Q$ -learning approaches have been proposed. A novel ML-based beam tracking and alignment framework

for a sparse and time-varying mm-wave channel was proposed by Booth et al. (2019). The channel tracking was performed using Bayesian learning and Kalman filtering after which the optimal beam selection strategy was obtained using MAB. A fast ML algorithm for beam selection in 5G mm-wave vehicular networks using contextual MAB (CMAB) was proposed by Asadi et al. (2018) and Sim et al. (2018). The proposed model considers the traffic pattern and different types of blockages in order to select the optimal beam in real-time without prior training of the model. In (Aykin et al., 2020), a beam tracking approach based on MAB is proposed to determine the optimal beams and data rates of the beams in a mm-wave communication system. The proposed model uses the beam quality information, and the feedback obtained from users during initial access to determine the optimal beam and transmission rate during the next transmission.

The authors Li et al. (2020) proposed an online learning algorithm for optimal beam selection in mm-wave vehicular networks using CMAB. The developed algorithm is able to predict the beam direction of the target mm-wave BS from the serving mm-wave BS depending on the current traffic pattern while considering the user QoS requirements. In (Xu et al., 2020a) multi-agent RL (MARL) approach for the joint optimization of user scheduling and beam selection in mm-wave networks was developed. The proposed method ensures that the delays associated with beam selection are minimized while ensuring that the users QoS are satisfied. The authors Chiang et al. (2020) proposed a framework for mm-wave beam prediction in multi-UAV communication systems using  $Q$ -learning. The proposed model exploits the received coupling coefficients (a pair of analog beamforming vector from the transmitter and receiver side of the channel) to determine the optimal beam that will maximize the received signal-to-interference-plus-noise-ratio.

A learning-based approach for optimizing beam search in mm-wave BSs in an indoor network environment while considering user mobility has been proposed by Chen et al. (2018b). The proposed approach uses multi-state  $Q$ -learning while exploiting user trajectory-based data from the radio. They argue that the proposed method is superior to traditional methods because it jointly considers both BS and beam selection, can be adapted to different indoor environments and user mobility and minimizes the delays due to beam search. A beam selection framework for high mobility vehicular networks which aims at enhancing data rate, minimizing the number of HOs and disconnection time was proposed by Van Huynh et al. (2020). The proposed

framework utilizes parallel  $Q$ -learning to determine the optimal beam for each vehicle. The algorithm leverages the possibility of simultaneously collecting information from multiple vehicles on the road to hasten its convergence to the optimal solution. In (Wang et al., 2019b), an RL framework for beam selection in NLOS scenarios was introduced. The proposed framework employs  $Q$ -learning to determine the optimal NLOS beam for each user based on the user's QoS requirement.

The user location can also be exploited in order to identify the optimal beam selection for user association. The authors Rezaie et al. (2020), proposed a beam selection strategy based on ML that considers the user position and receiver orientation to select the optimal beam pair, thereby reducing the overhead associated with beam alignment. Moreover, since their approach to beam selection is based on multi-classification, the neural network model is enriched with a large amount of CSI to enable it not only to select the strongest beam based on the magnitude of received signal but also an alternative beam. This makes the proposed approach resilient against blockages. A hierarchical learning-based beam selection scheme was proposed by Wu & Ai (2019) for multi-users in mm-wave vehicular networks. They developed a graph neural network (GNN) model for beam pair selection while considering CSI and user positions. A deep learning model based on CNN architecture was proposed by Tauqir & Habib (2019) for selecting the beam that gives the best communication performance to users in a massive MIMO system while considering user position. In (Heng & Andrews, 2019), the authors developed a learning-based beam alignment scheme for mm-wave systems that can determine the optimal BS while only exploiting the user position. The proposed scheme can predict the optimal BS and beam even with incomplete user location information with reduced search time. A position-based online learning framework for optimal beam pair selection and refinement was proposed by Va et al. (2019) while considering only user position. The beam selection and refinement problem was first modeled as a continuum-armed bandit problem after which a risk-aware greedy upper confidence bound (UCB) algorithm was developed for beam selection while a hierarchical optimistic optimization was used for beam refinement. The observed that where more information regarding the environment can be obtained from BS or user devices, the training overhead can be further reduced.

**b) Base station selection:** A proactive HO framework that enable users to switch connection to another BS before link disconnection was proposed by Alkhateeb et al. (2018). The proposed

method uses deep learning to predict obstacles and trigger HO before link disconnection occurs, thereby ensuring the reliability of the link and preventing data transmission delays due to link disconnection. In (Sun et al., 2019a) and (Sun et al., 2019a), a HO mechanism for selecting the optimal BS in mm-wave network based on MAB approach was developed in order to ensure the user has a longer connection time with the BS after HO. The considered the user's post-HO trajectory, and the blockages along the LOS to predict future HO. A RL framework for minimizing frequent HO while satisfying users QoS was proposed by Sun et al. (2018) using MAB. The proposed framework takes into account the channel conditions and user QoS requirements before triggering HO. Furthermore, two BS selection algorithms were also developed based on user density for both single-user and multiply user HO scenarios, respectively.

An intelligent HO decision framework for BS selection was proposed by Mollet et al. (2020b). The proposed framework uses a DDRL algorithm to learn the optimal BS for user association in order to minimize the number of HOs and optimize the average throughput along the user trajectory. A distributed learning framework for HO optimization in dense mm-wave networks was proposed by Sana et al. (2020a) and Sana et al. (2020b) in order to minimize frequent HO and optimize user throughput. The framework employs MARL where each user was modeled as an agent and takes an independent HO decision based on its local observation, thereby reducing signaling overhead. The authors Yajnanarayana et al. (2020) introduced a HO optimization algorithm based on RL for 5G systems. They modeled the HO problem as a CMAB, then developed a  $Q$ -learning solution. In (Klus et al., 2020), the authors proposed a deep learning model for user localization and proactive HO management, while considering user behaviour in the network. The proposed model uses the received signal measurements to reduce the number of unnecessary HOs and predict the user location while ensuring that the throughput of the network is maintained.

A joint optimization framework for minimizing HO frequency and maximizing user throughput was proposed by Guo et al. (2020). The HO and power allocation problem was modeled as a cooperative multi-agent task, after which a MARL framework using proximal policy optimization (PPO) was developed. The model training was performed in a centralized manner after which decentralized policies were obtained for each user. The authors Khosravi et al. (2020a), proposed a learning framework that jointly optimizes HO and beamforming for mm-wave networks. RL algorithm was employed to determine the optimal backup BSs along user trajectory

that will help reduce the overhead signaling during channel estimation for user association and minimize the number of HOs. This would ensure an enhanced data rate along the user trajectory. A learning-based load balancing HO mechanism was proposed by Khosravi et al. (2020b). The user association problem was modeled as a non-convex optimization problem, after which a deep deterministic policy gradient (DDPG) RL algorithm was applied to solve the optimization problem. The algorithm's goal is to associate all the users in different trajectories in the network environment to the optimal BSs in such a way that maximizes their sum rate as well as reduces the number of HO occurrences.

Table 6: Summary of the State-of-the-art ML-based HO Optimization in 5G mm-wave Communication Systems.

Paper	Visual data	Wireless data		ML algorithms
		Beam Selection	BS Selection	
(Koda et al., 2020b), (Koda et al., 2020a)	✓			DRL
(Klautau et al., 2019)	✓			CNN
(Dias et al., 2019)	✓			DNN
(Xu et al., 2020b)	✓			CNN
(Alrabeiah et al., 2020b)	✓			DNN
(Long et al., 2018)		✓		SVM
(Antón-Haro & Mestre, 2019)		✓		$k$ NN,SVM, ANN
(Ma et al., 2019)		✓		Random Forest
(Yang et al., 2020)		✓		SVM
(Wang et al., 2018)		✓		Random Forest
(Li et al., 2019)		✓		ANN
(Jagyasi & Coupechoux, 2020)		✓		DNN
(Booth et al., 2019)		✓		MAB
(Li et al., 2020)		✓		CMAB
(Xu et al., 2020a)		✓		MARL
(Chiang et al., 2020)		✓		$Q$ -learning
(Chen et al., 2018b)		✓		$Q$ -learning
(Van Huynh et al., 2020)		✓		$Q$ -learning
(Alkhateeb et al., 2018)			✓	DNN
(Sun et al., 2019a, 2020)			✓	CMAB
(Sun et al., 2018)			✓	MAB
(Mollet et al., 2020a)			✓	DDRL

## CHAPTER THREE

### MATERIAL AND METHODS

#### 3.1 System Models Implementation

This study relies on simulations to generate the data for both the model's training and evaluation. At some point, we used third-party software such as Wireless Insite, and for the rest of the work, we develop the environment either from Matlab or Python. The environment is based on a discrete-time event simulator, and the parameters and models used in this thesis are presented as follows.

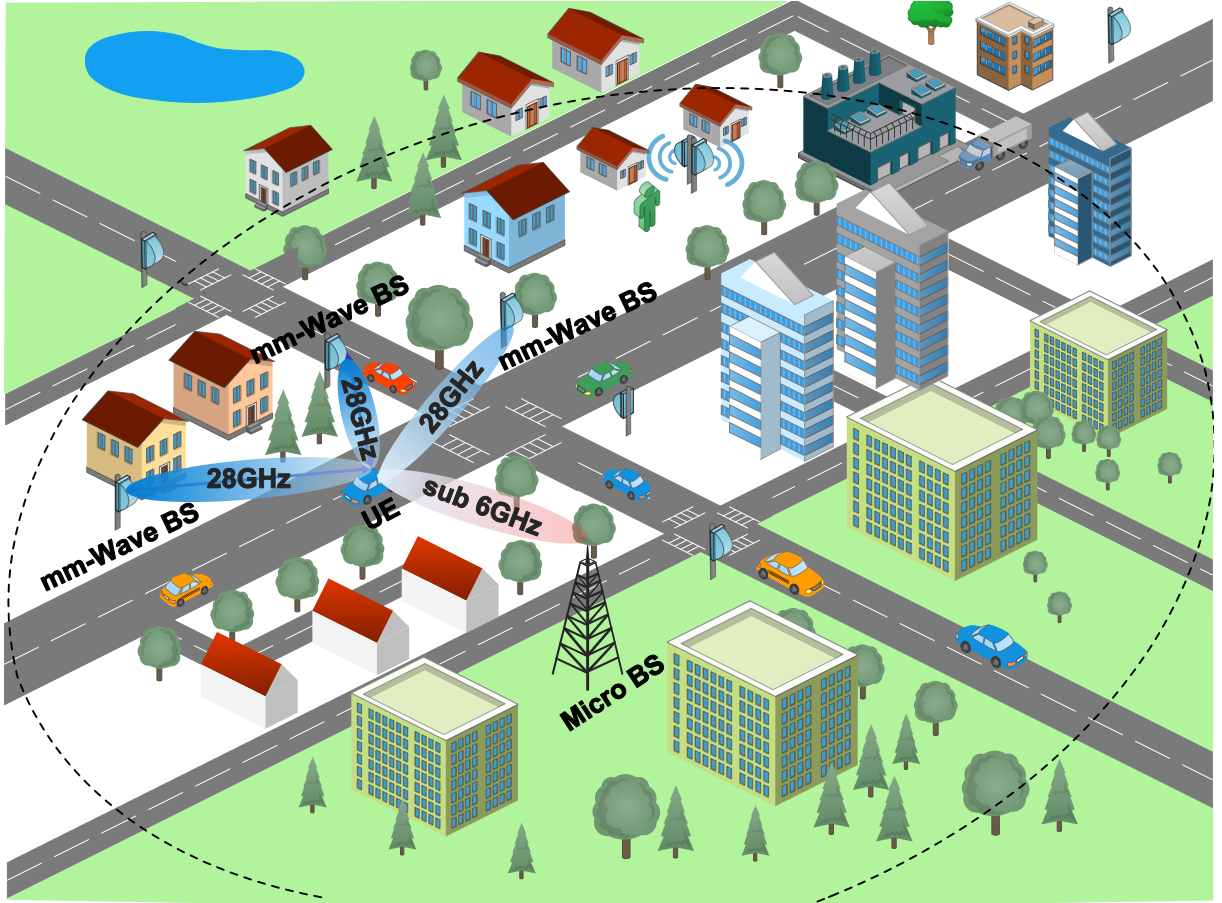


Figure 15: The system model of mm-wave UDN.

##### 3.1.1 5G Heterogeneous Network

We consider Fig. 15 as use-case for system model, which demonstrates a simplified 5G HetNet wireless network system with mm-Wave SCs placed close to one another. A cellular network

environment comprises a large number of mm-wave SCs and UEs, as shown in Fig. 15. The network considers the presence of LOS signals, NLOS signals, blockages, and building reflectors. The aim is to depict a natural urban environment with entirely distinct obstacles. Each mm-wave SC is equipped with  $X$  antennas; all SCs are assumed to be connected to a BS central controller (BSc). Multiple overlapping mm-wave SCs are randomly distributed in the network to provide high throughput via LOS links. Additionally, macro BSs (MBSs) exist in the network. They serve two purposes; firstly, to ensure reliable communication whenever there is no LOS link between the mm-wave SCs and the UE, and secondly, to facilitate the transmission of control signals to the BSc that acts as decision node in the network. Each mobile UE is assumed to be equipped with a single antenna.

### 3.1.2 Channel Model

The ray-tracing model is used to model the wireless channel in this study. The ray-tracing model is based on the superposition principle whereby all the reflected and LOS waves generated by a transmitter are aggregated between the transmitter and the receiver. More specifically, a geometric wide band mm-wave channel model (Rappaport et al., 2013) with  $N$  clusters is adopted, where each cluster  $n \in N$  is assumed to produce a single ray with a finite time delay  $\tau_n \in \mathbb{R}$  and angle of arrival (AoA) for elevation/azimuth  $\phi_n, \theta_n$ . We assume that each user has a single antenna and that the path loss between the UE and  $m^{th}$  BS is  $\lambda_m$ . The time delay channel between the  $m^{th}$  BS and the UE,  $s_{d,m}$ , can be expressed as:

$$s_{d,m} = \sqrt{\frac{X}{\lambda_m}} \sum_{n=1}^N \alpha_n p(dT_p - \tau_n) a_m(\theta_n, \phi_n), \quad (1)$$

where  $X$  symbolizes the number of antennas in the BS,  $a_m(\theta_n, \phi_n)$  represents the array response vector of the  $m^{th}$  BS at AoA  $(\theta_n, \phi_n)$ , and  $p(dT_r - \tau_n)$  denotes the pulse shaping function of the spacing signalling,  $T_p$  obtained at  $\tau$  seconds (Schniter & Sayeed, 2014). From the time delay channel in (1), the subcarrier,  $k$  and the frequency-domain channel model,  $s_{k,n}$  can be expressed as:

$$\mathbf{s}_{k,m} = \sum_{d=0}^{D-1} \mathbf{s}_{d,m} e^{-j \frac{2\pi k}{K} d}. \quad (2)$$

It is assumed that the block-fading channel model,  $\{s_{k,m}\}_{k=1}^K$  is constant within the coherence time of the channel,  $T_c$ , (Rappaport et al., 2013) which is dependent on the velocity of the user as well as the multi-path components of the channel.

### 3.1.3 Beamforming Model

We assume that the BSs are equipped with highly directional antennas with sectorized gain patterns. Since the BS uses high-frequency mm-waves, it is easy to exploit beam-forming techniques. The antenna gain from the main lobe is  $\Omega$  while that of the side-lobes is  $\omega$ . The directional gain of the antenna is a function of  $\theta$  as well as the steering angle and is given as:

$$G(\theta) = \begin{cases} \Omega & \text{if } |\theta| \leq \theta_b \\ \omega & \text{otherwise,} \end{cases} \quad (3)$$

with  $\theta_b$  denoting the beam-width of the antenna and  $\theta$  the angle between the UE and the mm-wave BS. Highly sophisticated beam-tracking is deployed in WI<sup>TM</sup> to ensure that the definitive association between the UE and mm-wave BS is successful. Therefore, the UE is always in the main lobe, with main lobe gain connected in a given antenna, and UE experiences no interference or any signals from other antennas.

### 3.1.4 SINR Model

One of the key metrics for measuring the performance of the wireless communication channels is the SINR, which can be expressed as:

$$SINR = \frac{S}{I + \sigma_u}, \quad (4)$$

where  $S$  represents the received signal power,  $\sigma_u$  denotes the noise component, and  $I$  denotes the interference power from all surrounding BSs except the S-BS. Equation 4 can be further expressed as

$$SINR = \frac{P_s \Omega_s L_s}{\sigma_u + \sum_{i=1}^M P_{(i=t \cap t \neq s)} \omega_i L_i}, \quad (5)$$

where  $P_t$  and  $P_s$  represent the power transmitted by surrounding and S-BSs respectively,  $\omega_i = \omega_t$  and  $\Omega_s$  depict the antenna gain of surrounding and S-BSs, and  $L_i = L_t$  and  $L_s$  is the path loss gain of the surrounding and S-BSs respectively. Because mm-wave antennas form directional beams (Eqn. 3), the contribution of inter-cell interference can be assumed to be negligible. Hence, SNR is considered in this work for the mm-wave channel while SINR is used for the sub-6 GHz channel.

### 3.1.5 Radio Link Failure Model

The term "*Radio Link Failure*" (RLF) refers to a state in which a wireless connection between a UE and a BS is considered lost. A connection is considered to be in RLF if the received

SNR ( $\gamma_r(t)$ ) is below a threshold  $\gamma_{th}$  for a certain period, denoted  $T_{RLF}$ . If by any chance the signal  $\gamma_r(t)$  exceeds the threshold  $\gamma_{th}$  during that time period— $T_{RLF}$ , the counter is reset and the connection is considered good. If the signal remains below  $\gamma_{th}$  after  $T_{RLF}$ , then communication is considered to be in RLF and the connection is dropped. As a result, the user has to initiate a new connection procedure—HO in this case—which can be time-consuming. *Time consumption during HO or reconnection time ( $t_d$ )* is defined as the minimum time required by the UE to establish a connection following a connection drop. Mathematically, RLF can be written as:

$$\gamma_r(t) < \gamma_{th}; \forall t \in [t_0 - T_{RLF}, t_0], \quad (6)$$

where  $\gamma_{th}$  denotes the minimum value, and  $T_{RLF}$  denotes the time in which the link is broken.

Table 7: Radio link failure parameters.

Parameters	Value
$\gamma_{th}$	15-25 [dB]
$T_{RLF}$	50 ms
TTT	{80, 100, 128, 160, 256, 320, 480, 512, 640, 1024} ms
Reconnection time $t_d$	0.7, 1, 2, 3 seconds

An additional vital parameter to note is the *time-to-trigger* (TTT) because it directly affects the RLF (3GPP, 2018b). TTT is defined as the length of time for which the RSRP, or  $\gamma_s$  of the S-BS is lower than that of the T-BS. The value set for the TTT can either increase or decrease the RLF. This is because a higher value of TTT might lead to a connection drop between the UE and BS before HO is triggered and hence RLF. In contrast, a lower value of TTT could lead to multiple unnecessary HOs. Both of these affect the QoS and QoE of the UE. Generally, this timer helps to avoid irregular measurements and HOs. Also, HO margin or hysteresis factor, which is defined as the minimum RSRP difference between the serving and T-BSs needed to start TTT count, is ignored in this work. The values used in this model are summarized in Table. 7.

### 3.1.6 User and Traffic Models

A user is represented by its equipment, called a user equipment, UE. User mobility is defined to follow a random waypoint model and bounce in a square area to ensure that all UEs stay

within the simulated area. When the user reaches the edge of the area, it merely "bounces" back towards the simulated region at a random angle.

While the total number of UEs remains constant in the simulation, two distinct classes of UEs are considered: random UEs and interested UEs. In the assisted UE-network HO strategy, interested UEs are considered as agents or systems used for learning purposes, while random UEs are used to generate random traffic within the simulated environment. UEs are considered to connect to the BSs directly and are in the active state unless stated otherwise. As explained before, a UE can also be in RLF if the received signal is too weak.

Varying size packets—headers included— are considered: 2 Mbps, 10 Mbps, and 100 Mbps. The packet size distribution follows Poisson distribution, which means there is a high chance of generating small packets than large ones. Two consecutive packets have completely distinct distributions. The inter-arrival time is non-periodic and also modelled following an exponential distribution. Three seconds is set as the mean time between two packets. It is important to note that the inter-arrival time distribution is unrelated to the packet size between two successive packets. The File Transfer Protocol (FTP) is chosen to simulate data transfer, with the minor modification of changing the packet size to realistically mimic various UE demands. The complete summary of the traffic model is provided in Table. 8.

Table 8: Traffic model parameters.

Parameters	Value
Traffic Type	FTP
Size Distribution	Exponential
File Size	2, 20, 100 Mbps
Size Probability	1/2, 3/10, 2/10
Mean Inter-arrival Time	Exponential
Inter-arrival Time Distribution	3 secs

### 3.2 Handover as Combinatorial Optimization Problems

Combinatorial optimization problems require the finding of an "optimal action" from a set of finite actions. The optimal action is quantified according to the evaluation function that maps

that action to either a reward or a penalty, and the objective is to select the action with maximum reward or minimum penalty. Most practically interesting combinatorial optimization problems (COPs from now on) are also very hard, in the sense that the number of objects in the set increases extremely fast due to even small increases in the problem size, making exhaustive search impractical. HO can be modelled as a combinatorial problem. This is because a UE that moves in the given trajectory normally experiences link failure, and the choice of the best S-BS among M potential T-BS is crucial to achieving a high system throughput. The reward here is the throughput that the UE can experience throughout the trajectory while minimizing unnecessary HOs.

This section presents the general conditions that trigger HO and essential criteria for choosing a S-BS. We also present the reward and cost functions that can be used as trade-offs for making decisions. The aim is to avoid a short-sighted decision by selecting a BS that temporarily offers a higher SNR but in which the UE-BS connection is lost a few seconds after HO. The findings from Mollel et al. (2019) show that HO event criteria alone from 3GPP (2018b) lead to unnecessary HOs and results in a lower system throughput.

### 3.2.1 Conditions for Initiating Handover based on 3GPP

According to the 3GPP (2018b), six events are defined for initiating conventional HOs, where events A2 and A3 are specified for intra-radio access technology (intra-RAT) HO. In contrast, inter-RAT HO is described in event B2. These are the criteria for initiating and terminating HOs:

Event A2 is entered when the S-BS becomes worse than the stipulated threshold in terms of RSRP or SNR value; the opposite is true for the leaving state i.e., event A2 is exited for RSRP or SNR values higher than the specified threshold. The authors Mollel et al. (2019) used the SINR value as threshold value and analyzed its effect on HO in mm-wave networks. Their study concluded that different use cases (service-aware) could have different HO rates in the same trajectory. In this study, we adopt SNR as the triggering criteria and define event A2 as:

$$\begin{cases} \gamma_s < \gamma_{th} & \text{initiate HO} \\ \gamma_s \geq \gamma_{th} & \text{otherwise,} \end{cases} \quad (7)$$

where  $\gamma_s$  is the SNR from the S-BS and  $\gamma_{th}$  is minimum SNR required by the UE to maintain a connection based on the service type. We remove the hysteresis margin since the parameter

itself needs optimization. However, removing the hysteresis margin might lead to an increase in the number of ping pong and other unnecessary HO, but the proposed model ensures that HOs are reduced without a hysteresis margin. Details regarding the working principles of the proposed model will be explained in a later section. Therefore, if the condition for HO is met, the HO process commences and the UE needs to select a potential T-BS.

The condition for entering event A3 is defined when the neighbour BS offers an SNR value greater than the HO margin compared to the S-BS. Regardless of how much power the UE receives from the S-BS, the event looks for the offset value between the serving and neighbour BSs, and if the condition is met, the HO process is initiated. In the proposed solution, we neglect this event as it can sometimes lead to unnecessary HOs. Finally, the condition for event B2 is the same as that of event A2 except that it involves inter-RAT HOs while A2 deals with intra-RAT HOs.

### 3.2.2 Handover Cost

It is essential to have an objective function which is necessary for the optimization problem. Here, we introduce the objective constraint, which is called the HO cost. It should be recalled that this thesis is based on 5G networks, and unlike LTE networks which use soft HO, 5G uses hard HO. The authors Christensen & Knappe (2016) provide further details regarding the complexity and the waste of resources in soft and softer HO. Therefore, this study focuses on hard HOs unless stated otherwise. Because hard HO is considered in this study, the optimal BS selection during HO becomes more critical since the connection is broken before a new connection is established.

After the HO process is initialized in Eqn. 7 and shown in Fig. 14, it takes time to complete the HO process for the UE to switch connection from the S-BS to the T-BS. During the HO process, nothing is transmitted between the UE and either the S-BS or T-BS. The time spent to complete a successful event, where the UE is switched from the S-BS to the T-BS without data transmission, is known as HO delay time  $t_d$ . The accumulation of  $t_d$  has a significant effect on the average throughput of the UE. The cumulative  $t_d$  and total number of HOs in a given trajectory contributes to HO cost ( $\beta_c$ ). HO cost is a function of the number of HOs and the HO delay time, and is expressed as presented by Arshad et al. (2016)

$$\beta_c = \min(\mathcal{H}_l \times v \times t_d, 1), \quad (8)$$

where  $\mathcal{H}_l$  is the total number of HOs per unit length (m),  $v$  is the velocity in ( $\text{ms}^{-1}$ ), and  $t_d$  is the HO time delay in (sec).

The factor  $\beta_c$  is evaluated as the total time wasted without useful data transmission due to HO operations such as signalling and radio link switching between S-BS and T-BSs. The network performance becomes zero if  $(\mathcal{H}_l \times v \times t_d) \geq 1$  because the UE spends the entire time transmitting HO signalling. The importance of  $\beta_c$  is observed in the average throughput equation, which is derived as follows from the Shannon capacity formula:

$$\mathcal{T} = B \times \log_2(1 + \gamma) \times (1 - \beta_c), \quad (9)$$

where  $B$  is the overall bandwidth,  $\gamma$  is the average experienced SNR, and  $\mathcal{T}$  is the system throughput. Hence, to obtain a high overall system throughput in a given trajectory, the factor  $\beta_c$  and  $\gamma$  play a vital role if  $B$  remains constant. Therefore, our objective is to maximize the system throughput and avoid unnecessary HOs by intelligently selecting BSs with longer duration of unobstructed LOS links. At the same time, to reduce redundant HOs, UE occasionally sacrifice a connection to a BS with the highest SNR that would potentially result in HO after few seconds. Yet, the average SNR must be above the threshold value to maintain the QoS of the user.

### 3.2.3 Trajectory and Service-aware Handover

Equation 9 has two parameters that can be varied as a trade-off to achieve the maximum average throughput. In this study, we introduce a service-aware concept whereby the UE maintains a connection to the S-BS if the  $\gamma_s$  experienced is above or equal to the threshold  $\gamma_{th}$ . This is to guarantee the QoS of the UE, and we call it a service-aware strategy which also agrees with event A2. That means, in the HO event, the UE does not need to select the T-BS that has the highest  $\gamma_s$ ; instead, the UE can select any BS with  $\gamma_s$  above  $\gamma_{th}$ . In addition, we introduce the trajectory-aware strategy, meaning the complete path is known to the UE. By knowing the trajectory of the UE and service type, the UE can carefully and intelligently select a BS with  $\gamma_s$  above  $\gamma_{th}$  that guarantees long connection duration, and this far-sighted view helps to minimize the possibility of incurring multiple HOs by selecting the optimal BSs. Nevertheless, the combined strategy comes at the cost of sometimes sacrificing connection to a BS that provides maximum SNR during HO (ignoring current instantaneous SNR) in favour of a BS that can maintain extended connectivity along the UE trajectory.

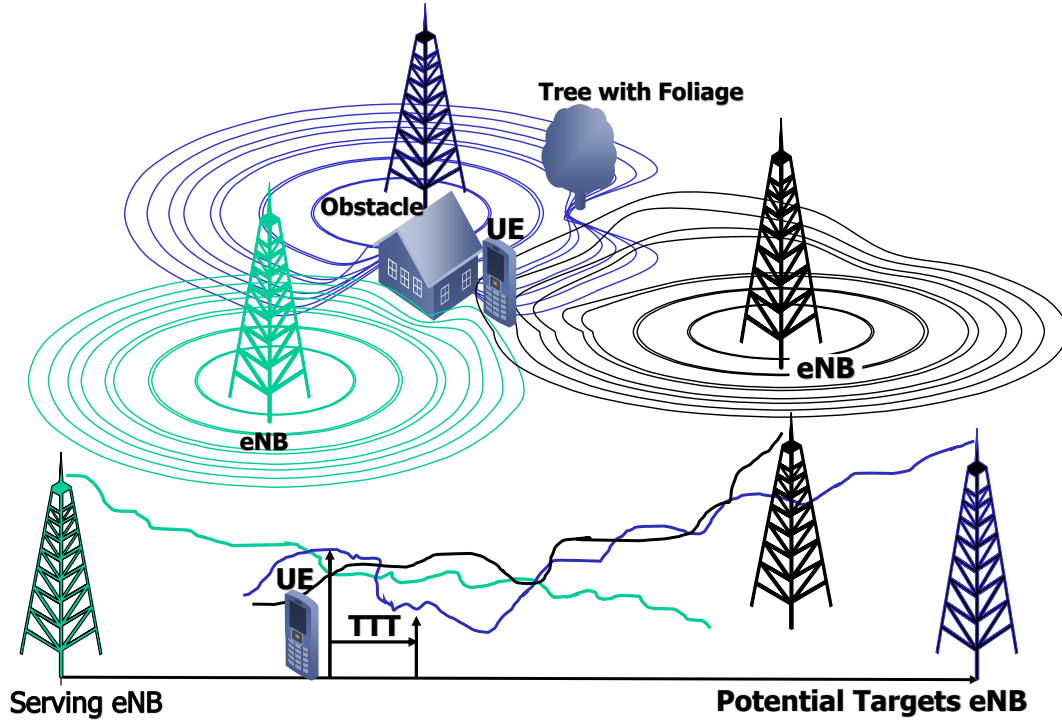


Figure 16: HO problem in the overlapping BSs coverage area

### 3.3 Benchmarks Solutions

In legacy networks, selecting a BS during a HO event mainly depended on using a heuristic approach known as Rate-based HO (RBH). Throughout this thesis, this method is considered one of the benchmark solutions. RBH relies on threshold values such as SNR or RSRP to select a BS when a UE needs to perform a HO operation. The reader is referred to (3GPP, 2018b) for further details regarding the implementation of RBH, as well as Section 3.2.1 for the HO event. Two additional benchmark schemes are considered in this thesis: a) smart HO policy (SHP), proposed by Sun et al. (2018) and b) the developed HO management approach for dense networks based on sparse network information, as shown in Section 3.3.1.

#### 3.3.1 Selection of Target HO BS in Sparse 5G Networks based on Clustering of UDNs

In this proposed scheme, we address the HO problem for a UE moving in a crowded area with overlapping BSs. The solution is based on clustering a dense network of BSs into sparse groups. Since clustering is a widely discussed algorithm that can be found in various literature, this section focuses on a concise discussion of the implementation, as there is a wealth of literature on the topic.

### (i) Problem Formulation

Figure. 16 shows the effect of BSs densification within an area of interest. A UE moving within the area is expected to be connected to one of the BSs. However, due to reflection, multiple obstacles, and the nature of mm-waves in dense mm-wave networks, the HO rate becomes higher than when the UE moves in a sparse mm-wave network (few BSs). The preceding statement is only valid if the T-BS is not chosen intelligently. However, proper BS selection results in lower HO events. The disadvantage of sparse networks is a considerably high occurrence of RLF and outage probability due to insufficient coverage compared to dense mm-wave networks which provide reliable coverage (Ge et al., 2016) for a given trajectory, given that both BSs in sparse and dense deployments have the same transmission power. Besides, LOS and NLOS signals and reflection increase signal fluctuations in dense mm-Wave networks. The UE might receive excellent signals from a distant cell due to having a LOS channel with the cell, while the nearest BS might not provide a high signal strength due to blockage by obstacles. This results in signal fluctuations and with poor TTT configurations, increase the HO and RLF rates.

### (ii) Solution based on Clustering

We assumed that for a given UE trajectory or area of interest where a UE is roaming, there are a number of cells that can serve the UE with minimum HO rate and provide it a satisfactory QoS. We refer to this network of a few BSs that can serve a UE with minimum HO as a sparse network. For both the sparse and dense cases, the cells are deployed and positioned at specific locations to improve the minimum received SNR in the area of interest (Li et al., 2016). During simulation, an average minimum SNR of 0 dB and 3 dB are assigned, following the specification for sparse and dense networks, respectively.

**a) Clustering Dense Networks into Sparse Networks:** We adopted the  $k$ -means algorithm to partition the dense network into sparse networks. The  $k$ -means algorithm partitions an unlabelled multidimensional data set into a set of  $k$  clusters,  $C_i = \{C_1, C_2, \dots, C_{k_i}\}$ , where the desired number of clusters,  $k$  also corresponds to the number of sparse networks generated. The objective here is to partition the network into a set of non-overlapping clusters so as to minimise the intra-cluster distance between the nodes (cells) and the centroid of the cluster while

maximising the inter-cluster distance (Omeke et al., 2021)

$$J_{\min} = \sum_{j=1}^k \sum_{c_n \in C_i} \|x_i - \mu_j\|^2, \quad (10)$$

where each cluster,  $C_i$  contains  $n_j$  cells,  $x_i$  represents the  $i$ -th cell in the network;  $\mu_j$  represents the geometric centroid of the cells in a given cluster (Omeke et al., 2021) and is given by:

$$\mu_j = \left( \frac{1}{N} \sum_{i=1}^N x_i, \frac{1}{N} \sum_{i=1}^N y_i \right). \quad (11)$$

The elbow approach is used to determine the optimal number of clusters,  $k$ . This is done by determining the intra-cluster sum of squares (ISS) for the data, where

$$\text{ISS} = \sum_{i=1}^n (x_i - c_i)^2. \quad (12)$$

The above approach was adopted to cluster the BSs in the dense network into a sparse network. After clustering, a number of points are sampled at regular intervals along the trajectory of the UE. The metric for choosing the sampling frequency could be based on time or distance (e.g. sample every two seconds or 5 meters) – similar results are obtained in both cases. The distance between each BS and each sampling point is measured, and the results summed to obtain an average sum BS-UE distance for each BS. The result obtained is then ranked in ascending order for each cluster and stored in a table. The table is propagated to all the clusters and used for selecting a T-BS.

**b) Selecting a target BS:** To select the best BS for HO, it is assumed that the UE trajectory is known to the network. Two types of BSs are generated, which are hereby referred to as normal and pillar BSs. A pillar BS is one that can maintain connectivity with the UE for a long duration. In a sparse network, pillar BSs are obtained by selecting BSs with minimum sum distances to the UE trajectory. This corresponds to BSs at the top of the ranked tables in the previous subsection. The aforementioned table contains a list of the BSs in each cluster, ranked in ascending order of their proximity to the UE trajectory (i.e., in order of increasing sum distances). The first (top of the table) BS in each table is known as the first-level pillar BS. The pillar level of BSs decreases as it moves away from the first level pillar BS down the table. To HO, the UE selects the BS with the minimum sum distance from the list of first-level pillar BS. The reason for doing this is that BSs with lower average sum distances to the UE have a higher probability of providing a higher SNR and QoS than BSs which have higher average

sum distances to the UE. Where two BSs have the same average sum distance from the UE, the one in the direction of travel is chosen since it is likely to provide good coverage over a longer interval. If the two BSs with the same average sum distance are both in the direction of travel, one is selected at random. For the same reason (based on direction of travel), the proposed scheme avoids selecting a BS that the UE had connected to before, since the UE is moving away from it and will likely experience RLF if it hands over to such BSs.

This model employs the conventional heuristic method to determine whether a UE needs to initiate HO based on Eqn. 7. However, unlike the legacy method which uses constant TTT (ref Section 3.1.5) and always selects the BS with the highest SNR or RSRP as the T-BS, this method uses a different value of TTT based on the BS level in the pillar-normal BS table. The assignment of TTT is an optimization problem, and for the sake of simplicity, this work assigns its values based on a previously conducted experiment. The values are assigned as follows: for BS-level 1, TTT = 1024 ms; BS-level 2, TTT = 640 ms, decreasing progressively until BS-level 9 where TTT = 100 ms, following Table 7. If there are still more BS in the table (more than 9), the remainder are all assigned a TTT value of 80 ms.

To select a T-BS, the developed scheme prioritizes BSs in the top-level of the pillar-normal BS table that satisfy the condition  $\gamma_s \geq \gamma_{th}$ . That is, instead of searching through the entire list of qualifying BSs, this approach prioritises BSs that are in close proximity to the UE trajectory. This has several advantages: it lowers the time taken to find candidate T-BSs for HO, reduces energy consumed in sampling candidate BSs since only a small subset of BSs in the sparse networks are prioritised as well as lowers the required computational resources since the search space is smaller. In addition, it guarantees that the selected BS will provide an acceptable level of QoS throughout the UE trajectory and eliminates frequent HOs or the ping-pong effect.

### 3.4 Proposed Solutions

Our problem mainly focuses on selecting a BS that can maintain more extended connectivity along the user trajectory during HO event. This section presents a novel solution for the HO decision optimization. The proposed solution is based on RL, and subsection 3.4.1 presents a general introduction to RL as well as the problem formulation while the learning algorithms are presented in the subsequent subsections.

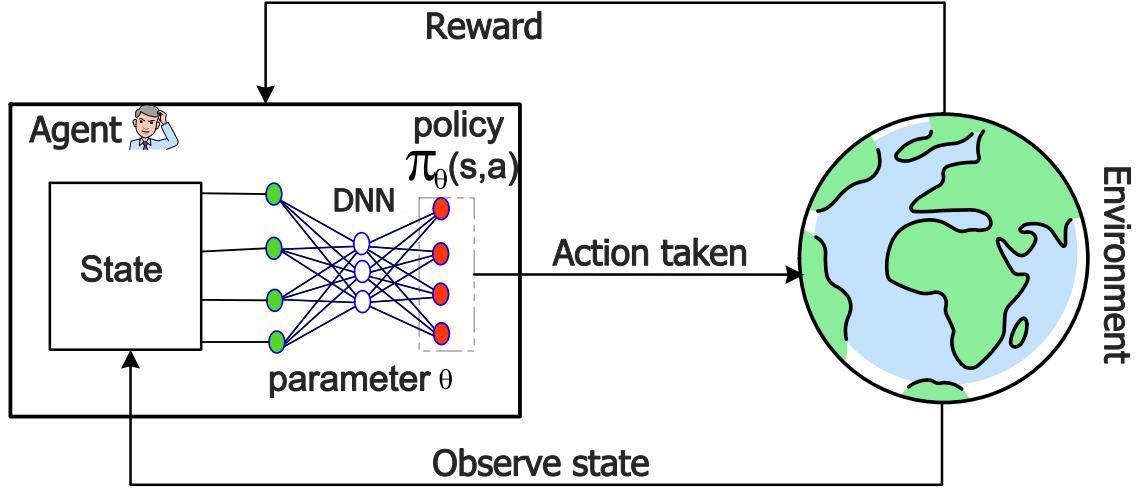


Figure 17: Overview of generic RL algorithm

### 3.4.1 Proposed Solution for Handover Management using RL

The problem of selecting a T-BS is modelled as a combinatorial problem, as discussed in Section 3.2. Whenever a HO event is initiated, there is potentially more than one possible optimal S-BS that the UE can choose to establish a connection with to guarantee reduced chances of the UE entering into a HO event again, i.e., to prevent frequent HOs.

#### (i) Reinforcement Learning (RL) Framework

Learning in RL takes place through a series of interactions between an agent and an environment, with a reward given for each interaction according to how it performs against a preset metric called the reward function. The authors Sutton & Barto (2018) have explained the relationship between an agent, action (interaction) and the environment. They further illustrated clearly and concisely how an agent learns the best policy through multiple interactions with the environment, as shown in Fig. 17. Here, we first define the main elements of RL. At a time  $t$ , the agent observes the state of the environment,  $s_t \in S$ , where  $S$  is a set of possible states. After observing the state  $s_t$ , the agent takes an action,  $a_t \in A(s_t)$  where  $A(s_t)$  is the set of possible actions at state  $s_t$ . Subsequent to the action  $a_t$  selected from state  $s_t$ , the agent receives an immediate reward  $r_{t+1}$  from the state-action pair  $(s_t, a_t)$ . The selected action in state  $s_t$  moves the agent to a new state  $s_{t+1}$  at time  $t + 1$ . It is important for the environment to have state transition dynamics such that  $P(s_{t+1}|s_t, a_t)$  exists, also time  $(t)$  is the arbitrary successive stages

of decision making and acting, which represent the situation and not the interval of real time in seconds.

The strategy for selecting an action by an agent in a given state is known as a policy  $\pi$ . As shown in Fig. 17, the agent learns the optimal policy  $\pi_*$  by first observing the current state  $s_t$ , and then taking action  $a_t$  following the current policy  $\pi$ . As a result, the environment state changes from  $s_t$  to  $s_{t+1}$  and the agent gets an immediate reward  $r(s_t, a_t)$ . The agent repeatedly updates policy  $\pi$  until it reaches an optimal policy  $\pi_*$ . The agent's goal is to achieve the optimal policy  $\pi_*$  by maximizing the cumulative reward. Mathematically, the cumulative reward is the sum of immediate reward and future rewards (Sutton & Barto, 2018) and is given as:

$$\mathbf{G}_t \triangleq R_t + \zeta R_{t+1} + \zeta^2 R_{t+2} + \dots = \sum_{k=0}^{\infty} \zeta^k R_{t+k}, \quad (13)$$

where  $R_t$  is the immediate reward per episode and  $k$  is the total number of episodes the agent must navigate to obtain a full understanding of the environment.  $\zeta \in (0, 1]$  is a discount factor for weighting future rewards; its purpose is to make the sum of the rewards finite. The episode is the complete sequence of states visited by an agent from the initial state to the terminal state. With sufficient experience through episodes iterations, the agent can learn an optimal decision policy  $\pi_*$  that would maximize the long-term accumulated reward.

## (ii) Problem Formulation

A significant number of BSs in the mm-wave environment are required to serve UE. However, the presence of obstacles and the characteristics of mm-wave channels increase the probability of having event A2. Consequently, when event A2 occurs, the UE is required to HO to maintain the connection. The BS that results in minimum HO cost is selected, as shown in Eqn. 8. Therefore, our proposed solution ensures that once event A2 is initiated, the UE switches to the BS with longer unobstructed time for its LOS connection or intelligently skips the HO while ensuring the maximum throughput along the UE trajectory.

Our problem of selecting the best BS during HO that can minimize the number of HO and maximize the average throughput falls into the category of model-free learning where an agent discovers its environment by trial-and error (Rummery & Niranjan, 1994). One of the most common RL algorithms in the model-free method is  $Q$ -learning which is an off-policy algorithm. In  $Q$ -learning, an action-value function  $Q(s, a)$  is defined as the long-term reward and

is given as:

$$\begin{aligned}
Q_\pi(s, a) &\triangleq \mathbf{E}_\pi [G_t | s_t = s, a_t = a] \\
&\triangleq \mathbf{E} [R_t + \zeta G_{t+1} | s_t = s, a_t = a] \\
&\triangleq \sum_{r,s} \left[ r + \gamma \sum_{a'} G_{t+1} | s_t = s, a_t = a \right],
\end{aligned} \tag{14}$$

where the optimal action-value function,  $Q(s, a) \triangleq \max_\pi Q_\pi(s, a)$ , obeys the Bellman optimality equation as shown in equation 15.

$$Q^*(s, a) = \mathbf{E}_{s'} \left[ r_{t+1} + \zeta \max_a Q^*(s', a') | s, a \right] \tag{15}$$

The  $Q$ -learning algorithm updates the corresponding element in the  $Q$ -Table episodically according to the equation,

$$Q(s, a) = (1 - \alpha)Q(s, a) + \alpha(R_{t+1} + \zeta \max_a Q_\pi(s', A)) \tag{16}$$

where  $\alpha$  is the learning rate, and  $s'$  is the next state after the agent follows a policy  $\pi$  in state  $s$ . Based on the number of decision points on a UE's trajectory that the UE might pass and decide whether to HO due to link failure, it becomes too difficult to individually learn all the action-values in all the states for the HO management problem. Instead, we can learn a parameterized value function  $Q(s, a; \theta_t)$ . The standard  $Q$ -learning update for the parameters after taking action  $A_t$  in-state  $S_t$  and observing the immediate reward  $R_{t+1}$  and resulting state  $S_{t+1}$  is then

$$\theta_{t+1} = \theta_t + \alpha \left( Y_t^Q - Q(S_t, A_t; \theta_t) \right) \nabla_{\theta_t} Q(S_t, A_t; \theta_t) \tag{17}$$

where  $\alpha$  is a scalar step size and the target  $Y_t^Q$  is defined as

$$Y_t^Q \equiv R_{t+1} + \zeta \max_a Q(S_{t+1}, a; \theta_t) \tag{18}$$

This update in Eqn. 17 resembles a stochastic gradient descent, and since it uses a neural network for value estimation, this type of  $Q$ -learning is commonly known as deep  $Q$ -learning (DQN) or deep reinforcement learning (DRL)<sup>11</sup>. DQN is explained in section 3.4.2, and in the latter section, we introduce DDRL, which improves the efficiency of DQN. Up till now, we have been presented the components of RL and introduced the HO management concept. In the following, elements of RL that are vital to HO are described to associate these elements to the problem.

---

<sup>11</sup>DQN and DRL are used interchangeably throughout this thesis unless stated otherwise. The same applies to DDQN and DDRL.

**a) Action:** An action represents the BS to connect to if A2 event occurs. We define an action in an action space ( $a \in A(s)$ ) as the scalar representation of the S-BS index at state  $s$ . The set  $A(s)$  comprises all BS in the environment.

**b) State Vector:** Traditionally, mobility management and other BS association strategy consider the location of the UE to associate it to the S-BS. This study, however, considers the combination of the SNR received by the UE from all surrounding BSs to represent the location of interest instead of the exact location of the UE (i.e. geo-coordinates of UE's location). Getting the exact location of the UE is impractical in reality; hence, we consider  $\gamma$  from all BSs along the UE trajectory as representative of the point of interest instead of geo-coordinates.

Therefore, we correlate the current position of the UE to the channel quality from all surrounding BSs. We also assume constant SNR values are realised from all BSs at a particular point based on the assumption that average SNR is used, and the combination of the average SNR is uniquely sufficient to represent the geo-location coordinates at the specific position.

Hence, at the point,  $p$  with a total number of  $M$  BS, the state vector for the UE is given as

$$s_p = \{\gamma_1, \gamma_2, \dots, \gamma_i, \dots, \gamma_M, \text{BS}_i\}, \quad (19)$$

where  $s_p \in S$  is the state at point  $p$ ,  $\gamma_n$  is the SNR value from the  $n^{th}$  BS, and  $\text{BS}_i$  is the index of the S-BS at point  $p$ . The BS index is in one-hot encoded vector. One-hot encoding (Mollet et al., 021a) is the vector representation of the integer variable into the binary value of all zero except the index of the integer. For instance, if  $\text{BS}_i$  is designated as BS three and there are total of five BSs, then the equivalent one-hot encoding vector becomes  $\text{BS}_i = [0, 0, 1, 0, 0]$ .

However, since the agent has only a single snapshot of the channel quality at a particular point, this partial observation presents a problem called perceptually aliased, i.e. it is impossible to understand the current status of the channel using only the current (single) observation. For example, if we consider only instantaneous SNR at the current point, then it becomes impossible to know the direction of the UE or obtain information regarding the previous S-BS. Therefore, Eqn. 19 is revised and the updated state equation considers a sequence of actions and observations, given by

$$s_t = \{s_p, s_{p-1}, s_{p-2}, s_{p-3}\}, \quad (20)$$

where  $s_p$  is the latest observation and  $s_{p-3}$  is the observation in the previous three steps. All

observations in the environment are assumed to terminate in a finite number of time-steps, i.e. at the end of the UE's trajectory.

**c) Reward Design:** The reward design is used to motivate the agent to take actions that would maximize the cumulative reward in the long run since our objective is to achieve maximum system throughput ( $\mathcal{T}$ ) for a given trajectory. Equation 9 shows that we can maximize  $\mathcal{T}$  by minimizing  $\beta_c$ . To minimize  $\beta_c$ , Eqn. (8) shows that for a given velocity ( $v$ ) and HO time delay ( $t_d$ ), the parameter  $\mathcal{H}_l$  should be as small as possible. The parameter,  $\mathcal{H}_l$  can be controlled by implementing a HO skipping policy. Technically, the agent initiates indirect TTT without setting a constant value, and this should be done intelligently to ensure that the UE can achieve maximum throughput even if it skips some unnecessary HOs. This method has been used for micro and macro BS before 4G, and the TTT parameter was manually determined. Additionally, to minimise the value of  $\beta_c$  while maximising  $\mathcal{T}$  during HO, the agent can select BSs that have few event A2 in the future, known as far-sighted HO decision, provided the constraint  $\gamma_s \geq \gamma_{th}$  is met.

In reward design, we avoid delayed rewards due to the problem of credit assignment (Sutton & Barto, 2018). Therefore, we introduce the immediate reward function as the instantaneous throughput and evaluate the immediate effect of the action taken to achieve the agent's goal. Such a reward is given as follows:

$$r(s_{t+1}, a, s_t) = \begin{cases} B \times \mathcal{R} \times (1 - \beta_c), & \text{if HO occurs} \\ B \times \mathcal{R}, & \text{otherwise} \end{cases} \quad (21)$$

where  $B$  is the maximum bandwidth allocated,  $\mathcal{R} = \log_2(1 + \gamma_s)$  is the spectral efficiency,  $\gamma_s$  is the average instantaneous SNR the UE experiences from the S-BS at  $s_t$  obtained from the simulated environment. For the proposed model to work accurately, information should be collected that the agent can use for decision making.

**d) Experience replay:** The objective of experience replay is to overcome the instability of the learning algorithm. Experience replay is used to update the deep Q-network in such a way that both current and previous experiences are considered in the supervised learning-based update process. This means that not only samples  $(s, a, r, s')$  obtained from current online learning network but also old experience tuples  $(s, a, r, s')$  are considered in the training process. Hence, experience replay store observed transitions for some time and sample uniformly from this memory bank to update the network. Using experience replay improves the performance of the

algorithm (van Hasselt et al., 2015).

### 3.4.2 DRL-based optimal BS selection

In this section, we present the challenges of applying tabular  $Q$ -learning to our problem, and then provide a detailed explanation of how deep  $Q$ -learning can help to solve these challenges. It is known that if the environment has many states such as if the number of states is of the order of hundreds of states and hundreds of actions per state, this would result in a  $Q$ -table with ten thousand cells, hence the learning process will quickly get out of control. This practically infinite number of states and actions creates two problems. The first problem is that the amount of memory required to store and update the state action table increases as the number of states increases, and secondly, the time spent to explore each state in order to populate the  $Q$ -table accurately (Sutton & Barto, 2018) becomes significantly high. Another limitation of  $Q$ -learning is that it only works in environments with discrete and finite state and action spaces, which implies that  $Q$ -learning is unable to estimate  $Q$ -values for any unlearned states.

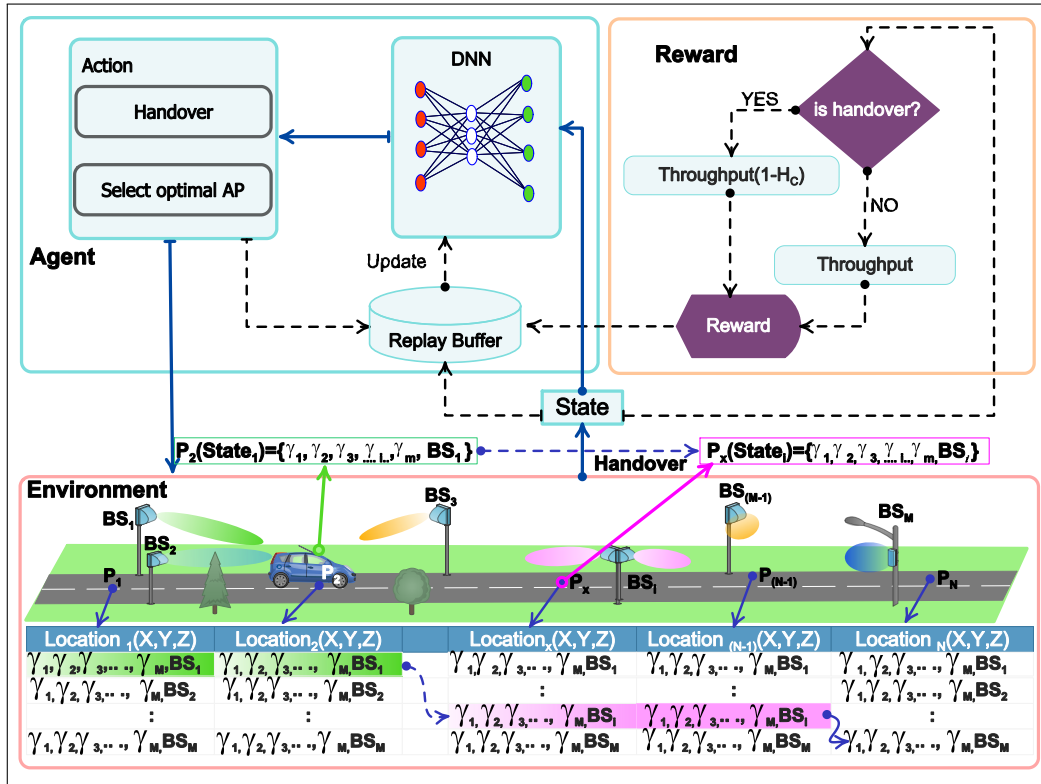


Figure 18: DRL-based framework comprising environment, states, actions, and rewards.

The DQN is a multi-layered neural network that maps an input state  $s$  to an action value  $Q$ , where  $\theta$  represents the parameters of the network. In the learning phase, we use offline learn-

---

**Algorithm 1:** Deep  $Q$ -learning with Experience Replay

---

```
1 Initialize replay memory  $\mathcal{D}$  to capacity  $\mathcal{N}$ ;  
2 Initialize action-value function  $Q$  with random weights;  
3 for  $episode = 1, M$  do  
4   Get Initialise state,  $s$ ;  
5   for  $t = 1, T$  do  
6     With probability  $\epsilon$  select a random action  $a_t$ ;  
7     otherwise select  $a_t = \max_a Q^*(s, a; \theta)$ ;  
8     Execute action  $a_t$  in environment and observe reward  $r_t$ ;  
9     Get new state,  $s'$ ;  
10    Store transition  $(s_t, a_t, r_t, s')$  in  $\mathcal{D}$ ;  
11    Sample random minibatch of transitions  $(s, a, r, s')$  from  $\mathcal{D}$ ;  
12    Set  $y_j = \begin{cases} r_j & \text{for terminal } t==T \\ r_j + \zeta \max_{a'} Q(s_{j+1}, a'; \theta) & \text{otherwise} \end{cases}$  ;  
13    Perform a gradient descent step on  $(y_j - Q(s_j, a_j; \theta))^2$  according to  
    equation 17  
14  end  
15 end
```

---

ing, whereby the agent gathers necessary information by simulating the UE trajectory in the environment, as shown in Fig. 18. The agent simulates the trajectory from starting point to the endpoint of the UE's path, and then performs HO in a trial-and-error fashion. It is worth noting that we assume trajectory aware HO. Therefore, the path taken by the UE is known, and during HO, the agent can select any BS from among  $M$  BSs. The objective is to maximize cumulative rewards. The complete DQN algorithm is presented in Algorithm 1

**a) Data Preprocessing:** Working directly with SNR information as observations can harm learning since we use neural networks that require data processing, so we apply an essential preprocessing step to improve the performance of the model by normalizing input features. The raw data is processed by first finding the maximum achieved and minimum allowed SNR and normalizing them using the max-min normalization technique in Eqn 22.

$$\frac{\gamma_i - \gamma_{\min}}{\gamma_{\max} - \gamma_{\min}} \quad (22)$$

Table 9: Parameters for designing and developing DQN model

Parameter	Value	Parameter	Value
Hidden layers	4	Neuron size	$512 \times 256 \times 128 \times 64$
Activation function hidden layers	relu	Activation function output layer	linear
Initial exploration training	1	Final exploration training	0.01
Optimizer	adam	Output layer	M BS
Learning rate, $\alpha$ and Discount Factor, $\zeta$	0.001 , 0.96	Minibatch size	32

The min-max normalization technique has a significant downside: it does not deal very well with outliers. We, therefore, set the ceiling  $\gamma_{max}$  and floor  $\gamma_{min}$  for our data. Thus, any  $\gamma_i$  above the ceiling or below the floor is truncated to the ceiling  $\gamma_{max}$  or floor  $\gamma_{min}$ , respectively. As described in Section 3.4.1, the final input representation of a state is obtained by inserting a serving BS index before HO takes place, and then this preprocessing observation is carried out on the last four history frames and stacked together to produce the  $Q$ -function input state.

**b) Model Architecture:** Our architecture considers the output layer, which has a separate unit for every action possible. That means that there are  $M$  units corresponding to the  $M$  BSs. Using the  $s$  state as an input, the output layer from the network estimates a value function,  $Q$ . The importance of this architecture is shown by its ability to produce an estimated  $Q$ -values of every action in a given state.

The DNQ architecture can be described as follows: the input to the neural network is the state described in Section 3.4.1.b and the number of input units reflects the state’s content. Then four fully connected linear layers with  $512 \times 256 \times 128 \times 64$  neural units are implemented as the hidden layer using a Rectified Linear Unit (ReLU). The output layer is a fully connected linear layer with a single output for each valid action. The number of valid actions depends on the number of  $M^{th}$  BS. We call this architecture DQN; a summary of the training and the architecture is presented in Table 9.

### 3.4.3 DDRL-based optimal BS selection

It was noted by van Hasselt et al. (2015) that for some games, the max operator in Eqn. 18 tends to select overestimated values resulting in over-optimistic values. To prevent this, the decoupled architecture is introduced to yield the selection and evaluation network. This is the underlying principle behind DDQ-learning (van Hasselt et al., 2015). The DDRL architecture works by exploiting the advantages of previous models, that is,  $Q$ -learning and DQN. DDRL is a DRL that uses and maintains two separate deep  $Q$ -Networks (DQN). According to van Hasselt et al. (2015), the two separate networks for DDRL include a target network and an online network. The target network, with parameter  $\theta^-$ , is the same as the online network except that its parameters are updated from the online network every  $\tau$  steps, such that  $\theta_t^- = \theta_t$  and kept fixed on all other steps. DDRL reduces overestimation by decomposing the max operation in the target network into action selection and action evaluation steps. Therefore, the greedy policy is evaluated according to the online network while values are estimated in the target network.

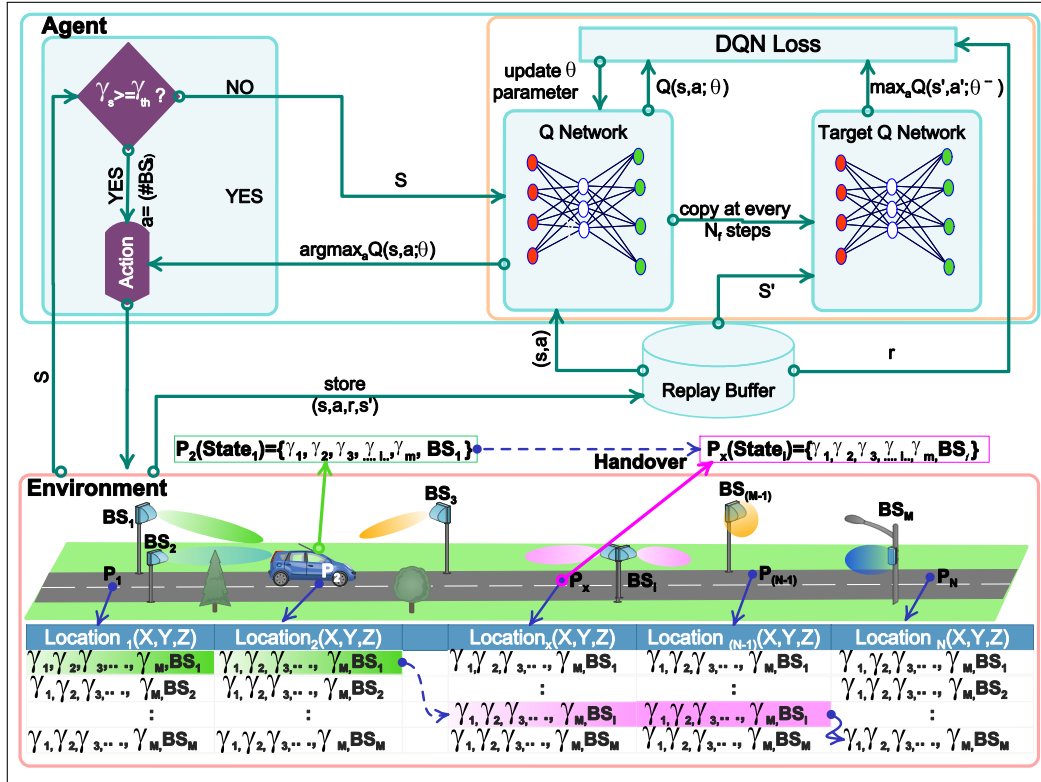


Figure 19: The structure of the proposed DDRL - with double Q- networks

Apart from the agent network having decoupled networks, the rest of the parameters and data preprocessing for DDQN are the same as those for DQN presented in Section 3.4.2. The only

---

**Algorithm 2:** Double Deep  $Q$ -learning

---

```
1 input :  $\mathcal{D}$ — empty replay buffer;  $\theta^-$  initial network parameters,  $\theta^-$ — copy of  $\theta$ ;  
2 input :  $N_r$ — replay buffer maximum size;  $N_b$ — training batch size;  $N^-$  - target network  
   replacement freq.;  
3 for  $episode = 1, M$  do  
4   Get Initialise state,  $s$ ;  
5   for  $t = 1, T$  do  
6     With probability  $\epsilon$  select a random action  $a_t$ ;  
7     otherwise select  $a_t = \max_a Q^*(s, a; \theta)$ ;  
8     Execute the action in the environment and observe the reward  $r_i$  and the next  
       state ( $s'_i$ ).;  
9     Store transition tuple  $(s, a, r, s')$  to  $\mathcal{D}$ , replacing the oldest tuple if  $|\mathcal{D}| \geq N_r$ ;  
10    Sample a minibatch of  $N_b$  tuples  $(s, a, r, s') \sim \text{Unif}(\mathcal{D})$ ;  
11    Construct target values, one for each of the  $N_b$  tuples;;  
12    Define  $a^{\max}(s'; \theta) = \arg \max_{a'} Q(s', a'; \theta)$   
       
$$y_j = \begin{cases} r & \text{if } s' \text{ is terminal} \\ r + \zeta Q(s', a^{\max}(s'; \theta); \theta^-), & \text{otherwise.} \end{cases};$$
  
13    Do a gradient descent step with loss  $\|y_j - Q(s, a; \theta)\|^2$ ;  
14    Replace target parameters  $\theta^- \leftarrow \theta$  every  $N^-$  steps;  
15  end  
16 end
```

---

component added is the  $\tau$  step which used to update the target network (the value was set to 40000 steps in this work). Figure. 19 shows the complete system environment for our proposed model based on DDRL. The pseudo-code for the DDQN algorithm and how it works in regard to the proposed solution is given in Algorithm. 2. It should be noted that vital elements in Algorithm. 2 are the same as in Algorithms 1.

Since it is uncertain that DQN can work efficiently with any RL problem, a comparative study of DQN and DDQN was performed to determine DQN efficiency in solving HO optimization decisions compared to DDQN. The experiments were conducted in the same setup for both DQN and DDQN. The main reason for doing this was to investigate if DDQN and DQN can

lead to the same conclusion. The results of these experiments are presented in Fig. ?? and Fig ?. Figure. ?? illustrates an example of DQN(red color)-induced overestimation during training. It is critical to remember that DQN and DDQN agents are trained under identical seed and environmental conditions. As shown by the red learning curves, DQN is often vastly over-optimistic about the value of the current greedy policy. The average value estimate is calculated for the trajectory during the training and evaluation phases. The equation for value estimates is as follows:

$$\frac{1}{T} \sum_{t=1}^T \max Q(S_t, a; \theta) \quad (23)$$

where T is the number of steps used for evaluation,  $\max Q(S_t, a; \theta)$  is the value estimate of the online network. The figure illustrates more serious over-estimations and emphasizes how highly unstable DQN is.

The average cumulative reward (score) against the training episodes for each training algorithm is shown in Fig. ?. Generally, it can be seen that there are increases in value estimates for DQN as shown in Fig. ?, which coincide with decreasing scores in Fig. ?. This demonstrates that overestimations have a detrimental effect on the quality of the resulting policies. From Fig. ?, it can be observed that the learning trend for DRL is quite similar to DDRL for some episodes, which are from the start of training step to nearly  $3 \times 10^7$  training steps. This implies that for both algorithms, an agent equally learns and improves its policy at least for a given range of the training steps. However, after certain steps—above  $3 \times 10^7$ , the learning curve for DRL (shown in red) starts to drop while the learning curve for DDRL (shown in blue) remains high with a stabilizing trend compared to that of DRL. The main reason for the decline in cumulative reward is because DRL usually tends to over-estimate value functions during the training phase. Therefore, judging from both curves, it is without a doubt that DDRL produces more accurate value estimates, and better policies because of higher return and more stable learning throughout the training process. For this reason, we do not show any further results to compare DDRL and DRL as DDRL is guaranteed to perform better than DRL in all cases.

## CHAPTER FOUR

### RESULTS AND DISCUSSION

#### 4.1 Performance Evaluation

This section presents a performance analysis of the proposed DDRL HO management solution for 5G UDNs. It comprises three parts; a) the first part describes the simulation environment, including the simulation setup, b) the performance metrics are presented in the second part, which provides a brief insight into the measurement parameters used, c) while the third part offers a comparative analysis that compares the proposed method with benchmarks solutions.

##### 4.1.1 Simulation Setup

In both experiments, we consider the system model as described in Section 3.1. We consider a square environment with an area of  $1000(\text{m}) \times 1000(\text{m})$ , with randomly deployed BSs. We then utilise the random waypoint model (Bettstetter et al., 2004) to generate user trajectory with an average velocity of  $8\text{ms}^{-1}$  for experiment 1. For other experiments, we vary the velocity with a constant threshold SNR ( $\gamma_{\text{th}}$ ) of 20 dB. The random way-point model was used to generate 10 trajectories and assign probability distributions to four directions (North, South, East, and West) as follows; trajectories 1 - 5 were assigned the distribution  $[0.25, 0.25, 0.25, 0.25]$  while the other trajectories were assigned  $[0.6, 0.2, 0.1, 0.1]$ . The cumulative HO time delay during HO process  $t_d$  was set to  $[0.5, 0.75, 1, 2]$  sec (Arshad et al., 2016). However, many previous studies used 1 sec as the average HO time delay between mm-wave BSs. In this study, we deliberately used more than one value of HO time delay to evaluate and quantify the effect of HO time delay associated with a HO event. The observation time for UE mobility for all trajectories is 10000 secs, and the status of CSI is recorded at every 10 ms for both the training and evaluation data sets. However, for evaluation, the starting position of the UE along the trajectory is time-shifted to more closely approximate real life scenarios. All the states fed to the DDRL are generated within this time period. Ultimately, we ignore the effect of interference as its effect is negligible, as discussed in Section 3.1.3. Hence, we consider SNR as the parameter of interest. We consider hyper-parameters for DDRL from (Koda et al., 2019) since systematic grid search incurs a high computational cost. The complete parameters for the radio network environment and the DDRL are given in Table. 10 and Table. 9, respectively.

Table 10: Environment simulation parameters

Parameter	Value
BS intensity	10 - 70 (BS/km <sup>2</sup> )
mm-wave frequency	28 GHz
mm-wave bandwidth	1 GHz
BS transmit power	30 dBm
Thermal noise density	$-174 \text{ dBm/Hz}$
Delay without data transmission $t_d$	0.75, 1, 2, 3 sec

#### 4.1.2 Performance Metrics

The proposed method was evaluated using performance metrics such as system throughput, average running time, number of ping-pongs and number of HOs. These are summarized as follows:

##### (i) Experienced SNR

Experienced SNR is the average instantaneous SNR experienced by the user in a given trajectory. The average SNR, denoted as  $\gamma_{ave}$ , is given as:

$$\gamma_{ave} = \frac{1}{T} \sum_{t=1}^T \gamma_{s=t}, \quad (24)$$

where  $T$  is the total number of locations where the UE exchanges data with the BS and  $\gamma_{s=t}$  is the SNR value experienced from the serving BS at index  $t$  of such locations.

##### (ii) Number of HOs

Number of HOs ( $n_{HOs}$ ) is the sum of HO events a UE experiences from the starting point to the destination in a given path. Mathematically, this is given by:

$$n_{HOs} = \sum_{t=1}^T \mathbb{1}_{HO}, \quad (25)$$

where  $\mathbb{1}_{HO}$  is an indicator function ( $\mathbb{1}_{HO} \doteq 1$  if the HO occurs and 0 otherwise) and  $T$  is the total number of locations where the UE exchanges data with the BS.

### (iii) Ping-pong HO rate

Ping-pong HO rate is defined as the total number of ping-pong events to the total number of HOs. Ping-pong effect is described in Section 2.3.2 as an event in which the number of HOs over a specified period ( $T_{pp}$ ) exceeds a predetermined threshold (for this study more than 1 HO in  $T_{pp}$ ). This study considers ( $T_{pp}$ ) equal to 100 ms, and ping-pong HO rate is described mathematically as

$$\text{Ping-pong rate} = \frac{\sum_{t=0, nT_{pp}, \dots}^T \left[ \sum_{i=t}^{t+T_{pp}} \mathbb{1}_{HO} > 1 \right]}{n_{HOs}}; \forall t \in [t_0, t_0 + T_{pp}], \quad (26)$$

where  $n = 1, 2, \dots, k$ ; given that  $k$  satisfies  $kT_{pp} \leq T$ , and  $n_{HOs}$  is the number of HOs the UE experiences along the trajectory.

### (iv) Average Throughput

Average throughput is the average data rates that are delivered to a UE or rate of successful message delivery over a communication channel. The average throughput, denoted as  $\mathcal{T}$ , is defined as:

$$\mathcal{T} = B \times \log_2(1 + \gamma_{ave}) \times (1 - \beta_c), \quad (27)$$

where  $B$  is the maximum bandwidth allocated,  $\beta_c$  is the HO cost defined in Eqn. 8, and  $\gamma_{ave}$  is the average instantaneous SNR the UE experiences throughout the trajectory in the simulated environment.

### (v) Average execution time

The execution time is defined as the average time spent by the model to produce results based on given inputs.

## 4.1.3 Comparative Analysis

This section compares the proposed solution to the heuristic RHB (3GPP, 2018b) approach, the smart handoff policy (SHP) (Sun et al., 2018), and a benchmark HO policy named clustering dense network to sparse networks to minimize HO in 5G mm-wave networks (CDSHO) presented in Section 3.3.1 and in (Mollel et al., 2019).

(i) **Analysis of User-experienced SNR ( $\gamma$ ) against HO interruption time ( $t_d$ )**

In the first experiment, we consider the velocity of the UE to be  $8 \text{ ms}^{-1}$  for UE trajectories, and then we analyse the relationship between the the experienced average SNR received  $\gamma_{ave}$  by the UE for the proposed model against  $t_d$ . Figure 20 shows the relationship between the  $\gamma_{ave}$  against HO interruption time  $t_d$  as parameter in RL reward function for  $\lambda=10$  and  $50 \text{ BSKm}^{-2}$ , respectively. It can be seen clearly from Fig. 20 that as  $t_d$  increases,  $\gamma_{ave}$  decreases but at a very slow rate for both  $\lambda=10$  and  $50 \text{ BSKm}^{-2}$ . This behaviour illustrates how the agent can sacrifice maximization of  $\gamma_{ave}$  by selecting a BS that does not necessarily give  $\gamma_{max}$  for some time (but which may lead to a HO after only a short connection period) but the one which leads to maximum throughput (by providing a longer connection period, even if at a lower  $\gamma_{ave}$ ), since throughput maximization is an objective function of the algorithm.

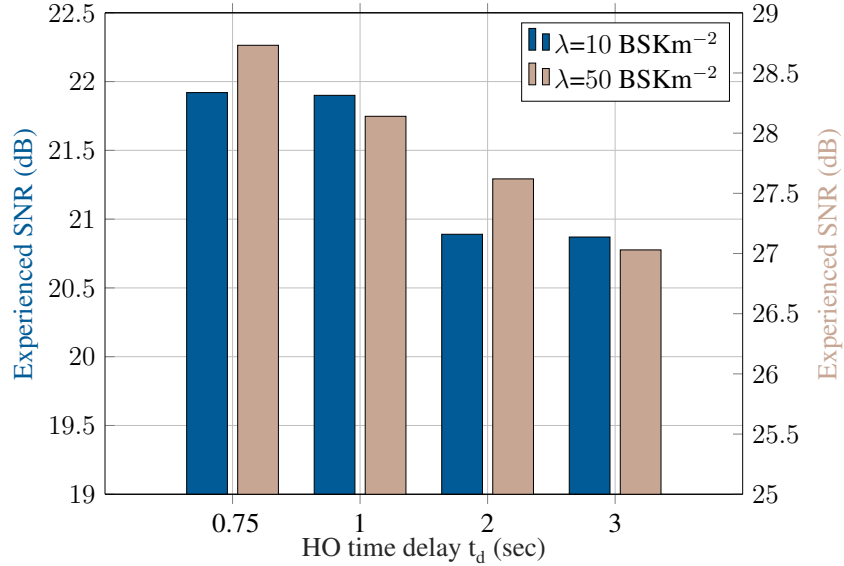


Figure 20: Number of HOs and  $\gamma_{ave}$  as functions of  $t_d$ , for  $\gamma_{th} = 20 \text{ dB}$  and  $\lambda=10 \text{ BSKm}^{-2}$

It worth noting that in real environments,  $t_d$  varies randomly due to factors such as how sophisticated the network equipment is, the channel response, and how easy it is to discover T-BSs. Since  $t_d$  is part of the reward, it has a significant influence in determining the learning and selection of BS that will maximise the cumulative reward. Also, it should be noted that  $t_d$  is a random parameter that is not constrained by design but dictated by the real environment. However, for the sake of RL design, one can select  $t_d$  as a critical parameter in order to control the trade-off. It can be clearly seen from Fig. 20 that changing the reward signal encourages the agent to select actions that substantially reduce the number of HOs by sacrificing achieving  $\gamma_{max}$ .

## (ii) Analysis of HO Interruption time $t_d$

This experiment aims to examine the relationship between the number of HOs and the SNR threshold ( $\gamma_{th}$ ) for various values of  $t_d$  as a reward in the proposed DDRL model, as illustrated in Fig. 21 and Fig. 22 for  $10 \text{ BSKm}^{-2}$  and  $20 \text{ BSKm}^{-2}$ , respectively. In this experiment, the following set of parameters were considered for the simulation: a) average UE velocity of  $8 \text{ ms}^{-1}$ , b) simulation time of 10000 sec.

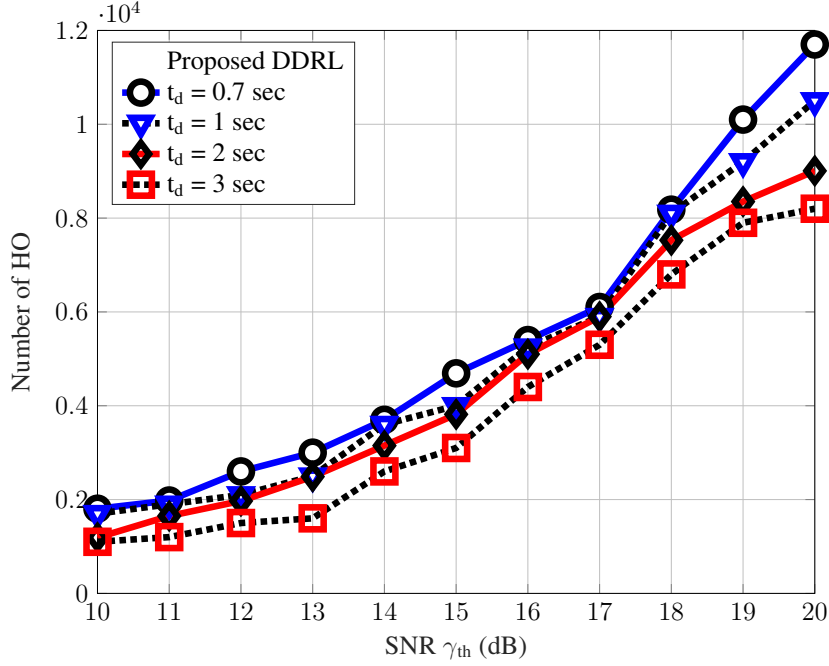


Figure 21: Number of HOs as a function of  $\gamma_{th}$  for  $t_d = 0.7, 1, 2, 3$  sec and  $\lambda = 10 \text{ BSKm}^{-2}$

From both figures, it can be understood that different values of  $t_d$  have a different contributions towards reinforcing the agent to reduce the number of HOs. However, for specific values of  $\gamma_{th}$ , there is a slight difference in the number of HOs for values of  $\gamma_{th}$  ranging from 10 - 18 when  $\lambda = 10 \text{ BSKm}^{-2}$  and between 15 and 23 when  $\lambda$  equals  $50 \text{ BSKm}^{-2}$ . On the other hand, with increasing  $\lambda$ , the contribution of  $t_d$  in reinforcing the agent to reduce HO is observed. Throughout this thesis,  $t_d = 3$  sec unless otherwise stated.

Another interesting point observed in both figures is the fluctuation of the curves, meaning that for different values of  $t_d$  and same values of  $\gamma_{th}$ , we can have the same or different number of HOs. This is because during the evaluation, we set the value of exploration  $\epsilon$  to  $2 \times 10^{-3}$ , although the trend can also be understood from the explanation given earlier. Therefore, an agent may ignore selecting a BS that can result in maximum reward and choose to explore the

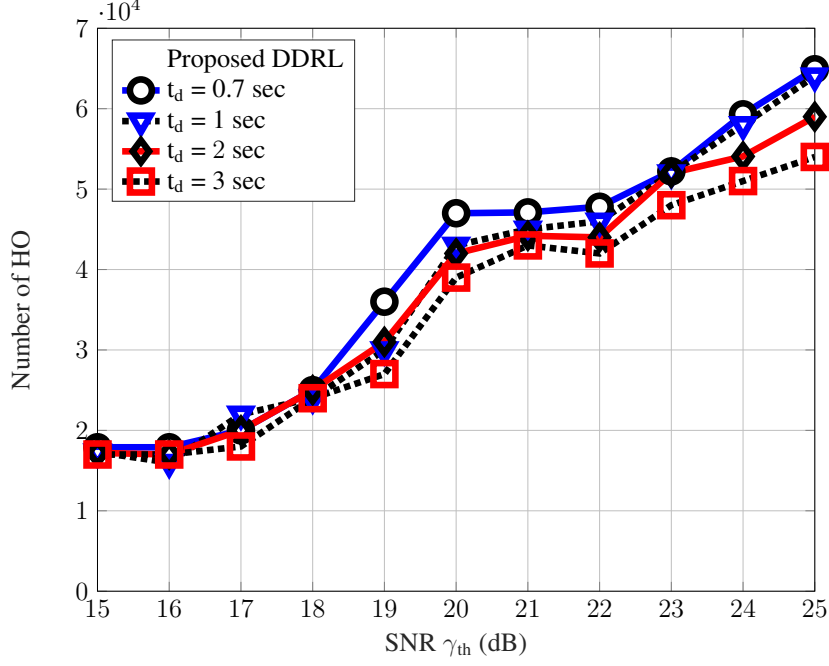


Figure 22: Number of HOs as a function of  $\gamma_{th}$  for  $t_d = 0.7, 1, 2, 3$  sec and  $\lambda = 50 \text{ BSkm}^{-2}$

environment by selecting the BS with  $\gamma_{max}$  or selecting a BS randomly. This effect can be neglected if the agent can set the  $\epsilon$  to 0 after the learning phase. The factors for selecting  $t_d$  may vary depending on the agent's objective. The agent does not consider factors such as network response and intuitively, the network with the lower  $t_d$  provides a good QoS and QoE. However, the agent can decide to use a higher  $t_d$  in the training phase if the objective is to minimize HOs due to other advantages that the agent could gain by such a choice.

### (iii) Effect of UE velocity on HOs

For the third experiment, we assess the performance of the proposed model by comparing it with other benchmark solutions in terms of the number of HOs and the system throughput. The parameters for the experiment were set as follows:  $t_d = 3$  sec,  $\gamma_{th} = 20$  dB and  $\lambda = 50 \text{ BSkm}^{-2}$  while other parameters remain unchanged. Figure. 23 and Fig. 24 demonstrate the average system throughput and the number of HOs for the three benchmarks as well as the proposed HO management policies against UE velocity. From Fig. 23, it can be seen that the proposed DDRL outperforms other policies. The general trend shows a slight and gradual increase in number of HOs for the benchmark models as velocity increases.

Since the CDSHO has fewer BSs to select from at any given point, it automatically reduces the number of HOs. In this regard, in the CDSHO solution, UEs are usually handed over to the sub-

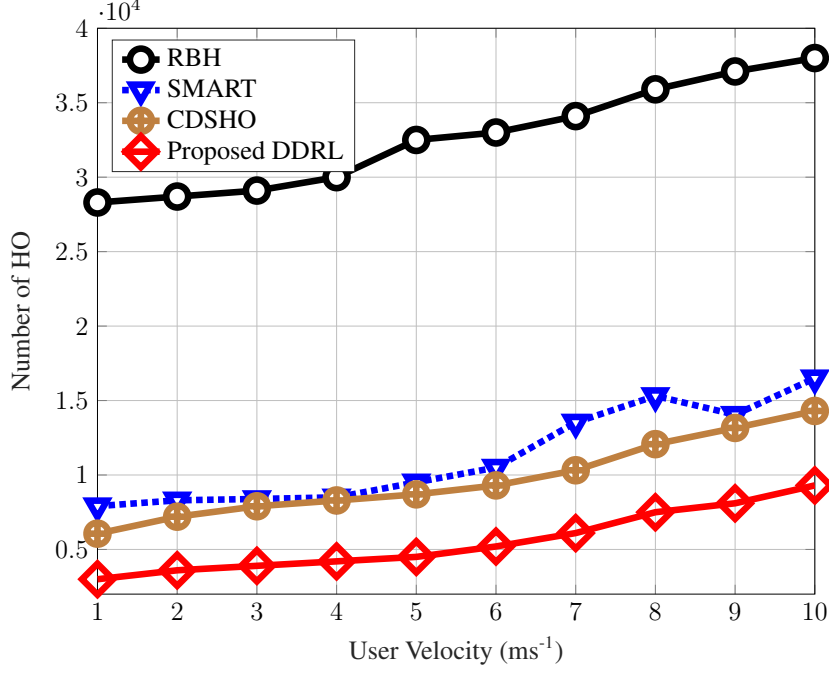


Figure 23: Relationship between the number of HOs and UE velocity

6 GHz channel if non of the mm-wave BSs in the cluster can accommodate the UE. In addition, this has repercussions in terms of throughput since the capacity of sub-6 GHz is lower than that of mm-wave links. The SHP technique relies on bandit RL which is an online learning method. Hence, it will require some time to figure out the best BS to select at any given time in order to reduce the number of HOs. This is the leading cause of the high number of HOs encountered in the SHP solution. The RBH, on the other hand, is expected to have a high number of HOs since it always selects the BS that provides a high data rate during HO decisions.

The most significant result is shown in Fig. 24 where the proposed model far outperforms other models in terms of throughput, especially for high mobility UEs. Additionally, it can be observed that as the UE velocity increases, the system throughput decreases due to rapid changes in the channel quality. On the one hand, as expected, CDSHO shows a lower throughput compared to other solutions. This is expected since the objective of CDSHO is to minimise HOs for the UE without considering the trade-off of reducing QoS (throughput). That is, the CDSHO algorithm limits the UE to select a target BSs from a specific cluster only. If none of the BSs within the cluster satisfies the minimum  $\gamma_{th}$ , the UE HO to the sub-6 GHz macro-cell instead of considering BSs in nearby clusters with higher throughput, which may result in lower average system throughput. In addition, the SHP and RBH schemes select T-BSs on a shortsighted basis. For SHP, it aims to reduce the number of HOs but neither considers the UE trajectory

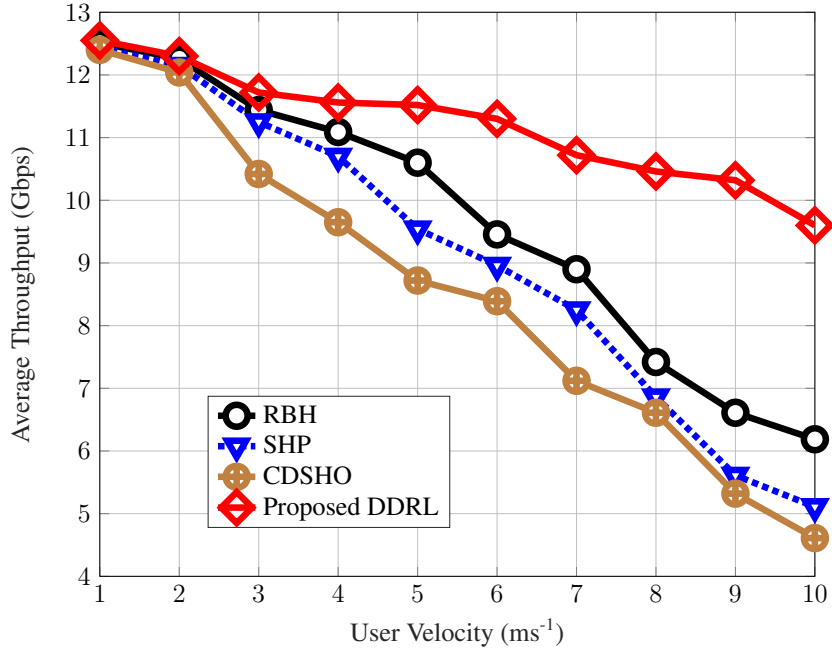


Figure 24: HO performance showing the relationship between average system throughput and UE velocity

nor maintains an extended connectivity of the chosen link while RBH only selects the BS with the maximum  $\gamma_s$  whose connection to the UE might break after a few seconds. In contrast, the proposed DDRL offers a high throughput compared to other benchmarks since it incorporates trajectory awareness and training based on searching for the optimal combination of BS that can maintain more extended connectivity, leading to both high a QoE and QoS.

#### (iv) Time complexity

The fourth experiment compares the average running time per single HO decision to the number of mm-wave BSs for each of the four HO strategies. The simulation parameters are as follows:  $t_d = 3$  sec, UE velocity =  $8 \text{ ms}^{-1}$ , and  $\gamma_{th} = 20$  dB. As shown in Fig. 25, all the policies exhibit a similar pattern. It can be observed that the proposed model takes more time to make a HO decision compared to RBH; whereas it takes a significantly lower time compared to SHP and CDSHO. Additionally, there is a linear relationship between an increase in the number of SCs and the running time for all policies. For a small number such as  $\lambda = 10 - 30 \text{ BSkm}^{-2}$ , the time complexity of CDSHO is higher than the other three HO policies. However, as the BS density increases ( $\lambda \geq 40 \text{ BSkm}^{-2}$ ), the time complexity of CDSHO becomes lower than SHP. This is because the search space in CDSHO increases slowly due to clustering while that of SHP keeps

increasing. It is worth noting that SHP is based on bandit algorithms, and as such, the time for convergence increases as the number of actions increases. The proposed model outperforms SHP and CDSHO since there is no computational complexity in training the neural network where the only computation required is for predicting an output based on the given input. For RBH, it is only an *if statement*, and as expected, the time for making the decision is lower than other policies. The simulation shows that the maximum time required to make a HO decision

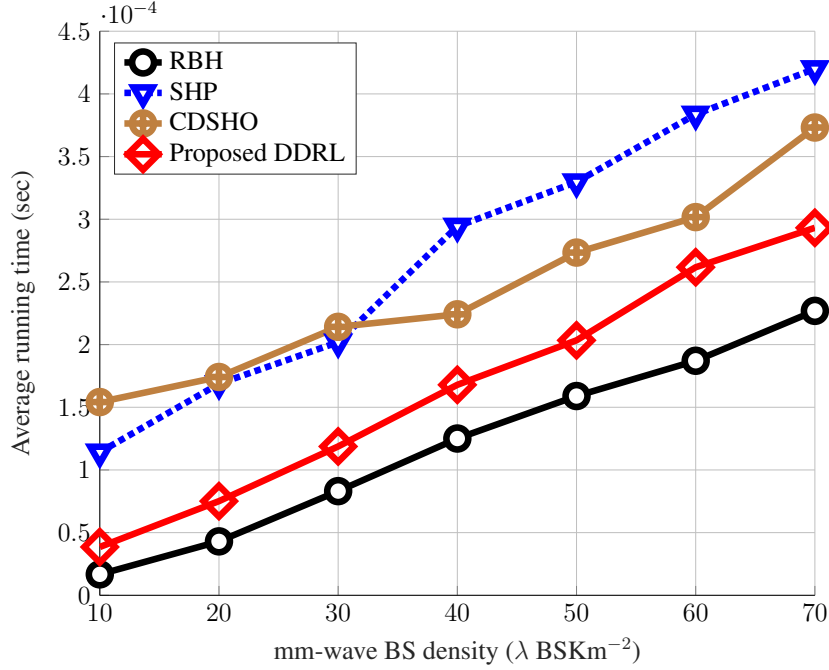


Figure 25: Average running time as a function of number of mm-wave BSs

is 42 ms, which occurs when  $\lambda = 70$  BS km $^{-2}$  in the SHP HO policy. This finding suggests that the decision time for HO is acceptable, as the selection of BS occurs within the minimum permissible  $T_{RLF}$  of 50 ms, as shown in Table. 7. Nonetheless, as the number of SCs increases, the running time for making HO decisions increases proportionately and UEs can frequently encounter RLF. The solution to this dilemma is beyond the scope of this work.

#### (iv) Effect of UE speed on the Ping-Pong rate

The analysis of the ping-pong rate considers the following simulation parameters:  $T_{pp} = 100$  ms,  $t_d = 3$  sec,  $\lambda = 70$  BS km $^{-2}$ . In this experiment, it is observed in Fig. 26 that the ping-pong rate of the proposed method is low compared to the benchmark methods. This is because the proposed method has an intelligent way of indirectly initializing TTT, which means avoiding taking HO decisions unnecessarily (skipping HO). In addition, once HO occurs, the proposed approach

selects the BS that can sustain communication for an extended period of time; this ensures a low number of HOs and a low ping-pong rate. It is known that the velocity and high density of SCs density significantly affects the ping-pong rate since the dwell time for the UE is low, but the proposed method shows a slight increase in the ping-pong rate as the velocity increases, as seen in Fig. 26.

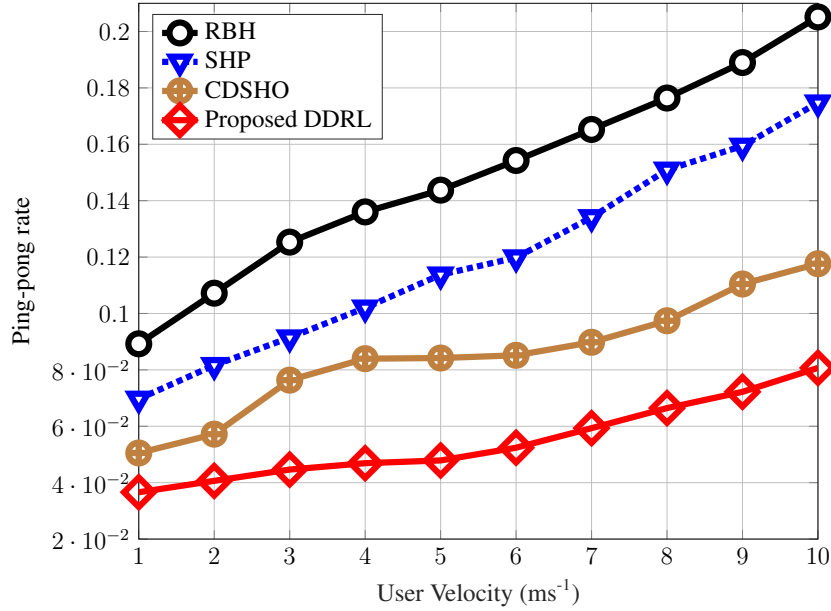


Figure 26: Ping Pong rate as a function of UE velocity

#### (v) Effect of BS Density

In this last experiment, we compare the performance of the benchmark HO policies with the proposed HO policy by varying the number of mm-wave BSs. The parameters for this experiment are as follows: UE velocity =  $8 \text{ ms}^{-1}$ ,  $\gamma_{\text{th}} = 20 \text{ dB}$  while the other parameters remain unchanged. The results of the evaluation are shown in Fig. 27 and in Fig. 28. It can be clearly observed that the proposed DDRL solution outperforms the other HO policies.

From Fig. 27, it can be seen that the proposed HO policy has outstanding effect in reducing the number HOs by 20% - 69% , 7% - 49%, and 5.7% - 21% compared to RBH, SHP, CDSHO, respectively. This is because the proposed intelligent method considerably reduces the number of unnecessary HOs (ping-pong and other HOs). In addition, the CDSHO benchmark offers the next best performance after the proposed DDRL algorithm in terms of low number of HOs for different BS densities. As previously mentioned, this approach achieves excellent performance because it restricts UEs to a few BSs and switches to sub-6GHz BSs when  $\gamma_s$  of T-BS is less than

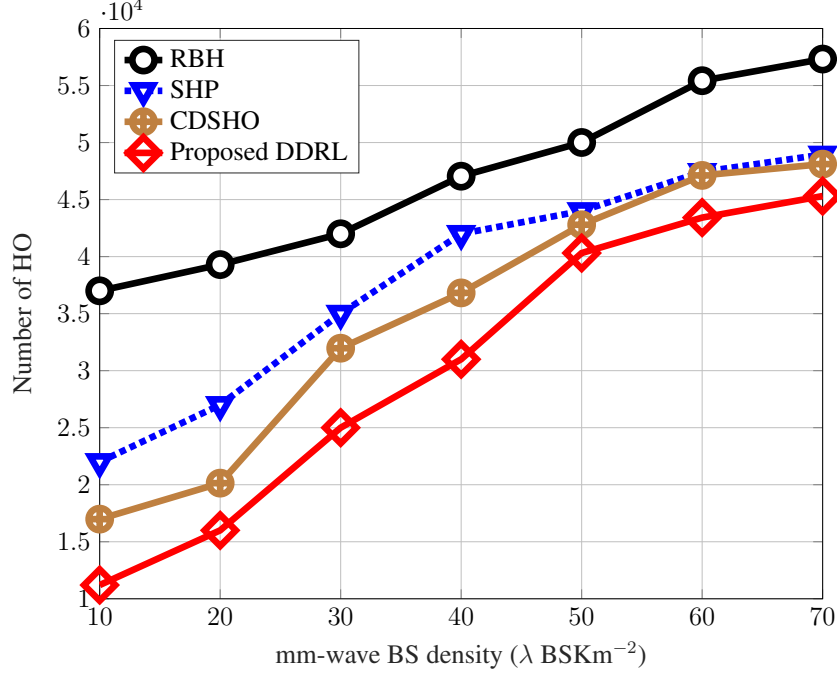


Figure 27: Frequency of HOs as a function of mm-wave BS density

$\gamma_{th}$ . Meanwhile, SHP incurs a high HO rate compared to the proposed model since it is based on the bandit algorithm. Though the bandit algorithm and DDRL works on the same principle to some extent (trial-and-error), DDRL uses offline learning while the bandit algorithm uses online learning which requires the online agent to undergo trial during learning while deciding at the same time. This contributes to the agent having a high HO rate before it learns the best action for the given state.

As expected, Fig. 28 shows that the proposed model outperforms other models in terms of system throughput, with a margin of 19% - 40%, 24% - 37%, and 25% - 43% compared to RBH, SHP, and CDSHO, respectively. In terms of throughput performance, RBH comes second. This is because during a HO decision, the RBH approach chooses the T-BS with the highest SNR, which results in high throughput. However, the high number of HO results in a low average throughput compared to the proposed DDRL because the average throughput incorporates a HO cost, which is a function of the number of HOs.

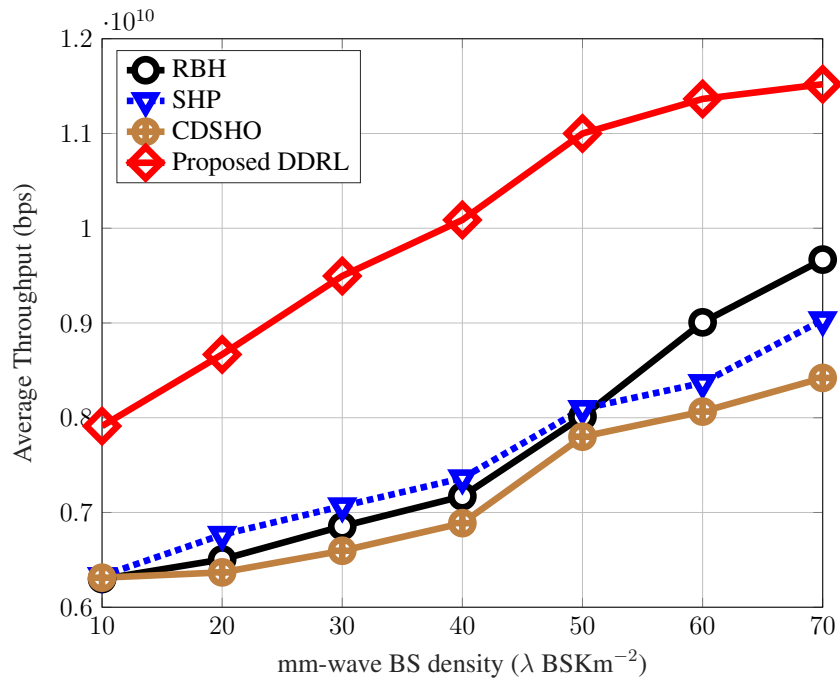


Figure 28: Average system throughput as a function of mm-wave BS density

## CHAPTER FIVE

### SUMMARY AND CONCLUSION

#### 5.1 Summary

This thesis proposed a flexible and dynamic solution to HO management in 5G mm-wave communication by demonstrating how ML-based algorithms can be used to develop an optimal solution to a problem at hand. In this regard, the proposed solution, referred to as intelligent HO management, was investigated considering various metrics, including its impact on reliability, the number of HOs, execution time, and throughput.

##### 5.1.1 Intelligent Handover Scheme

This thesis proposed a general framework for the optimization of HO decisions in UDNs. The proposed model was developed using  $Q$ -learning algorithms. The learning objective was to maximize user throughput while mitigating the negative impacts of HOs on the user QoE — to reduce intermittent connections and excessive HOs. A solution based on DRL was first developed, followed by a solution based on DDRL. The convergence efficiency of these two models was evaluated, upon which it was discovered that DRL tends to overestimate action value. Therefore, DDRL was chosen in this thesis for comparison with other state-of-the-art solutions.

##### 5.1.2 Benefits of the Developed Intelligent Handover Scheme

It was shown that the proposed intelligent solution can significantly enhance the system throughput and simultaneously reduce unnecessary HOs. The decision time of the proposed model was also evaluated for varying numbers of BSs and compared with other benchmark solutions. The results confirm the superiority of the proposed novel solution over the benchmarks. However, it is worth noting that at low UE speeds, both the proposed method and the benchmark solutions show similar performance in terms of throughput. This is because, at low speeds, there are fewer HOs and therefore, much intelligence is not required to decide whether to HO or not to minimize the HO cost. By and large, HO has little effect on stationary UEs.

This thesis also investigated the impact of HO interruption time  $t_d$  as a factor in reward design on the number of HOs for varying BS densities. The findings indicated that  $t_d$  has far less

contribution towards minimizing the number of HOs for small SNR threshold than anticipated and effect is observed for higher SNR threshold. This might be caused by the small contribution of  $t_d$  to the HO cost. Thus, for a given threshold  $\gamma_{th}$  and a fixed number of BSs, the number of HOs decreases slightly as  $t_d$  increases. Meanwhile, the number of HO events dramatically increases as  $\gamma_{th}$  increases.

Finally, it was demonstrated that the proposed model outperforms benchmark models under every given scenario. In particular, it outperforms benchmark models in terms of the number of HOs for a small number of BSs and by user throughput when applied to a large number of BSs. Hence, the proposed solution can be implemented in 5G UDNs where an enormous number of BSs are expected to be deployed.

## 5.2 Conclusion

Different HO management frameworks were examined in this thesis and a novel HO optimization scheme using ML presented and discussed.

Chapter one introduced the research area and outlined the background and rationale for the present study. It provided a brief overview of the key enablers for 5G and discussed the challenges posed by these enablers in mobility management. The chapter subsequently described the primary objective of the thesis, underscored the significance of the work, and ended with a presentation of the scope of the thesis.

Chapter two gave a comprehensive discussion of the general architecture of 5G networks, mobility and HO management in 5G, as well as a thorough review of various ML-based HO management schemes in 5G. The main takeaway of this review was that most of the existing ML solutions perform better than heuristic approaches.

Chapter three presented novel methods for HO management. First, the models used for simulation were presented, and in order to incorporate more realistic features, ray-tracing model was employed to capture real-world data in the simulation. Two novel solutions were developed, one based on clustering, used as a benchmark and the other based on RL as a proposed HO management solution. In terms of solutions based on RL, DRL was studied and its shortcomings presented, leading to the development and use of DDRL.

Chapter four presented the results obtained and offered extensive analyses of the results. Because DRL has more drawbacks compared to DDRL, this section concentrated on the model's analysis and performance based on benchmarks and DDRL. The results obtained from the DDRL-based HO management model resulted in higher throughput for a given trajectory, and thus DDRL was used in the proposed approach. Moreover, the proposed model has the advantage of being computationally adaptable and straightforward to implement in mm-wave 5G wireless systems. Simulations were used to validate the model and the findings indicate a significant improvement in throughput with fewer HOs. Additionally, the model was shown to be scalable in terms of the number of BSs and proven to be adaptable to current wireless network architectures. To the best of my knowledge, this is the first work that incorporated HO cost into the RL reward design. In addition, this is the first attempt at estimating UE location by mapping the UE position to the instantaneous SNR, which was used in defining the state of the DDRL model.

Optimization of HO decisions is a critical challenge in mm-wave mobile wireless communication. Consequently, it is critical to properly manage the decision regarding which BS the UEs hand over to, in order to reduce the adverse effects resulting from the use of mm-waves. The main contribution of this thesis is the development of an intelligent HO optimization model for 5G mm-wave communications. The proposed model is based on RL and can handle the trade-off between selecting the BS a) with the highest instantaneous SNR, and b) that can prolong the BS-UE connection, thereby resulting in higher QoS and QoE for UEs. This model can be applied to both mm-wave links and sub-6 GHz links.

Future works will focus on evolving HetNet architectures for 6G systems using THz communication links. While the vision of using THz frequency for SCs in future wireless systems would certainly improve the UE throughput and overall device capability, HO will still be a significant issue due to the propagation characteristics associated with high-frequency bands. Remarkably, the effect of high propagation loss associated with high frequencies necessitates the use of robust LOS even for short-distance communications in order to efficiently receive THz signals at the UE. If THz frequency is used by SCs in 6G networks, HO will become a more serious issue. As a result, future work will investigate more heterogeneous architectures involving sub-6 GHz, mm-wave, and THz cells, and suggest suitable intelligent HO schemes.

## REFERENCES

- 3GPP (2017a). 5G; Study on Scenarios and Requirements for Next Generation Access Technologies . Technical Report (TR) 138.913, 3rd Generation Partnership Project (3GPP), <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3144>. Version 14.2.0.
- 3GPP (2017b). LTE; General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access. Technical Specification (TS) 23.401, 3rd Generation Partnership Project (3GPP), <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=849>. Version 14.4.0.
- 3GPP (2018a). 5G;NR;Physical layer procedures for control. Technical Specification (TS) 38.213, 3rd Generation Partnership Project (3GPP), <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3215>. Version 15.3.0.
- 3GPP (2018b). 5G;NR;Radio Resource Control (RRC);Protocol specification. Technical Specification (TS) 38.331, 3rd Generation Partnership Project (3GPP), <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3197>. Version 15.3.0.
- 3GPP (2018e). NR; User Equipment (UE) procedures in idle mode and in RRC Inactive state. Technical Specification (TS) 38.304, 3rd Generation Partnership Project (3GPP), <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3192>. Version 15.1.0.
- 3GPP (2020). 5G;NR;Overall description:Stage-2. Technical Specification (TS) 38.300, 3rd Generation Partnership Project (3GPP), <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3191>. Version 15.9.0.
- 3GPP (2018d). LTE; Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description; Stage 2 . Technical Specification (TS) 36.300, 3rd Generation Partnership Project (3GPP),

<https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=2430>. Version 15.3.0.

3GPP (2019). 5G;System architecture for the 5G System (5GS). Technical Specification (TS) 23.501, 3rd Generation Partnership Project (3GPP), [https://www.etsi.org/deliver/etsi\\_ts/123500\\_123599/123501/15.05.00\\_60/ts\\_123501v150500p.pdf](https://www.etsi.org/deliver/etsi_ts/123500_123599/123501/15.05.00_60/ts_123501v150500p.pdf). Version 15.5.0.

3GPP (2018c). 5G;NR;Requirements for support of radio resource management. Technical Specification (TS) 38.133, 3rd Generation Partnership Project (3GPP), <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3204>. Version 15.2.0.

Abbasi, I. and Shahid Khan, A. (2018). A Review of Vehicle to Vehicle Communication Protocols for VANETs in the Urban Environment. *Future Internet*, 10(2):14.

Abubakar, A. I., Omeke, K. G., Ozturk, M., Hussain, S., and Imran, M. A. (2020). The Role of Artificial Intelligence Driven 5G Networks in COVID-19 Outbreak: Opportunities, Challenges, and Future Outlook. *Frontiers in Communications and Networks*, 1:4.

Agiwal, M., Roy, A., and Saxena, N. (2016). Next Generation 5G Wireless Networks: A Comprehensive Survey. *IEEE Communications Surveys Tutorials*, 18(3):1617–1655.

Ahmadi, S. (2019a). Chapter 1 - 5G network architecture. In Ahmadi, S., editor, *5G NR*, pages 1 – 194. Academic Press.

Ahmadi, S. (2019b). Chapter 2 - new radio access layer 2/3 aspects and system operation. In Ahmadi, S., editor, *5G NR*, pages 195 – 284. Academic Press.

Ahmed, A. A. and Alzahrani, A. A. (2019). A comprehensive survey on handover management for vehicular ad hoc network based on 5G mobile networks technology. *Transactions on Emerging Telecommunications Technologies*, 30(3):e3546.

Akpakwu, G. A., Silva, B. J., Hancke, G. P., and Abu-Mahfouz, A. M. (2018). A Survey on 5G Networks for the Internet of Things: Communication Technologies and Challenges. *IEEE Access*, 6:3619–3647.

- Ali, M., Mumtaz, S., Qaisar, S., and Naeem, M. (2017). Smart heterogeneous networks: a 5G paradigm. *Telecommunication Systems*, 66.
- Alkhateeb, A., Beltagy, I., and Alex, S. (2018). Machine learning for reliable mmwave systems: Blockage prediction and proactive handoff. In *2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pages 1055–1059.
- Alrabeiah, M. and Alkhateeb, A. (2020). Deep Learning for mmWave Beam and Blockage Prediction Using Sub-6 GHz Channels. *IEEE Transactions on Communications*, 68(9):5504–5518.
- Alrabeiah, M., Booth, J., Hredzak, A., and Alkhateeb, A. (2020a). ”ViWi Vision-Aided mmWave Beam Tracking: Dataset, Task, and Baseline Solutions”. *arXiv e-prints*, page arXiv:2002.02445.
- Alrabeiah, M., Hredzak, A., and Alkhateeb, A. (2020b). Millimeter Wave Base Stations with Cameras: Vision-Aided Beam and Blockage Prediction. In *2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring)*, pages 1–5.
- Alrabeiah, M., Hredzak, A., Liu, Z., and Alkhateeb, A. (2019). ViWi: A Deep Learning Dataset Framework for Vision-Aided Wireless Communications. *arXiv e-prints*, page arXiv:1911.06257.
- Alsaeedy, A. and Chong, E. (2018). Tracking Area Update and Paging in 5G Networks: a Survey of Problems and Solutions. *Mobile Networks and Applications*, 24.
- Antonioli, R., Parente, G., Silva, C., Sousa, D., Rodrigues, E., Maciel, T., and Cavalcanti, F. (2018). Dual Connectivity for LTE-NR Cellular Networks: Challenges and Open Issues. *Journal of Communication and Information Systems*, 33(1).
- Antón-Haro, C. and Mestre, X. (2019). Learning and Data-Driven Beam Selection for mmWave Communications: An Angle of Arrival-Based Approach. *IEEE Access*, 7:20404–20415.
- Arshad, R., ElSawy, H., Sorour, S., Al-Naffouri, T. Y., and Alouini, M. (2016a). Cooperative Handover Management in Dense Cellular Networks. In *2016 IEEE Global Communications Conference (GLOBECOM)*, pages 1–6.

- Arshad, R., ElSawy, H., Sorour, S., Al-Naffouri, T. Y., and Alouini, M. (2016b). Handover Management in 5G and Beyond: A Topology Aware Skipping Approach. *IEEE Access*, 4:9073–9081.
- Arshad, R., ElSawy, H., Sorour, S., Al-Naffouri, T. Y., and Alouini, M.-S. (2016). Handover management in dense cellular networks: A stochastic geometry approach. In *2016 IEEE international conference on communications (icc)*, pages 1–7. IEEE.
- Arulkumaran, K., Deisenroth, M. P., Brundage, M., and Bharath, A. A. (2017). Deep Reinforcement Learning: A Brief Survey. *IEEE Signal Processing Magazine*, 34(6):26–38.
- Asadi, A., Müller, S., Sim, G. H., Klein, A., and Hollick, M. (2018). FML: Fast Machine Learning for 5G mmWave Vehicular Communications. In *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications*, pages 1961–1969.
- Asadi, A., Wang, Q., and Mancuso, V. (2014). A Survey on Device-to-Device Communication in Cellular Networks. *IEEE Communications Surveys Tutorials*, 16(4):1801–1819.
- Astely, D., Dahlman, E., Fodor, G., Parkvall, S., and Sachs, J. (2013). LTE release 12 and beyond [Accepted From Open Call]. *IEEE Communications Magazine*, 51(7):154–160.
- Attiah, M., Isa, M., Zakaria, Z., Abdulhameed, M., Mohsen, M., and Ali, I. (2020). A survey of mmWave user association mechanisms and spectrum sharing approaches: an overview, open issues and challenges, future research trends. *Wireless Networks*, 26.
- Aykin, I., Akgun, B., Feng, M., and Krunz, M. (2020). MAMBA: A Multi-armed Bandit Framework for Beam Tracking in Millimeter-wave Systems. In *IEEE INFOCOM 2020 - IEEE Conference on Computer Communications*, pages 1469–1478.
- Balakrishnan, R. (2015). *Handover management in heterogeneous networks for 4G and beyond cellular systems*. PhD thesis, Georgia Institute of Technology.
- Bettstetter, C., Hartenstein, H., and Pérez-Costa, X. (2004). Stochastic properties of the random waypoint mobility model. *Wireless Networks*, 10(5):555–567.
- Bian, C., Yang, Y., Gao, F., and Li, G. Y. (2020). "FusionNet: Enhanced Beam Prediction for mmWave Communications Using Sub-6GHz Channel and A Few Pilots". *arXiv e-prints*, page arXiv:2009.02655.

- Björnson, E., Hoydis, J., and Sanguinetti, L. (2017). Massive MIMO Networks: Spectral, Energy, and Hardware Efficiency. *Foundations and Trends® in Signal Processing*, 11(3-4):154–655.
- Booth, M. B., Suresh, V., Michelusi, N., and Love, D. J. (2019). Multi-Armed Bandit Beam Alignment and Tracking for Mobile Millimeter Wave Communications. *IEEE Communications Letters*, 23(7):1244–1248.
- Briso-Rodríguez, C., Guan, K., Xuefeng, Y., and Kürner, T. (2017). Wireless Communications in Smart Rail Transportation Systems. *Wireless Communications and Mobile Computing*, 2017.
- Bui, N., Cesana, M., Hosseini, S. A., Liao, Q., Malanchini, I., and Widmer, J. (2017). A Survey of Anticipatory Mobile Networking: Context-Based Classification, Prediction Methodologies, and Optimization Techniques. *IEEE Communications Surveys Tutorials*, 19(3):1790–1821.
- Busari, S. A., Mumtaz, S., Al-Rubaye, S., and Rodriguez, J. (2018). 5G Millimeter-Wave Mobile Broadband: Performance and Challenges. *IEEE Communications Magazine*, 56(6):137–143.
- Chataut, R. and Akl, R. (2020). 5G and beyond Networks—Overview, Recent Trends, Challenges, and Future Research Direction. *Sensors (Basel, Switzerland)*, 20.
- Chen, M., Challita, U., Saad, W., Yin, C., and Debbah, M. (2019). Artificial Neural Networks-Based Machine Learning for Wireless Networks: A Tutorial. *IEEE Communications Surveys Tutorials*, 21(4):3039–3071.
- Chen, R., Long, W., Mao, G., and Li, C. (2018a). Development Trends of Mobile Communication Systems for Railways. *IEEE Communications Surveys Tutorials*, 20(4):3131–3141.
- Chen, Y., Cheng, W., and Wang, L. (2018b). Learning-assisted beam search for indoor mmWave networks. In *2018 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, pages 320–325.

- Chiang, H., Chen, K., Rave, W., Marandi, M. K., and Fettweis, G. (2020). Multi-UAV mmWave Beam Tracking using Q-Learning and Interference Mitigation. In *2020 IEEE International Conference on Communications Workshops (ICC Workshops)*, pages 1–7.
- Christensen, B. and Knape, O. (2016). Optimization of algorithms for mobility in cellular systems. Student Paper.
- Da Silva, I. L., Mildh, G., Säily, M., and Hailu, S. (2016). A novel state model for 5G Radio Access Networks. In *2016 IEEE International Conference on Communications Workshops (ICC)*, pages 632–637.
- Dahlman, E., Parkvall, S., and Sköld, J. (2018a). Chapter 14 - scheduling. In Dahlman, E., Parkvall, S., and Sköld, J., editors, *5G NR: the Next Generation Wireless Access Technology*, pages 275 – 299. Academic Press.
- Dahlman, E., Parkvall, S., and Sköld, J. (2018b). Chapter 6 - Radio-Interface Architecture. In Dahlman, E., Parkvall, S., and Sköld, J., editors, *5G NR: the Next Generation Wireless Access Technology*, pages 73 – 102. Academic Press.
- Demarchou, E., Psomas, C., and Krikidis, I. (2018). Mobility management in ultra-dense networks: Handover skipping techniques. *IEEE Access*, 6:11921–11930.
- Dey, A. (2016). Machine learning algorithms: a review. *International Journal of Computer Science and Information Technologies*, 7(3):1174–1179.
- Dias, M., Klautau, A., González-Prelcic, N., and Heath, R. W. (2019). Position and LIDAR-Aided mmWave Beam Selection using Deep Learning. In *2019 IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pages 1–5.
- Ericsson (2020). Ericsson Mobility Report. Technical report, Ericsson, Available at <https://www.ericsson.com/49da93/assets/local/mobility-report/documents/2020/june2020-ericsson-mobility-report.pdf> [Accessed 11. 09. 2020].
- Fourati, H., Maaloul, R., and Chaari, L. (2020). A survey of 5G network systems: challenges and machine learning approaches. *International Journal of Machine Learning and Cybernetics*, pages 1–47.

- Gandotra, P. and Jha, R. K. (2016). Device-to-Device Communication in Cellular Networks: A Survey. *Journal of Network and Computer Applications*, 71:99 – 117.
- Ge, X., Tu, S., Mao, G., Wang, C., and Han, T. (2016). 5g ultra-dense cellular networks. *IEEE Wireless Communications*, 23(1):72–79.
- Ghafoor, K. Z., Guizani, M., Kong, L., Maghdid, H. S., and Jasim, K. F. (2020). Enabling Efficient Coexistence of DSRC and C-V2X in Vehicular Networks. *IEEE Wireless Communications*, 27(2):134–140.
- Guo, D., Tang, L., Zhang, X., and Liang, Y. C. (2020). Joint optimization of handover control and power allocation based on multi-agent deep reinforcement learning. *IEEE Transactions on Vehicular Technology*, 69(11):13124–13138.
- Hamed, A. and Rao, R. (2018). Spatial spectral and energy efficiencies of cellular networks limited by co-channel interference and path loss in Nakagami-m fading environment. *EURASIP Journal on Wireless Communications and Networking*, 2018.
- Heng, Y. and Andrews, J. G. (2019). Machine Learning-Assisted Beam Alignment for mmWave Systems. In *2019 IEEE Global Communications Conference (GLOBECOM)*, pages 1–6.
- Huq, K. M. S., Busari, S. A., Rodriguez, J., Frasca, V., Bazzi, W., and Sicker, D. C. (2019a). Terahertz-Enabled Wireless System for Beyond-5G Ultra-Fast Networks: A Brief Survey. *IEEE Network*, 33(4):89–95.
- Huq, K. M. S., Busari, S. A., Rodriguez, J., Frasca, V., Bazzi, W., and Sicker, D. C. (2019b). Terahertz-Enabled Wireless System for Beyond-5G Ultra-Fast Networks: A Brief Survey. *IEEE Network*, 33(4):89–95.
- Hussain, F., Hassan, S. A., Hussain, R., and Hossain, E. (2020). Machine Learning for Resource Management in Cellular and IoT Networks: Potentials, Current Solutions, and Open Challenges. *IEEE Communications Surveys Tutorials*, 22(2):1251–1275.
- Islam, S. M. R., Kwak, D., Kabir, M. H., Hossain, M., and Kwak, K. (2015). The Internet of Things for Health Care: A Comprehensive Survey. *IEEE Access*, 3:678–708.

- Jagyasi, D. and Coupechoux, M. (2020). DNN Based Beam Selection in mmW Heterogeneous Networks. In *International Conference on Network Games, Control and Optimisation (Netgcoop)*, Cargèse, France.
- Jameel, F., Faisal, Haider, M. A. A., and Butt, A. A. (2017). Massive MIMO: A survey of recent advances, research issues and future directions. In *2017 International Symposium on Recent Advances in Electrical Engineering (RAEE)*, pages 1–6.
- Khosravi, S., Ghadikolaei, H. S., and Petrova, M. (2020b). Learning-based Load Balancing Handover in Mobile Millimeter Wave Networks. *arXiv e-prints*, page arXiv:2011.01420.
- Khosravi, S., Ghadikolaei, H. S., and Petrova, M. (2020a). Learning-based Handover in Mobile Millimeter-wave Networks. *arXiv e-prints*, page arXiv:2003.11009.
- Kim, J., Kim, D., and Choi, S. (2017). 3GPP SA2 architecture and functions for 5G mobile communication system. *ICT Express*, 3(1):1 – 8.
- Kim, Y., Lee, H., Hwang, P., Patro, R. K., Lee, J., Roh, W., and Cheun, K. (2016). Feasibility of Mobile Cellular Communications at Millimeter Wave Frequency. *IEEE Journal of Selected Topics in Signal Processing*, 10(3):589–599.
- Klaine, P. V., Imran, M. A., Onireti, O., and Souza, R. D. (2017). A Survey of Machine Learning Techniques Applied to Self-Organizing Cellular Networks. *IEEE Communications Surveys Tutorials*, 19(4):2392–2431.
- Klautau, A., González-Prelcic, N., and Heath, R. W. (2019). LIDAR Data for Deep Learning-Based mmWave Beam-Selection. *IEEE Wireless Communications Letters*, 8(3):909–912.
- Klus, R., Klus, L., Solomitchii, D., Valkama, M., and Talvitie, J. (2020). Deep Learning Based Localization and HO Optimization in 5G NR Networks. In *2020 International Conference on Localization and GNSS (ICL-GNSS)*, pages 1–6.
- Koda, N., Nakashima, K., Yamamoto, K., Nishio, T., and Morikura, M. (2020a). Handover Management for mmWave Networks With Proactive Performance Prediction Using Camera Images and Deep Reinforcement Learning. *IEEE Transactions on Cognitive Communications and Networking*, 6(2):802–816.

- Koda, Y., Nakashima, K., Yamamoto, K., Nishio, T., and Morikura, M. (2020b). Cooperative Sensing in Deep RL-Based Image-to-Decision Proactive Handover for mmWave Networks. In *2020 IEEE 17th Annual Consumer Communications Networking Conference (CCNC)*, pages 1–6.
- Lai, W. K., Shieh, C.-S., Chou, F.-S., Hsu, C.-Y., and Shen, M.-H. (2020). Handover Management for D2D Communication in 5G Networks. *Applied Sciences*, 10(12):4409.
- Lashari, M. H., Memon, A. A., Shah, S. A. A., Nenwani, K., and Shafqat, F. (2018). IoT Based Poultry Environment Monitoring System. In *2018 IEEE International Conference on Internet of Things and Intelligence System (IOTAIS)*, pages 1–5.
- Lee, S., Bae, M., and Kim, H. (2017). Future of IoT networks: A survey. *Applied Sciences*, 7(10):1072.
- Li, X., Guo, D., Grosspietsch, J., Yin, H., and Wei, G. (2016). Maximizing mobile coverage via optimal deployment of base stations and relays. *IEEE Transactions on Vehicular Technology*, 65(7):5060–5072.
- Li, X., Zhou, R., Angela Zhang, Y.-J., Jiao, L., and Li, Z. (2020). Smart vehicular communication via 5G mmWaves. *Computer Networks*, 172:107173.
- Li, Y., Pateromichelakis, E., Vucic, N., Luo, J., Xu, W., and Caire, G. (2017). Radio Resource Management Considerations for 5G Millimeter Wave Backhaul and Access Networks. *IEEE Communications Magazine*, 55(6):86–92.
- Li, Z., Ma, X., Chen, W., Kuang, N., and Zhang, B. (2019). Neural Network Enhanced Analog Beam Selection Scheme for Terahertz Systems. In *2019 IEEE/CIC International Conference on Communications Workshops in China (ICCC Workshops)*, pages 136–141.
- Liang, J.-M., Chang, P.-Y., Chen, J.-J., Huang, C.-F., and Tseng, Y.-C. (2018). Energy-efficient DRX scheduling for D2D communication in 5G networks. *Journal of Network and Computer Applications*, 116:53 – 64.
- Lin, C., Kao, W., Zhan, S., and Lee, T. (2019). BsNet: A Deep Learning-Based Beam Selection Method for mmWave Communications. In *2019 IEEE 90th Vehicular Technology Conference (VTC2019-Fall)*, pages 1–6.

- Liu, G. and Jiang, D. (2016). 5G: Vision and requirements for mobile communication system towards year 2020. *Chinese Journal of Engineering*, 2016(2016):8.
- Long, N. (2018). Resource Allocation for Energy Efficiency in 5G Wireless Networks. *EAI Endorsed Transactions on Industrial Networks and Intelligent Systems*, 5:154832.
- Long, Y., Chen, Z., Fang, J., and Tellambura, C. (2018). Data-Driven-Based Analog Beam Selection for Hybrid Beamforming Under mm-Wave Channels. *IEEE Journal of Selected Topics in Signal Processing*, 12(2):340–352.
- Luong, N. C., Hoang, D. T., Gong, S., Niyato, D., Wang, P., Liang, Y., and Kim, D. I. (2019). Applications of Deep Reinforcement Learning in Communications and Networking: A Survey. *IEEE Communications Surveys Tutorials*, 21(4):3133–3174.
- Ma, X., Chen, Z., Li, Z., Chen, W., and Liu, K. (2019). Low Complexity Beam Selection Scheme for Terahertz Systems: A Machine Learning Approach. In *2019 IEEE International Conference on Communications Workshops (ICC Workshops)*, pages 1–6.
- Mahira, A. G. and Subhedar, M. S. (2017). Handover decision in wireless heterogeneous networks based on feedforward artificial neural network. In *Computational Intelligence in Data Mining*, pages 663–669. Springer.
- Meng, Y., Zhang, W., Zhu, H., and Shen, X. S. (2018). Securing Consumer IoT in the Smart Home: Architecture, Challenges, and Countermeasures. *IEEE Wireless Communications*, 25(6):53–59.
- Mezzavilla, M., Goyal, S., Panwar, S., Rangan, S., and Zorzi, M. (2016). An MDP model for optimal handover decisions in mmwave cellular networks. In *2016 European conference on networks and communications (EuCNC)*, pages 100–105. IEEE.
- Mocrii, D., Chen, Y., and Musilek, P. (2018). Iot-based smart homes: A review of system architecture, software, communications, privacy and security. *Internet of Things*, 1-2:81 – 98.
- Mollet, M., Ozturk, M., Kisangiri, M., Kaijage, S., Onireti, O., Imran, M. A., and Abbasi, Q. H. (2019). Handover management in dense networks with coverage prediction from sparse networks. In *2019 IEEE Wireless Communications and Networking Conference Workshop (WCNCW)*, pages 1–6.

- Mollet, M. S., Abubakar, A. I., Ozturk, M., Kaijage, S., Kisangiri, M., Zoha, A., Imran, M. A., and Abbasi, Q. H. (2020b). Intelligent handover decision scheme using double deep reinforcement learning. *Physical Communication*, 42:101133.
- Mollet, M. S., Abubakar, A. I., Ozturk, M., Kaijage, S. F., Kisangiri, M., Hussain, S., Imran, M. A., and Abbasi, Q. H. (2021b). A Survey of Machine Learning Applications to Handover Management in 5G and Beyond. *IEEE Access*, 9:45770–45802.
- Mollet, M. S., Kaijage, S., and Kisangiri, M. (2021a). Deep Reinforcement Learning based Handover Management for Millimeter Wave Communication. *International Journal of Advanced Computer Science and Applications*, 12(2).
- Mollet, M. S., Kaijage, S., Kisangiri, M., Imran, M. A., and Abbasi, Q. H. (2020a). Multi-user position based on trajectories-aware handover strategy for base station selection with multi-agent learning. In *2020 IEEE International Conference on Communications Workshops (ICC Workshops)*, pages 1–6.
- Morocho Cayamcela, M. E. and Lim, W. (2018). Artificial Intelligence in 5G Technology: A Survey. In *2018 International Conference on Information and Communication Technology Convergence (ICTC)*, pages 860–865.
- Mumtaz, S., Jornet, J., Aulin, J., Gerstacker, W., Dong, X., and ai, b. (2017). Terahertz Communication for Vehicular Networks. *IEEE Transactions on Vehicular Technology*, 66:5617–5625.
- Mustafa, H. A., Shakir, M. Z., Imran, M. A., Imran, A., and Tafazolli, R. (2015). Coverage gain and device-to-device user density: Stochastic geometry modeling and analysis. *IEEE Communications Letters*, 19(10):1742–1745.
- Nishio, T., Koda, Y., Park, J., Bennis, M., and Doppler, K. (2020). When Wireless Communications Meet Computer Vision in Beyond 5G. *arXiv e-prints*, page arXiv:2010.06188.
- Niu, Y., Li, Y., Jin, D., Su, L., and Vasilakos, A. V. (2015). A survey of millimeter wave communications (mmWave) for 5G: opportunities and challenges. *Wireless networks*, 21(8):2657–2676.
- Nokia (2017). 5G deployment below 6 GHz Ubiquitous coverage for critical communication and massive IoT. White paper, Nokia, Available at

[https://www.rrt.lt/wp-content/uploads/2018/10/Nokia\\_5G\\_Deployment\\_below\\_6GHz\\_White\\_Paper\\_EN.pdf](https://www.rrt.lt/wp-content/uploads/2018/10/Nokia_5G_Deployment_below_6GHz_White_Paper_EN.pdf) [Accessed 1. 10. 2020].

- Koda, Y., Nakashima, K., Yamamoto, K., Nishio, T., and Morikura, M. (2019). End-to-End Learning of Proactive Handover Policy for Camera-Assisted mmWave Networks Using Deep Reinforcement Learning. *arXiv preprint arXiv:1904.04585*.
- Oguma, Y., Nishio, T., Yamamoto, K., and Morikura, M. (2016). Proactive Handover Based on Human Blockage Prediction Using RGB-D Cameras for mmWave Communications. *IEICE Transactions on Communications*, E99.B(8):1734–1744.
- Omeke, K. G., Mollel, M. S., Ozturk, M., Ansari, S., Zhang, L., Abbasi, Q. H., and Imran, M. A. (2021). Dekcs: A dynamic clustering protocol to prolong underwater sensor networks. *IEEE Sensors Journal*, 21(7):9457–9464.
- Pal, D., Papasratorn, B., Chutimaskul, W., and Funilkul, S. (2019). Embracing the Smart-Home Revolution in Asia by the Elderly: An End-User Negative Perception Modeling. *IEEE Access*, 7:38535–38549.
- Rangan, S., Rappaport, T. S., and Erkip, E. (2014). Millimeter-Wave Cellular Wireless Networks: Potentials and Challenges. *Proceedings of the IEEE*, 102(3):366–385.
- Rappaport, T. S., Sun, S., Mayzus, R., Zhao, H., Azar, Y., Wang, K., Wong, G. N., Schulz, J. K., Samimi, M., and Gutierrez, F. (2013). Millimeter wave mobile communications for 5g cellular: It will work! *IEEE access*, 1:335–349.
- Rappaport, T. S., Xing, Y., MacCartney, G. R., Molisch, A. F., Mellios, E., and Zhang, J. (2017). Overview of millimeter wave communications for fifth-generation (5G) wireless networks—with a focus on propagation models. *IEEE Transactions on Antennas and Propagation*, 65(12):6213–6230.
- Ren, Z., Wu, S., and Zhao, A. (2018). Coexist Design of Sub-6GHz and Millimeter-Wave Antennas for 5G Mobile Terminals. In *2018 International Symposium on Antennas and Propagation (ISAP)*, pages 1–2.
- Rezaie, S., Manchón, C. N., and de Carvalho, E. (2020). Location- and Orientation-Aided Millimeter Wave Beam Selection Using Deep Learning. In *ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*, pages 1–6.

- Roh, W., Seol, J., Park, J., Lee, B., Lee, J., Kim, Y., Cho, J., Cheun, K., and Aryanfar, F. (2014). Millimeter-wave beamforming as an enabling technology for 5G cellular communications: theoretical feasibility and prototype results. *IEEE Communications Magazine*, 52(2):106–113.
- Rummery, G. A. and Niranjan, M. (1994). *On-line Q-learning using connectionist systems*, volume 37. University of Cambridge, Department of Engineering Cambridge, England.
- Sana, M., De Domenico, A., Strinati, E. C., and Clemente, A. (2020a). Multi-Agent Deep Reinforcement Learning For Distributed Handover Management In Dense MmWave Networks. In *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 8976–8980.
- Sana, M., De Domenico, A., Yu, W., Lohan, Y., and Calvanese Strinati, E. (2020b). Multi-Agent Reinforcement Learning for Adaptive User Association in Dynamic mmWave Networks. *IEEE Transactions on Wireless Communications*, 19(10):6520–6534.
- Schniter, P. and Sayeed, A. (2014). Channel estimation and precoder design for millimeter-wave communications: The sparse way. In *2014 48th Asilomar Conference on Signals, Systems and Computers*, pages 273–277. IEEE.
- Schumacher, A., Merz, R., and Burg, A. (2019). 3.5 GHz Coverage Assessment with a 5G Testbed. In *2019 IEEE 89th Vehicular Technology Conference (VTC2019-Spring)*, pages 1–6.
- Shafi, M., Molisch, A. F., Smith, P. J., Haustein, T., Zhu, P., De Silva, P., Tufvesson, F., Ben-jebbour, A., and Wunder, G. (2017). 5G: A Tutorial Overview of Standards, Trials, Challenges, Deployment, and Practice. *IEEE Journal on Selected Areas in Communications*, 35(6):1201–1221.
- Shayea, I., Ergen, M., Hadri Azmi, M., Aldirmaz Çolak, S., Nordin, R., and Daradkeh, Y. I. (2020). Key Challenges, Drivers and Solutions for Mobility Management in 5G Networks: A Survey. *IEEE Access*, 8:172534–172552.
- Sim, G. H., Klos, S., Asadi, A., Klein, A., and Hollick, M. (2018). An Online Context-Aware Machine Learning Algorithm for 5G mmWave Vehicular Communications. *IEEE/ACM Transactions on Networking*, 26(6):2487–2500.

- Sim, M. S., Lim, Y., Park, S. H., Dai, L., and Chae, C. (2020). Deep Learning-Based mmWave Beam Selection for 5G NR/6G With Sub-6 GHz Channel Information: Algorithms and Prototype Validation. *IEEE Access*, 8:51634–51646.
- Srinidhi, N., Dilip Kumar, S., and Venugopal, K. (2019). Network optimizations in the Internet of Things: A review. *Engineering Science and Technology, an International Journal*, 22(1):1 – 21.
- Sultan, K., Ali, H., and Zhang, Z. (2018). Big Data Perspective and Challenges in Next Generation Networks. *Future Internet*, 10:56.
- Sun, L., Hou, J., and Shu, T. (2019a). Optimal Handover Policy for mmWave Cellular Networks: A Multi-Armed Bandit Approach. In *2019 IEEE Global Communications Conference (GLOBECOM)*, pages 1–6.
- Sun, L., Hou, J., and Shu, T. (2020). Spatial and Temporal Contextual Multi-Armed Bandit Handovers in Ultra-Dense mmWave Cellular Networks. *IEEE Transactions on Mobile Computing*, pages 1–1.
- Sun, Y., Feng, G., Qin, S., Liang, Y., and Yum, T. P. (2018). The SMART Handoff Policy for Millimeter Wave Heterogeneous Cellular Networks. *IEEE Transactions on Mobile Computing*, 17(6):1456–1468.
- Sun, Y., Peng, M., Zhou, Y., Huang, Y., and Mao, S. (2019b). Application of Machine Learning in Wireless Networks: Key Techniques and Open Issues. *IEEE Communications Surveys Tutorials*, 21(4):3072–3108.
- Sutton, R. S. and Barto, A. G. (2018). *Reinforcement Learning: An Introduction*. A Bradford Book, Cambridge, MA, USA.
- Tabassum, H., Salehi, M., and Hossain, E. (2019). Fundamentals of Mobility-Aware Performance Characterization of Cellular Networks: A Tutorial. *IEEE Communications Surveys Tutorials*, 21(3):2288–2308.
- Tauqir, H. P. and Habib, A. (2019). Deep Learning Based Beam Allocation in Switched-Beam Multiuser Massive MIMO Systems. In *2019 Second International Conference on Latest trends in Electrical Engineering and Computing Technologies (INTELLECT)*, pages 1–5.

- Tayyab, M., Gelabert, X., and Jäntti, R. (2019). A Survey on Handover Management: From LTE to NR. *IEEE Access*, 7:118907–118930.
- Tian, Y., Pan, G., and Alouini, M.-S. (2020). Applying Deep-Learning-Based Computer Vision to Wireless Communications: Methodologies, Opportunities, and Challenges. *arXiv e-prints*, page arXiv:2006.05782.
- Usman, M. A., Philip, N. Y., and Politis, C. (2019). 5G Enabled Mobile Healthcare for Ambulances. In *2019 IEEE Globecom Workshops (GC Wkshps)*, pages 1–6.
- Uwaechia, A. N. and Mahyuddin, N. M. (2020). A Comprehensive Survey on Millimeter Wave Communications for Fifth-Generation Wireless Networks: Feasibility and Challenges. *IEEE Access*, 8:62367–62414.
- Va, V., Shimizu, T., Bansal, G., and Heath, R. W. (2019). Online Learning for Position-Aided Millimeter Wave Beam Training. *IEEE Access*, 7:30507–30526.
- van Hasselt, H., Guez, A., and Silver, D. (2015). Deep reinforcement learning with double Q-learning. *CoRR*, abs/1509.06461.
- Van Huynh, N., Nguyen, D. N., Hoang, D. T., and Dutkiewicz, E. (2020). ”Optimal Beam Association for High Mobility mmWave Vehicular Networks: Lightweight Parallel Reinforcement Learning Approach”. *arXiv e-prints*, page arXiv:2005.00694.
- Van Quang, B., Prasad, R. V., and Niemegeers, I. (2012). A survey on handoffs—lessons for 60 GHz based wireless systems. *IEEE Communications Surveys & Tutorials*, 14(1):64–86.
- Wang, J., Jiang, C., Zhang, H., Ren, Y., Chen, K.-C., and Hanzo, L. (2019a). Thirty Years of Machine Learning: The Road to Pareto-Optimal Wireless Networks. *arXiv e-prints*, page arXiv:1902.01946.
- Wang, R., Onireti, O., Zhang, L., Imran, M. A., Ren, G., Qiu, J., and Tian, T. (2019b). Reinforcement Learning Method for Beam Management in Millimeter-Wave Networks. In *2019 UK/ China Emerging Technologies (UCET)*, pages 1–4.
- Wang, T., Wang, S., and Zhou, Z. (2019c). Machine learning for 5G and beyond: From model-based to data-driven mobile wireless networks. *China Communications*, 16(1):165–175.

- Wang, Y., Klautau, A., Ribero, M., Narasimha, M., and Heath, R. W. (2018). Mmwave vehicular beam training with situational awareness by machine learning. In *2018 IEEE Globecom Workshops (GC Wkshps)*, pages 1–6.
- Wu, Y. and Ai, B. (2019). A learning-based beam training method for multi-user milli-meter wave systems. In *IET 8th International Conference on Wireless, Mobile Multimedia Networks*, pages 38–42.
- Xiao, M., Mumtaz, S., Huang, Y., Dai, L., Li, Y., Matthaiou, M., Karagiannidis, G. K., Björnson, E., Yang, K., I, C., and Ghosh, A. (2017). Millimeter Wave Communications for Future Mobile Networks. *IEEE Journal on Selected Areas in Communications*, 35(9):1909–1935.
- Xu, C., Liu, S., Zhang, C., Huang, Y., and Yang, L. (2020a). Joint User Scheduling and Beam Selection in mmWave Networks Based on Multi-Agent Reinforcement Learning. In *2020 IEEE 11th Sensor Array and Multichannel Signal Processing Workshop (SAM)*, pages 1–5.
- Xu, L. D., He, W., and Li, S. (2014). Internet of Things in Industries: A Survey. *IEEE Transactions on Industrial Informatics*, 10(4):2233–2243.
- Xu, W., Gao, F., Jin, S., and Alkhateeb, A. (2020b). 3D Scene Based Beam Selection for mmWave Communications. *IEEE Wireless Communications Letters*, pages 1–1.
- Yajnanarayana, V., Rydén, H., and Hévízi, L. (2020). 5G Handover using Reinforcement Learning. In *2020 IEEE 3rd 5G World Forum (5GWF)*, pages 349–354.
- Yan, L., Ding, H., Zhang, L., Liu, J., Fang, X., Fang, Y., Xiao, M., and Huang, X. (2019). Machine Learning-Based Handovers for Sub-6 GHz and mmWave Integrated Vehicular Networks. *IEEE Transactions on Wireless Communications*, 18(10):4873–4885.
- Yang, H., Alphones, A., Xiong, Z., Niyato, D., Zhao, J., and Wu, K. (2020). Artificial-Intelligence-Enabled Intelligent 6G Networks. *IEEE Network*, 34(6):272–280.
- Yang, Y., Gao, Z., Ma, Y., Cao, B., and He, D. (2020). Machine Learning Enabling Analog Beam Selection for Concurrent Transmissions in Millimeter-Wave V2V Communications. *IEEE Transactions on Vehicular Technology*, 69(8):9185–9189.

- Yang, Y., He, Y., He, D., Gao, Z., and Luo, Y. (2019). Machine Learning Based Analog Beam Selection for 5G mmWave Small Cell Networks. In *2019 IEEE Globecom Workshops (GC Wkshps)*, pages 1–5.
- Yilmaz, T., Gokkoca, G., and Akan, O. B. (2016). *Millimetre Wave Communication for 5G IoT Applications*, pages 37–53. Springer International Publishing, Cham.
- Yue, G., Yu, D., Cheng, L., Lv, Q., Luo, Z., Li, Q., Luo, J., and He, X. (2019). Millimeter-Wave System for High-Speed Train Communications Between Train and Tracksides: System Design and Channel Measurements. *IEEE Transactions on Vehicular Technology*, 68(12):11746–11761.
- Zaidi, S. M. A., Manalastas, M., Farooq, H., and Imran, A. (2020). Mobility Management in Emerging Ultra-Dense Cellular Networks: A Survey, Outlook, and Future Research Directions. *IEEE Access*, 8:183505–183533.
- Zanella, A., Bui, N., Castellani, A., Vangelista, L., and Zorzi, M. (2014). Internet of Things for Smart Cities. *IEEE Internet of Things Journal*, 1(1):22–32.
- Zang, S., Bao, W., Yeoh, P. L., Chen, H., Lin, Z., Vucetic, B., and Li, Y. (2017). Mobility handover optimization in millimeter wave heterogeneous networks. In *2017 17th International symposium on communications and information technologies (ISCIT)*, pages 1–6. IEEE.
- Zappone, A., Di Renzo, M., and Debbah, M. (2019). Wireless Networks Design in the Era of Deep Learning: Model-Based, AI-Based, or Both? *IEEE Transactions on Communications*, 67(10):7331–7376.
- Zhang, C., Patras, P., and Haddadi, H. (2019). Deep Learning in Mobile and Wireless Networking: A Survey. *IEEE Communications Surveys Tutorials*, 21(3):2224–2287.
- Zhang, J. and Letaief, K. B. (2020). Mobile Edge Intelligence and Computing for the Internet of Vehicles. *Proceedings of the IEEE*, 108(2):246–261.
- Öztürk, M. (2020). *Cognitive networking for next generation of cellular communication systems*. PhD thesis, University of Glasgow.

## APPENDICES

### Appendix 1: Python codes for System Model

#### Deploying BS using PPP

```
import numpy as np
import scipy.stats
import matplotlib.pyplot as plt
#Simulation window parameters
xMin=0;xMax=1;
yMin=0;yMax=1;
xDelta=xMax-xMin;yDelta=yMax-yMin; #rectangle dimensions
areaTotal=xDelta*yDelta;
#intensity (ie mean density) of the Poisson process
def bs_generation(lambda0 = 10 ):
    result_out = np.zeros((lambda0,2))
    #Point process parameters
    np.random.seed(4)
    #Simulate Poisson point process
    #Poisson number of points
    numbPoints = scipy.stats.poisson( lambda0*areaTotal ).rvs()
    #x coordinates of Poisson points
    xx = xDelta*scipy.stats.uniform.rvs(0,100,((numbPoints,1)))+xMin
    #y coordinates of Poisson points
    yy = yDelta*scipy.stats.uniform.rvs(0,100,((numbPoints,1)))+yMin
    if len(xx) > lambda0:
        #randomRows = np.random.randint(len(xx), size=lambda0)
        randomRows = np.random.choice(len(xx), lambda0, replace=False)
        for i in range (lambda0):
            result_out[i,0]= int((xx[randomRows[i]]))
            result_out[i,1]= int((yy[randomRows[i]]))
    else:
        for i in range (xx):
            result_out[i,0]= int((xx[i]))
            result_out[i,1]= int((yy[i]))

    return result_out
```

## Pathloss Model

```
import numpy as np
def pathlossmmwave(distance = 1):
    alpha1 = 61.4
    beta1 = 2
    mu, sigma = 0, 5.8 # mean and standard deviation
    path_loss = alpha1 + 10*beta1*np.log10(distance) + np.random.normal(mu
                                                                    , sigma, 1)
    return path_loss
```

## Appendix 2: Python codes for Agent $Q$ -Neural Network

### Agent $Q$ -Network Function

```
# -*- coding: utf-8 -*-
"""
Created on Fri Aug 16 13:43:21 2019

@author: 2427060M
"""

from keras.models import Sequential
from keras.layers import Dense
from keras.optimizers import Adam

class DeepNeuralNetowk:

    def __init__(self, observation_space, action_space):
        self.model = Sequential()
        self.model.add(Dense(512, input_shape = (observation_space, ),
                                activation = "relu"))
        self.model.add(Dense(256, activation = "relu"))
        self.model.add(Dense(128, activation = "relu" ))
        self.model.add(Dense(64, activation = "relu"))
        self.model.add(Dense(action_space, activation = "linear"))
        self.model.compile(loss = "mse", optimizer = Adam(lr = 0.001))
        self.model.summary()
```

## Appendix 3: Python codes for Main function

### Main function

```
# -*- coding: utf-8 -*-
"""
Created on Fri Aug 16 14:53:08 2019
@author: Michael Mollel
Env = Tensorflow 1.15, Keras 2.2.4
Functions = to run this file you also need to import files found in the
*https://github.com/msamwelmollel/HOMAN_V1*
"""
import numpy as np
import argparse
from data_models.ddqn_data_model import DDQNTrainer
from import_trajectory_information import process_sinr
from keras.utils import to_categorical
from data_models.reward.reward_logger1 import RewardLogger
from data_models.user_log.user_log1 import print_data
from data_models.average_module import average_data
from data_models.handover_performance import handover_performance as hper
from data_models.handover_performance import handover as handover_report

def solution_v1(user, delay_cost, min_SNR, bs_intensity):

    ENV_NAME = 'handover'
    antennas = bs_intensity
    intensity = bs_intensity

    model_path = './data_models/models/model_save_intensity_'+str(intensity
                                                                    )+'_min_sinr_'+str(min_SNR)+'
                                                                    _delay_cost_'+str(delay_cost)+'
                                                                    _user_'+str(user)+'_.h5'

    SNR, SNR_h, POSITION = process_sinr(intensity, antennas, user)
    action_space = SNR.shape[1]
    INPUT_SHAPE = 2 * action_space
    data_model=DDQNTrainer(INPUT_SHAPE, action_space, model_path)
```

```

reward_path= "./data_models/reward/intensity_"+str(intensity)+"_user_"+
              str(user)+"_reward_delay_"+str(
              delay_cost)+"_SNR_"+str(min_SNR)+
              ".csv"

position_path= "./data_models/reward/intensity_"+str(intensity)+"_user_
              "+str(user)+"_position_delay_"+
              str(delay_cost)+"_SNR_"+str(
              min_SNR)+".csv"

reward_logger = RewardLogger(ENV_NAME,reward_path,position_path)

run = 0
total_step = 0

for j in range(1000):
    terminal = False
    step, score, hd, lf = 0, 0, 0, 0
    run += 1
    action= np.argmax(SNR[0,:])
    current_state=np.reshape(np.append(SNR[0,:], to_categorical(action,
        num_classes=SNR.shape[1])), [
        1, INPUT_SHAPE])
    current_state_h=np.reshape(np.append(SNR_h[0,:], to_categorical(
        action, num_classes=SNR_h.
        shape[1])), [1, INPUT_SHAPE])

    for i in range(0,1000):
        total_step += 1
        step += 1
        if i == 999:
            terminal = True

        if current_state[:, action] >= min_SNR:
            action = action
        else:
            action = data_model.move(current_state)
            hd = hd + 1

```

```

next_state = np.reshape(np.append(SNR[i+1,:], to_categorical(
    action, num_classes=SNR.
    shape[1])), [1,
    INPUT_SHAPE])

if next_state[:, action] < min_SNR:
    lf = lf+1

next_state_h = np.reshape(np.append(SNR_h[i+1,:],
    to_categorical(action,
    num_classes=SNR_h.shape[1
    ])), [1, INPUT_SHAPE])

if np.all(current_state[:, SNR.shape[1]:] == next_state[:, SNR.
    shape[1]:]):
    reward = np.multiply(1,np.log10(1+(10**((current_state[:,
    int(action)]/10))))

else:

    reward = (np.multiply(1,np.multiply(np.log10(1+(10**((
    current_state_h[:,int
    (action)]/10))),1)))
    *(1-delay_cost)

score += reward

data_model.remember((current_state), action, reward, (
    next_state), terminal)

current_state = next_state
current_state_h = next_state_h
data_model.step_update(total_step)

if ((j % 1 == 0) and (i % 999 == 0)):
    handover_report(delay_cost,intensity, user)
    a,b,c,d,e,f,g,h = hper(min_SNR, delay_cost, intensity, user
    )

```

```

        print("Run Episode :"+      str(j)      +" Position i = :"+
              str(i)      +" total step
              "+ str(total_step) +
              "   Handover Q      :"+
              str(c)      +"
              handover N      :"+str(
              f)      +"      MSNRQ      :"+
              +      str(d)      +      "
              MSNRN      :"+      +
              str(g)      +"      REward
              "+ str(score))

    if j % 100== 0:
        average_data(j, intensity,delay_cost)

    if terminal:
        break

if __name__ == "__main__":
    bs_intensity = [10,20,..., 100]
    min_snr= [20]
    delay_cost = [3] # t_{d} = 0.75, 1, 2
    for bs_intensity in bs_intensity:
        for delay_cost in delay_cost:
            for min_snr in min_snr:
                for user in range(5,6):
                    solution_v1(user, delay_cost, min_snr, bs_intensity)

```

## RESEARCH OUTPUTS

### LIST OF PUBLICATIONS AND POSTERS

#### International Journals

1. M. S. Mollel *et al.*, "A Survey of Machine Learning Applications to Handover Management in 5G and Beyond," in *IEEE Access*, vol. 9, pp. 45770-45802, 2021, doi: 10.1109/ACCESS.2021.3067503.
2. K. G. Omeke *et al.*, "DEKCS: A Dynamic Clustering Protocol to Prolong Underwater Sensor Networks," in *IEEE Sensors Journal*, vol. 21, no. 7, pp. 9457-9464, 1 April 2021, doi: 10.1109/JSEN.2021.3054943.
3. M. S. Mollel *et al.*, Intelligent handover decision scheme using double deep reinforcement learning, *Physical Communication*, Volume 42, 2020, 101133, ISSN 1874-4907, <https://doi.org/10.1016/j.phycom.2020.101133>.
4. Michael S. Mollel, Shubi Kaijage and Michael Kisangiri, "Deep Reinforcement Learning based Handover Management for Millimeter Wave Communication" *International Journal of Advanced Computer Science and Applications(IJACSA)*, 12(2), 2021. <http://dx.doi.org/10.14569/IJACSA.2021.0120298>

#### International Conferences

1. K. G. Omeke, M. S. Mollel, L. Zhang, Q. H. Abbasi and M. A. Imran, "Energy Optimisation through Path Selection for Underwater Wireless Sensor Networks," *2020 International Conference on UK-China Emerging Technologies (UCET)*, 2020, pp. 1-4, doi: 10.1109/UCET51115.2020.9205429.
2. A. Turkmen *et al.*, "Coverage Analysis for Indoor-Outdoor Coexistence for Millimetre-Wave Communication," *2019 UK/ China Emerging Technologies (UCET)*, 2019, pp. 1-4, doi: 10.1109/UCET.2019.8881890.

#### Workshop and National Communication

1. M. S. Mollel, S. Kaijage, M. Kisangiri, M. A. Imran and Q. H. Abbasi, "Multi-User Position Based on Trajectories-Aware Handover Strategy for Base Station Se-

lection with Multi-Agent Learning,” *2020 IEEE International Conference on Communications Workshops (ICC Workshops)*, 2020, pp. 1-6, doi: 10.1109/ICCWorkshops49005.2020.9145184.

2. M. Mollel *et al.*, ”Handover Management in Dense Networks with Coverage Prediction from Sparse Networks,” *2019 IEEE Wireless Communications and Networking Conference Workshop (WCNCW)*, 2019, pp. 1-6, doi: 10.1109/WCNCW.2019.8902854.