Design and Performance Evaluation of Seamless Handover for Fifth Generation Mobile Communication Systems

SUBMITTED TO AUCKLAND UNIVERSITY OF TECHNOLOGY

IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Supervisors

Professor Peter H. J. Chong

Doctor Hakilo Sabit

2022

By

Masoto Chiputa

School of Engineering, Computer and Mathematical Sciences

Abstract

The fifth (5G) generation mobile network adopted the LTE-mmWave architecture to provide a Heterogeneous mobile network environment. The LTE-mmWave architecture is flexible, and often reliable. It provides sufficient bandwidth and works over a diverse topography. It is also cost effective for mobile companies transitioning from 4G to 5G. This architecture, however, faces a lot of operational and optimization challenges. Particularly, despite its abundant bandwidth hence high data rate, it sometimes provides patchy coverage owing to mmWave propagation characteristics. mmWave propagation characteristics exhibit high propagation losses and blockage sensitivity. Further, users located just a couple of hundred meters away from mmWave gNBs (BSs) and in NLOS scenarios can be noise limited. In fact, for certain communication scenarios, using more bandwidths of mmWave bands has proven to be counterproductive. It is found out that channel estimation penalties end up exceeding the gains of using more bandwidth. This behavior of mmWaves is in contrast with that of legacy cellular systems using only LTE links. Usually, using more microwave bandwidth, for instance in LTE networks favors longer transmission range and high data rates. Given such conflicting LTE and mmWave performance factors, and as mobile networks transition from 4G to 5G, smart Handoffs (HO) are needed in the LTE and mmWave HetNet to ensure seamless, uniform and continuous communication for 5G's mobile users are attained. Unfortunately, most classic HO schemes are unable to take into consideration the vulnerability or the outage unpredictability of mmWaves. Moreover, in spite the 4G to 5G technology transition, a

lot of HO schemes were optimized for old generation mobile networks using microwaves and other radio bands. Furthermore, while recent studies show that smart HO algorithms using AI are being introduced in 5G communication and beyond, they hardly consider the coexistence and competition of multiple performance deterioration factors and parameters in microwave and mmWave links. This leads to HO failures including too-early or too-late handoff, a handoff to the wrong cell or Ping-Pong handoff decisions. The after-effects of these HO decisions include low data rates, high latency, underutilization of mmWave Bandwidth and energy inefficiency of 5G among other. This research thus explored the suitability of using intelligent/ self-learning schemes to sustain 5G mobile network connectivity and performance for mobile users. The study thus reflects on the following:

(i) The prospects of using distance-dependent and mobility Pattern HO schemes to sustain 5G mmWave mobile network connectivity longer.

(ii) The prospects of using Deep Reinforcement Learning with Jump Markov modeling to predicting not just the immediate behavior but also the abrupt and gradual performance changes of the mmWave links before HO execution.

(iii) Using mean field Game Theory and Jump Markov Modeling to predict connectivity sustainability of a target link after HO execution. The scheme extends Markov principles into Game Theory to prediction of random and abrupt mmWave link changes.

(iv) The prospects of route selection schemes in autonomous cars using not just the shortest distance when selecting the best route to destination but also considering 5G mmWave network connectivity guarantees. This ensures mmWave network availability on chosen routes is continuous and improve safety/infotainment data transmission.

Various network simulators are examined to build realistic scenario but for this research, the NS-3 simulator was picked due to its versatility. The NS-3 simulator can integrate with external AI and route mobility applications, e.g., Open AI Gym and Google Maps. This made the simulations closer to more realistic scenarios encountered in mobile network operation. The performance parameters studied are throughput, latency, HO failure rate, energy and spectral efficiency, etc. From the results, we show that the rate of convergence of our learning schemes and making of HO decisions is faster as a result of

utilizing, for instance, jump Markov modelling. This is deduced from the increase in dwell time, reduction of HO failure rate, improvement of energy and spectral efficiency and data rate. In conclusion, the reliability of mmWave connectivity hence 5G mobile access can be improved using smart HO schemes. Furthermore, smart algorithms that can seamlessly be integrated together to be smarter in terms of data rate, HO and energy efficiency if optimized.

Index Terms: Smart handover, Cognitive Radio, fifth Generation (5G) Machine Learning, millimeter Waves, Statistical data modelling.

Confirmation of Authorship

I Masoto Chiputa hereby affirm the originality of this work. To the best of my knowledge, none of the material used has ever been written to a substantial extent of publishing or conferring any degree or diploma by any university or other institution of higher learning.

_____

Signature of Student

Acknowledgements

To begin with, I'd like to profoundly thank my primary supervisor, Professor Peter Chong. From the supreme of his guidance and support during the course research, I have gained so much knowledge and wisdom. His patience, insights in our discussions, and inspiration made my research enjoyable. Professor Peter Chong inspired me to gain new skills beyond research. His openness, and enthusiasm to guide me not just academically but in many spheres of life are admirable. Having a great supervisor with so much patience and understanding has been unequivocally great and rewarding.

I would also like to greatly thank Doctor Hakilo Sabit. His insightful comments, support, and cooperation during my study tremendously helped my progress. In addition, may I extend my appreciation to the Wireless Innovation in Engineering (WISE) group's discussions and events. Furthermore, I'd also like to thank the AUT PhD Engineering and Research group, friends and colleagues.

Lastly but not the least, may I extend my sincere thanks to my family, especially my mother Loveness Sikapulwe, my dear elder sisters Cindy and Nancy and their families, my young sisters and brothers too many to mention, my son Gift and my extended family their unending love, concern, and support both emotionally and financially is unequivocal. I couldn't study without them.

Contents

List of Figures

List of Table

List of Acronyms

AI	Artificial Intelligence
BS	Base Station
DQN	Deep Q-Network
DDPG	Deep Deterministic Policy Gradient
DRL	Deep Reinforcement Learning
DC	Dual Connectivity
DNN	Deep Neural Network
EPC	Evolved Packet Core
EM	Expected Maximisation
EE	Energy Efficiency
4G	Fourth Generation
5G	Fifth Generation
gNB	Next Generation Node Base
GHz	Giga Hertz
GUI	Graphical User Interface
IoT	Internet of Things
KMO	Kaiser-Meyer-Olkin
LTE	Long Term Evolution
JMLS	Jump Markov Linear System
MHz	Mega Hertz
mmWave	Millimetre Wave
NR	New Radio
NS-2	Network Simulator-2
NS-3	Network Simulator-3
OMNeT++	Objective Modular Network Testbed in C++
NetSim	Network Simulator
SINR	Signal to Interference plus Noise Ratio
PDCP	Packet Data Convergence Protocol layer

NSA	Non-Stand Alone
mm-BS	Millimeter Wave BS.

List of Variables

Notation	Description
n	The number of Handover
d_{max}	The maximum distance
T_{gen}	The elapsed time between the reception of the first diagnosis message and the generation of a test request
T_f	The time needed to propagate a dissemination message
T_{out}	The time needed to collect test replies from neighbour nodes
d_{st}	The depth of the spanning tree
$\Delta\theta_0$	The change in Angle of Arrival
k	The number of iterations
ζ	The number of diagnosis periods triggered due to the movements of the active mobile hosts
l	The link
η	The direction
s	The state

$C_{i,s}$	A list of ordered nodes tested by node i in a cluster of size 2^{s-1}, in a given testing round.
V	velocity
δ_G	The minimum vertex degree of G
Δ_G	The maximum vertex degree of G
σ_v	The maximum number of faulty nodes among v's neighbors
α_v	The number of replies to wait for to make a decision at v
R_v^u	The result of a tested node v for a tester node u
FF_v	The set of fault-free nodes at a node v
F_v	Set of faulty nodes at v
m	The number of incorrect outputs (Extremely large)
k	The number of possible incorrect outputs
q	The HO
ID	Node identifier
L	The set of links in a network
t	A specific time

Chapter 1

Introduction

The mobile technology evolved using microwave bands over the last century. However, the microwave bandwidth has been depleting while the demand for mobile data continues to explode. There is need for more bandwidth to meet: gigabit data rates, the microsecond latency demands, improve reliability, security, and scalability in future mobile networks [1]. To that effect, research studies to provide Gigabit [1] and Terabit-per-second (Tbps) rates are underway [2]. Driving the gigabit data rate research are studies on millimetre (mmWave) bands. Unlike the Fourth Generation (4G) Long-Term Evolution (LTE) using microwave bands, the mmWave bands unleash mega bandwidth in the Fifth Generation (5G) mobile. This is tens of GHz bandwidth than what is obtainable from legacy systems (only in hundreds of MHz) such as LTE [1]. The mmWave frequency bands however incur higher propagation losses, offer short and patchy coverage. Additionally, mmWave propagation can easily be blocked. In addition, they are also poorly reflected/refracted on walls and encounter non-line sight (NLoS) scenarios [3],[4] too often. To mitigate the mmWave challenges, 5G technology integrates mmWave and LTE technologies. The

microwave bands in form of LTE radio access network (RAN)) offer long range coverage while mmWave bands simultaneously offer gigabit data rate [5]. The LTE-mmWave framework provides flexible, reliable, and secure communication for 5G and beyond. It however requires smart Handoff (HO) schemes. Smart HO schemes ensure a seamless and continuous communication platform among 5G mobile users [2], [4],[5] as they interchangeably use LTE and mmWave technologies.

Technically the HO process in mobile networks is termed as a process of switching users with an ongoing data transmission from the serving channel to another free channel link normally known as the 'target channel link.' The challenge with existing HO schemes is that most of their traditional/classic algorithms were optimized for legacy communication systems using microwave and UHF bands [6]. As a result, when applied in 5G, they manifest into too-early, too-late or wrong HO. And sometimes they cause ping pong HO [7]- [12] scenarios. At worst, most tradition HO schemes result into link failure [13]. The 5G mobile network thus require a new framework of HO mechanisms that takes into consideration the vulnerability of mmWaves. This will ensure mobile users benefit from their large bandwidth (commonly referred to as gigabit data rates). This thesis thus focuses on designing smart HO scheme for 5G's LTE-mmWave mobile framework. The HO designs take into account mmWaves' high sensitivity to blockage, short coverage range, propagation loss and the high contrast in data rate between LTE and mmWave links. Additionally, the designs also consider user mobility effects and the limited bandwidth hence data rate of LTE links. Additionally, we take advantage of LTE's long-range transmission capabilities (30km) compared to that of 5G's mmWaves in hundreds of meters only [11].

1.1 Background

They are many specific factors that would need triggering HO execution process in 5G mobile networks but not in 4G. These make most concepts used in designing 4G HO schemes infeasible in 5G mmWave mobile networks. Some of the major factors include:

- **Range**: One of the most significant areas of concern is the mmWave spectrum. Signals at mmWave frequencies are particularly vulnerable to being obscured by shadows. Brick, for instance, is one type of material that can reduce the strength of a signal by as much as 40–80 decibels [8, 9], [6], – [10]. On the other hand, the human body can reduce the effectiveness of mmWave communication by 20–35 dB [19]. In addition, high levels of humidity and rain fade can

cause significant transmission issues for long-range mmWave. Furthermore, for users over ranges of one hundred meters away from BSs or in NLOS are noise limited. The use of more bandwidth in NLOS scenarios for instance, results in a diminishing return. It can even be counterproductive to an extent where the channel estimation penalty becomes greater than the gain from increasing bandwidth. This is in contrast to the interference-limited legacy cellular system such as the 4G LTE where it is advantageous to use more bandwidth for long range. Faced with such a callous situation, 5G HO systems are needed, for instance, to ascertain at what distance from serving BS is worthy sustaining the serving link connectivity of mmWave BS with huge bandwidth. In addition, they need to know whether to trigger a mmWave-to-mmWave HO (horizontal HO) or mmWave to LTE (vertical HO) or vice versa. Furthermore, HO processes need to ascertain connectivity sustainability levels of a target link post HO before triggering a HO. To that effect, despite the negativity of human body brokerage, just like many outdoor materials that are very reflective, 5G mobile networks strategically take advantage of signal scatterers in sustaining mmWave propagation [18], [20] over long transmission ranges.

- **Channel Fluctuation and Intermittent Connectivity**: In the case of velocity driven influence, it is well established that the channel coherence time is linear with carrier frequency [9]. This is a point of contention among different parties because mmWaves are needed to support mobile users. The mmWave channel changes in the order of hundreds of microseconds with user speed. This is significantly faster than those in users in 4G LTE. For example, a Doppler spread at 60 GHz for a user moving 60 km/h is over 3 kHz. This is in addition to the impact of shadowing: the appearance of obstacles in mmWave propagation turn to cause more dramatic swings in path loss [26]. This is besides the shadowing effect. Furthermore, mmWave communication is inherently short ranged. Technically this also implies that the relative path loss estimates, and cell association triggering time need to be fast enough to comfortably switch from one BS to another and sustain on going connectivity. At the system level, the HO schemes need to rapidly adapt to circumvent performance degradation due to high intermittence and sustain connectivity. A HO algorithm in fact if well designed can ensure swift switches between LTE and mmWave links and among mmWave links. This ensures both bands are fully utilized and the link between users and BSs is maintained to meet the desired quality of service or customer experience

envisioned for 5G. Smart HOs additionally facilitate continuous transmission for mobile users by switching to LTE links or mmWave links depending on the user behavior's impact on 5G links.

- **Directional Transmission**: The incorporation of mmWave technologies into cellular access networks is not a seamless process because the properties of mmWave technologies at physical layer has an effect on a great number of essential HO network functions [25]. The Cell Discovery process for instance is an essential part of HO process. Th accuracy of cell discovery makes the HO decision process more reliable. Contrary to previous cell discovery requirement, the 5G process requires synchronization of signals in a directional way than the typical broadcasting into cells. The 5G mmWave propagation has adopted directional transmission to increase range and reduce interference. The directional nature of mmWave communications, however, requires spatially aligned transmitters and receivers in exchange synchronization signals [17]. For HO execution, this involves a stage where mmWave BS and the user need to scan through different directions before finding one that allows it to communicate. This step, if not intelligently managed creates huge latency in the HO selection process and ultimately impair delay-sensitive applications [19]-[21]. Smart cell discovery processes are needed to maintain a reasonable HO time. Another effect of directional transmissions is that the links become directionally unique. This makes many HO technologies introduced in 4G including those using techniques like coordinated multipoint, intercellular interference coordination, and interference alignment, have limited benefits in 5G systems [22].

HO algorithms must ensure that the switch between the two link takes into consideration the above concerns, data rate disparities, connectivity reliability, spectrum and energy efficiency and most importantly that the switch itself is seamless and supports continuous communication. Unfortunately, most existing HO schemes are unable to take into consideration the vulnerability nor random behaviour mmWaves into consideration. Ultimately classic HO schemes can't sustain 5G mobile network connectivity. Users in 5G end up relying more on LTE links than mmWaves to sustain connectivity. Moreover, the low rates result into degraded 5G mobile performance and failure to meet key performance indicators. This motivated as to design smart HO schemes as detailed in the following sections.

1.2 Research Motivation and Research Problems

The 5G mmWave mobile network's prone to abrupt changes and link failure opens a lot of research challenges and opportunities. Particularly, the abrupt and random performance of mmWave links and failure can hurt 5G mobile network's goal to offer gigabit data rates. Handover intelligence is essential to preventing 5G mmWave link failure and providing uniform performance [7]. To that effect, research on Smarter HO is inevitable to enhancing data rates, energy efficiency and mobile scalability among many others. To that effect, the author [8] for instance developed a 5G HO model based on user-BS distance and LOS probability functions. This model improves spectral efficiency when compared to traditional models. However, link brokage owing to mobile user activities and the patchy coverage makes the solution inefficient [24]. Another author in [26] proposed a cell discovery algorithm for HO schemes. It reduces antenna search space by excluding angular sectors depending on the user location data. This solution however prevents NLOS users from communicating via reflected paths. It is not energy efficient too. Authors [17] - [19] propose context-aware cell discovery algorithms. These can handle obstacles and reflected paths. While authors in [16] and [22] proposed an obstacle-free HO strategy using the user-side context data and direction. Using a high SNR regime, [13] proposes a HO scheme that transforms the HO process into a resource allocation problem and uses a concave optimization method to optimise mmWave link selection. The challenge in these solutions however seem to be similar. They are only precise at high SNRs. In reality however, mmWave systems suffer from pathloss and shadowing. The high SNR value expectations are unrealistic and evoke unwanted HO. Using machine, biological, and Quantum Physics methods, [26, 27, 28] developed 5G HO schemes. Machine based algorithms can adapt to network conditions, so they're useful in 5G mmWave systems. They can also readjust time to trigger given the changing conditions such as link quality and channel gain. Thus, need little to no human intervention. 5G performance however is harmed by their long learning processes and training overheard. Furthermore, most of these solutions were optimized for LTE systems using microwave propagation. Thus, there is appetite to formulate feasible HO schemes that can efficiently handle mmWave challenges and ultimately help 5G mobile networks attain their Key Performance indicators.

While 5G has many problems to prompt research as highlighted in the paragraphs. Some of the specific Handoff problems in 5G mobile networks include:

- 5G LTE-mmWave mobile networks turn to prefer more handoffs to LTE links compared to mmWave links. The domination of LTE links in 5G LTE-mmWave Dual Connectivity owing to their long-range transmission capability in spite their low data rates result into mediocre average data rates for 5G mobile communication. Considering the high propagation loss and blockage sensitivity [15] in mmWave, smart HO algorithms, that are mmWave-propagation limitations aware are needed to find more reliable mmWave links. This can minimize HOs and HO failures whilst ensuring high and uniform data rates are attained.
- For mmWave systems, users located at least hundreds of meters range and in NLOS are usually noise limited. Unlike LTE links using microwave bands, were using more bandwidths or power improves network performance [3]-[5], mmWave propagation encounter high transmission cost in terms of spectrum and energy. The problem is exacerbated by the fact that the Distance aware [30] HOs or Context aware HOs [31] designed for 5G do not take stock of such factors limitations. Further, most HOs do not take advantage of the reflective nature of some mmWave signal [32]. This leads to not just too many link failures but wasteful HOs and inefficient transmission. The ultimate consequences of poor HOs, too early or late HO tend to outweigh the benefits of using mmWaves in mobile networks.
- Most HO schemes using Machine learning or statistical modelling schemes require huge number of training data set to come up with reliable HO policies. Some in fact turn to need offline training to formulate credible HO decisions. Further even those that have integrated HO methods, rarely take stock of the competing deterioration factors of mmWaves transmission. For instance, in [6], the author integrated Mobility load balancing (MLB) and Mobility Robustness Optimization (MRO) HO schemes using Reinforcement Learning to solve 5G mobile HO challenges. The scheme does not consider mmWave deficiencies such as blockage. It only simultaneously focuses on load balancing and user speed triggered HOs. Further, the scheme relies on offline training. Offline data, however, is not always available. Further, mmWaves have random behavior that is rarely repetitive and would require huge training data sets in offline training.
- For the most HO schemes, the latency disparities between an omni directional transmission cell search in LTE [41] and directional cell search process in mmWave links is

ignored. This can cause link failure and unreliability in the way mmWave link are selected due to its longer cell searching period Having reliable knowledge and analysis on mmWave network availability can help issue swift but reliable HO decision. For instance, knowing the route to destination the user is taking can help the HO scheme prepare HO cells in advance and reduce the search period. This can be helpful especially for autonomous cars that could largely rely on mmWaves for continuous connectivity and huge sharing/gathering of data of their environment.

1.3 Research Objectives and Contribution

This thesis provides several original research HO solution to the many 5G LTE-mmWave HO challenges enlisted above. Our main contributions are given based on the research Problems in the preceeding sections and the set objectives as reflected below. We consider:

(i) **The prospects of using statistical models to design 5G HO schemes that do not just consider the distance in LOS but also intelligently consider reflected signals and the likely network pattern guarantee longer continuous user-mmWave connectivity after HO**.

The contribution to objective (i) is: we introduced a network availability pattern aware Handoff Scheme to improve target mmWave link behaviour predictions in the HO execution. Considering the mmWave performance can drastically and abruptly change within a split of a second, i.e. from excellent (best data rate) to worst e.g., outage state. The proposed HO scheme does not just predict the initial behaviour but also the likely network performance deterioration pattern or behaviour of a mmWave target link after a HO. The scheme employed Jump Markov System modelling to predict abrupt and drastic changes of mmWave target links after a HO.

(ii) **The prospects of predicting mmWave network patterns in target links by using online Artificial Intelligence (AI) learning schemes to estimate (before a HO selection process) the reliability of the mmWave link after a HO execution**.

The contribution to objective (ii) is: we introduced an integrated AI and statistical modelling HO model with online learning capabilities. The scheme capitalizes on statistical data modelling's novice policies to learn and formulate stable policies faster DRL. Here DRL leverages the statistical data variance of many mmWave link behaviour to predict the extent to which a HO parameter value can change. The proposed HO model does not need as much training data as DRL

to make good policies or converge. It combines approximation and accuracy to sustain connectivity. The proposed HO scheme was focused on solving problem number two.

(iii) The prospects of using machine learning schemes with meta data to the reduce training overheard and time for HO decisions.

The contribution to objective (i) is: we introduced the Mean-Field Game Theory and Jump Markov System modelling to predict the mmWave links' likely behaviour pattern after HO. The study introduces a user-behaviour tailored decentralized HO framework for 5G networks. The proposed HO scheme was focused on solving problem number (iii)

(iv) **The prospects of using not just aspects of shortest distance to a given destination but also mmWave network aware schemes in route selection by users such as autonomous cars. Knowing the network availability on the route prior to selecting it, improves prospected of continuous routing of important safety messages and improves user's quality of experience**.

With respect to objective number (iv): we introduced a Mean field Game Theory scaled DRL HO schemes to sustain user connectivity aware route selection mechanism. The scheme helps tries to select routes to destination based on mmWave network availability instead of just the shortest distance. For autonomous cars that need continuous and sustainable communication up to destination for safety messaging the schemes, this is a good start point. In addition, the schemes employ fast cell selection to reduce HO delay. The proposed HO scheme averts unwanted HO trials when a short residence time is expected on a particular target link.

1.3.1 Originality of the Thesis

It is worth reiterating that none of the mentioned works have been used in a similar fashion to solve HO problems in 5G HO algorithms.

1.4 Organization of Thesis

The thesis has 8 chapters. Figure 1.3 illustrates the structure flow of the thesis.

Chapter 2: This chapter reviews the fifth Generation (5G) mobile networks, and the importance of smart Handover (HO) in mobile networks. It avails the challenges of classic and smart HO schemes when applied in 5G mobile networks. The chapter also outlines qualitative and

quantitative statistical and machine learning mechanisms that are applied in other fields but could help improve HO processes in 5G mobile networks.

Chapter 3: This chapter details the objectives in our methodology and designs. It goes on to justify the rationale behind the research methodology and tools. Further, it outlines the validity of the verification processes adhered to in this research.

Chapter 4: Here, we scrutinize the problem of mmWave link reliability in 5G HO using statical data modelling HO schemes. The proposed HO scheme learns to predict the chances of the target link being in LOS and NLOS after the HO is executed using the Jump Markov Linear System. It leverages JMLS' ability to model continuous dynamics between modes and the abrupt changes to predict abrupt changes in mmWave target links.

Chapter 5: This chapter addresses the problem of link reliability in HO selection using Deep Reinforcement Learning in mobile wireless networks. We introduced a probabilistic model for mobile link behaviour estimations by the HO in NLOS and LOS scenario using the JMLS. The schemes integrate the two learning schemes to reduce training time and overhead in DRL.

Chapter 6: In this chapter, a Game-Theory -JMLS HO scheme was designed, and results compared to see if it is better than the DRL-JMLS HO model. The HO schemes proved to be energy efficient, had a better Data rate. The Handoff failure rate and system Over head proved to be minimal in spite being more environment aware than earlier proposed HO schemes. In addition, it improved DRL in all aspect of convergence, sample size analysis and parameter complexity understanding and interpretation.

Chapter 7: This chapter presents a network connectivity aware route selection model for autonomous cars. This is unlike traditional models where the route to destination is dependent on shortest route. The scheme helps cars to select routes to destination that will ensure the mmWave network availability on autonomous cars is continuous and sustainable up to destination. We integrate the mean field Game Theory and DRL concepts to learn available route to destination and the likely mmWave network behaviour. This is vital and improves user experience of service when using autonomous cars

Chapter 8: This is the last chapter. It concludes the thesis by providing a summary of the contributions of this research. It further provides suggestions and directions for future research.

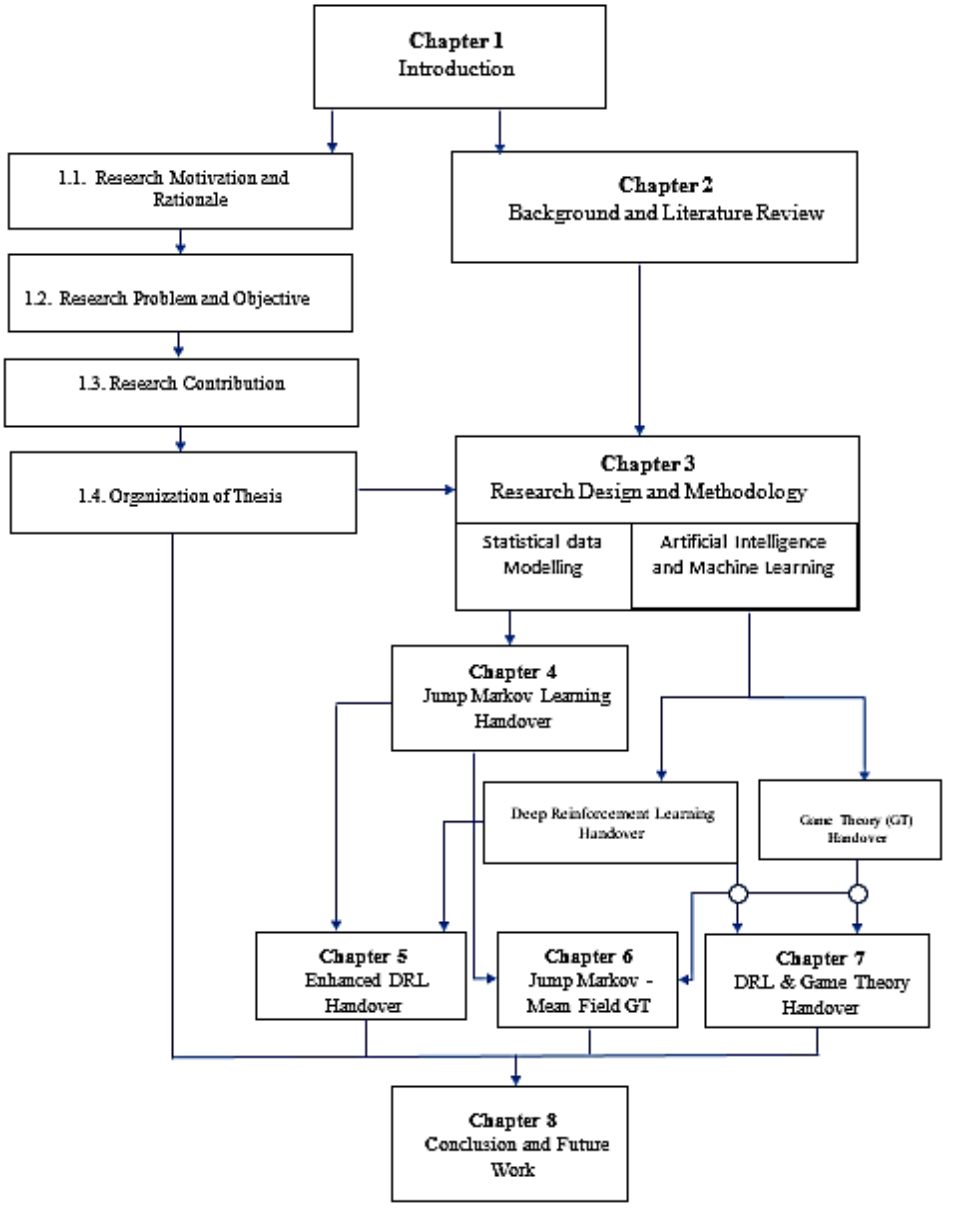

Figure 1.1: The structure of the thesis.

1.5 Research Publication List

1. **M. Chiputa**, P.H.J Chong et al**.,** "Cognitive Radio Technologies In 5g mm Wave Cellular Networks: Challenges And Solutions", *Advances in Engineering Research*, Victoria M. Petrova, New York, NY, USA, Nova Science Publishers, 2022 ,pp. 1-23.
2. **M. Chiputa,** M. Zhang, G. Ali, P.H.J. Chong, H. Sabit, A. Kumar, H. Li, "*Advanced Antenna Techniques for IoT and 5G Applications Sensors, MDPI Sensors Journal*, vol. 22, no.3, pp. 145-167, Jan 2022.
3. **M. Chiputa**, P.H.J. Chong and A. Kumar, "An Autonomous Deterioration Prediction Based Handover Model for 5G Networks," *2021 IEEE International Conference on Computing (ICOCO)*, 2021, pp. 168-173. doi: 10.1109/ICOCO53166.2021.9673555.
4. **M. Chipta**, M. Uttarwar and P. H. Joo Chong, "Intelligent Handover using User-Mobility Pattern Analysis for 5G Mobile Networks," *proceedings 2021 Conference on Information Communications Technology and Society (ICTAS)*, 2021, pp. 5-10, doi: 10.1109/ICTAS50802.2021.9395027.
5. **M. Chiputa**, P.H.J Chong, S.U. Rehman, A.Kumar ,"Investigating Mobility Robustness in 5G Networks Using User-Adaptive Handoff Strategies," in *Springer Ad Hoc Networks. ADHOCNETS 2019. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, vol 306.,2019. Cham. https://doi.org/10.1007/978-3-030-37262-0-8

1.6 Summary

This chapter introduces the importance of 5G mobile networks and smart HOs. It then highlights HO challenges in 5G mobile networks, the rational and motivation of the research. It goes on to state the HO problems in 5G with mmWave networks, the objectives of the research and highlights the contributions before summarizing the structure of the whole thesis writeup.

Chapter 2

Literature Review

The rationale for exploring the general Handover (HO) problem and known solutions in mobile networks were outlined in chapter 1. As alluded in chapter 1, the major goal of this thesis is to improve millimetre wave connectivity and dependability in fifth generation (5G) mobile networks via offering Handover schemes that meet 5G mobile network standards. To accomplish the goal, a thorough comprehension of the HO problem and analysis in 5G mmWave mobile networks is required. To that effect, Chapter 1 gave a background to the HO problems discussed in this thesis. It particularly introduces the reader to mobile communication HO and provides the necessary context for understanding the research challenge, opportunities and limitations. Chapter 2 presents a set of fundamental HO concepts, definitions and works relevant to 5G HO. It describes the 5G mobile network objectives, the role of HO and primary mmWave impairments, including weaknesses and consequences of those weakness and failures on mobile network. It then describes several machine learning and statistical techniques commonly used to avert mmWave challenges and enhance the reliability of the 5G mobile network. In addition, it reviews literature on Cognitive Radio (CR) techniques, discussing their strengths, weaknesses and how they can be applied to

improve 5G HO applications. The literature presented in chapter 2 literally identifies research gaps in the HO research area for 5G mobile networks and pioneers the way for developing competitive HO models for 5G mobile wireless networks. The chapter information is organised as follows. Section 2.1 introduces 5G networks, standards, problems, and ambitions. Section 2.2 discusses the function of 5G HO approaches, detailing their benefits and drawbacks. Sections 2.3, 2.4, and 2.5 examine user-aware, statistical, and machine learning HO challenges, as well as comparative methodologies and constraints in 5G wireless networks. Section 2.5 concludes the chapter by summarising the results and highlighting some of the pertinent research questions that need to be addressed further.

2.1 Overview of 5G networks

The fifth generation (5G) communication paradigm is profoundly altering the wireless communications and the telecom system as a whole. Smart cities, agriculture and energy sector are among some of the verticals that will tremendously benefit from 5G. To meet the expectations, 5G is slated to use multiple spectrum bands for mobile systems. This includes bands below and those above 1GHz up to 6 GHz. And most importantly launches the use of the millimetre wave (mmWave) frequency range. The current dwindling availability of bandwidth prompted the adoption of the mmWave spectrum to meet 5G expectations. One of the most anticipated outcomes of 5G New Radio (NR) systems in mmWave frequencies is their ready applicability for increased mobile broadband services [123]. Through dense spatial reuse via beamforming, smart beam switching, HO and relay techniques, the mmWave spectrum's is envisaged to provide enhanced user experiences with ultra-high data rates and low latency at both the cell site and device [126]. The utilisation of unused mmWave channels (technically 30 to 300 GHz, but also down to 24 GHz) will provide a substantial amount of bandwidth; approximated to be around 100 GHz earmarked for 5G mobile communication network alone [89]. This new platform using the mmWave spectrum, in conjunction with low- and mid-band spectrum provides breakthrough technologies that will exploit mobile technology to its fullest extent.

However, deploying mobile communications using the mmWave spectrum necessitates dealing with the harsh propagation conditions associated with high frequency bands. These include increases in pathloss, decreases in dispersion, less channel variety, and greater obstruction leading non-line-of-sight pathways. These important difficulties should be addressed for effective 5G

mobile communication. Furthermore, due to the use of greater bandwidths, the effect of noise power is a greater concern than for microwave systems. Without a doubt, various challenges must be addressed in order to efficiently utilise this spectrum.

Comprehension and eventual adoption of 5G mmWave technologies can successfully open up fascinating new frontiers in mobile broadband communications. The following sections provides information on primary elements that are crucial for 5G mmWave mobile communications and HO designs and application. This chapter outlines our comprehensive knowledge on mmWave propagation elements and challenges. It goes on to discuss the key difficulties in existing HO solutions, the limits connected with the utilisation of new HO schemes in 5G with the mmWave spectrum. Most prominently, this chapter details the lessons being learnt in the ongoing mmWave HO dispositions and suggests new HO solutions for future 5G mmWave HO technology.

2.1.1 5G mmWave Spectrum adoption

The mmWave spectrum range is between 30 ~ 300 GHz. [51,52]. The 5G mobile technology, however, includes spectrums in the range of 24 GHz and 28 GHz. This is owing to the fact that they all have similar working properties. Given the above mmWave spectrum range, research and development of mmWave 5G system has evolved well. The efforts that national and regional institutions have made in actualising the deployment of 24 GHz and above bands for 5G operationalisation are visible. Though not thorough yet, the global availability of 5G mmWave mobile technology is taking shape as evidenced in Table 2.1 below.

Table 2.1: Overview of the millimetre Wave Spectrum usage in Different Countries [1].

Territory	Spectrum	Licensed usage
Greece	24.5 – 26.5 GHz	Fixed Radio Access Network
China	26.55 – 27.75 GHz	Mobile and Fixed Radio Access Network
Italy	26 GHz	Open to all 5G use
Japan	27.0 – 28.2 GHz	Open to all 5G use
Korea	29.1 -29.5 GHz	Open to all 5G use
United Kingdom	24.5 – 26.6 GHz	localized spectrum licensing (technology neutral, indoors only)
Uruguay	27.5 – 28.35 GHz	Mobile Services
USA	24GHz, 28GHz	Technology Neutral

The millimetre wave spectrum is used in three major areas. The first one is the Licensed spectrum for mobile and fixed access. Then there is the point-to-point backhaul bands that include E band (70/80 GHz) and lightly licenced ones such as the 11, 13, 18, and 23 GHz. There is also unlicensed range between 57-71 GHz and the 14 GHz contiguous spectrum assigned for point-to-point and point-to-multipoint communications. It is fifteen times more than unlicensed Wi-Fi spectrum. Then for dynamic shared access between commercial and federal users the 37-37.6 GHz bands provide about 600 MHz of bandwidth. According to the FCC part 40 [2.2] the U.S. Transmit power limits for these bands according to the EIRP. EIRP is technically the sum total of transmit power plus antenna gains with an array of n row and m columns antenna gain (G_{Ant}) defined as:

$$G_{Ant} = G_{Element}(m) + 10\log(n * m),\ G_{Element}(n) \text{ (as stated by vendors).} \quad (2.1)$$

FCC defines the following EIRP:

- 75 dBm/100 MHz for fixed BS
- 43 dBm for mobile user
- 55 dBm for mobile BS

Further for the UEs, they are four power classes for mmWave bands set by 3GPP [2.4]: Power class 1 (PC1) is for fixed RAN, while PC3 is for handheld mobile phones. PC2 is for vehicular and PC4 for non-handheld.

To reflect distinct use situations, the enlisted PCs above use different percentile CDF definitions. The lowest EIRP represents the maximum gain angle in the spherical field. For phones, the 50th percentile EIRP is more representative of general use, indicating that the EIRP will be less than 11.5 dBm in 50% of the angular direction.

2.1.2 Millimeter Wave Challenges

While theoretical analysis indicates that the potential market for mmWave is high, particularly in terms of extreme data rates, several vital challenges for using mmWave in mobile networks emerge. We give a brief highlight to these topics.

- ***Free Space Pathloss****:* A signal's power density is attenuated or reduced as it travels, which is known as pathloss. It is a crucial factor in the analysis and link budget design for all wireless communication systems. MmWaves have a smaller range due to increased

pathloss than lower frequencies. The radiated power density in the absence of air absorption/scatter decreases by $1/d^2$, where d is the transmission distance and $1/d^2$ degradation is frequency independent. Using the Friis equation, i.e., providing the fundamental correlation amongst transmit and receive power under LOS scenarios. The free-space pathloss (LFS) i.e., loss in the Friis equation owing to distance and wavelength, λ for an isotropic unity-gain antenna, i.e., G_T gain at transmitter and G_R gain at receiver is:

$$LFS = {}^{P_T}/_{P_R} = \left({}^{4d}/_{\lambda}\right)^2. \quad (2.2)$$

In dB, the loss, L in (2.2) for a non-unity gain antenna, ($G_T \neq G_R$) is [56]:

$$L(dB) = -G_T[dB] - G_R[dB] + 20\log_{10}(d[km]) + 20\log_{10}(f[GHz]) + 32.44\,, \quad (2.3)$$

where, d is in kilometres, and f is frequency in GHz. As observed in (2.3), frequency-dependent loss has significant influence. For example, the free space pathloss at 28 GHz is 29 dB higher than at 1GHz for the same distance. Equally, at 38 GHz, it is 31.6 dB higher.

- ***Blockage (Shadowing Effects):*** The fact that the mmWave wavelength is less than a centimeter in length, the majority of the objects in the environmental turn to appear to be proportionally larger. This makes mmWave signals more sensitive to environmental impediments than sub-6 GHz communications. Consequently, when mmWave signals come into contact with these objects, they tend to be completely/partially absorbed, scattered, diffracted, or reflected. In fact, the interaction between the rise in scattering, and decrease in diffraction causes more sever shadowing. To that effect, mmWave communication systems typically employ antenna arrays with high gain and narrow beamforming to mitigate shadowing. The effects of shadowing are represented as a large-scale variation surrounding the pathloss and can be expressed as:

$$L_{Large-scale}(dB) = L_{LOS}(dB) + S_{\sigma}[dB], \quad (2.4)$$

where $L_{LOS}(dB)$ is the pathloss component. S_σ is the shadowing loss and $\boldsymbol{\sigma}$ is its standard deviation. When modelling the mmWave channel and defining time variance in mmWave channel, it is essential to keep in mind that shadowing plays are very vital part [6].

- ***Penetration Loss:*** By comparing the power levels of unobstructed and blocked paths, penetration loss provides mechanisms for quantifying the overall attenuation. Typically, this exceeds the free space pathloss. Studies on mmWave penetration losses at 28 and 38 GHz have revealed a comprehensive picture of its influence on mmWave communication. Results demonstrate qualitatively that mmWaves penetrate polystyrene with minimal reflections. The reflection effect on wood and walls is negligible, whereas that of tinted glass and brick is strong and results in a significant loss of light penetration. Quantitively, For example, materials such as colored glass can have a penetration loss of up to 40 dB, whereas the penetration loss of brick is estimated to be around 28 dB. Clear glass and walls have minimal acoustic transmission losses between 3.6 and 6.6 dB. The indoor loss of tinted glass was approximately 24.5 dB, approximately 3.5 dB less than the outdoor loss. Considering the effects of reflection coefficients and penetration losses are normally higher in mmWave transmission is important too. Similar observations are noted at 60 GHz and 72 GHz when compared to those at 28 GHz. In actuality, the loss of penetrability increases with frequency. Depending on the polarization of mmWave signals, at 30 GHz to 50 GHz, research has revealed that material penetration losses fluctuate. For instance, a penetration loss difference of approximately 1.77 dB/cm was observed for solids in [56]. Owing to lack of ability for mmWave to penetrate building materials, some studies have demonstrated a penetration loss of approximately 74 dB when travelling from indoors to outdoors and vice versa. The difference between indoor and outdoor penetration loss enables frequency reuse in two environments. To ensure indoor-to-outdoor coverage [4], heterogeneous networks, repeaters, and relays are required due to the enormous penetration losses.
- ***Foliage Loss:*** For mmWave links foliage and more broadly, vegetation clutter is another challenge. A study for urban and sub-urban settings avails a deep understanding of the influence of foliage on mmWave when contrasted with sub 6 GHz frequency band signals. It is characterized by Fresnel zones around transmitter and receiver [4]. The studies [10] conducted for 38 GHz outdoor channel revealed that wet foliage can be a cause of multipath

reflection. At 29 GHz, results show that fading owing to foliage can reach to about 10 dB and for 5 GHz it is assumed to approximately reach 2dB. Additionally, at 57.6 GHz – foliage covering a meagre distance of 5m can cause an attenuation of around 40 dB greater than that of 9.6 GHz. The foliage covering 20-80m or equivalent to just 4-14 trees with heavy foliage, causes around 50-80 dB attenuation [114]. Further, for frequencies from 28.8 GHz up to 57.6 GHz, measurements result in [88] revealed a significant foliage loss.

- ***Body and Hand Losses:*** The characteristics of mmWave signal propagation are significantly influenced by humans. Typically, the received power with and without humans present is used to compare body or hand loss. For a single person blocking mmWave signals, the technique revealed body losses of between 25 and 40 dB at 15, 21.5, 60, or 73 GHz with horn antennas [5] [11]. This assertion was supported by research reported in [45], in which the authors stated that measurements made at 28 GHz on UEs with phase array antennas in a mobile human environment revealed a mean body loss of 8.5 dB. According to reports, mobile users experience body loss that lasts at least 100 milliseconds [11]. There is a decrease of 100 milliseconds in a more mobile environment. This aspect significantly affects the robustness of human body or hand blockage mitigation mechanisms, such as beamforming/HO solutions. The time-variation of the blockage is significant. For example, beamforming algorithms like beam switching must be activated within 100 milliseconds in pedestrian scenarios to prevent link failure [11].
- ***Scattering:*** mmWave transmissions exhibit multi-path propagation as a result of scattering. Compared to lower frequencies, mmWave propagation environments have many more sources of scattering due to the pervasive presence of objects larger than the wavelength. When line-of-sight (LOS) conditions apply, reception is greatly aided by objects like lamp posts and the sides of buildings. Rough outdoor materials like brick walls or tree bark also contribute significantly to scattering because their surface area is so large in comparison to the wavelength. Humans, other people, and the objects and places around them all contribute, albeit to varying degrees. The received signal strength is reduced when the signal is scattered off of furniture under LOS conditions, but the opposite is true under NLOS conditions. With NLOS, a larger spread amplifies any refracted component to a more noticeable delayed spread [4].

- ***Atmospheric Loss:*** Atmospheric losses at mmWave frequencies are primarily caused by oxygen (O_2) and water vapor. For the conditions shown in Fig. 4.2, their typical contributions to attenuation are depicted. Loss magnitude is location-specific and influenced by the local concentration of oxygen and water vapor. For instance, at 20 degrees Celsius and sea level, the 60 GHz peak is about 13 dB/km. Since this is a relatively new frequency, it is easy to see why satellite providers and other incumbents have avoided using it. There are also prominent peaks at 180 and 315 GHz, with attenuations of a few decibels per kilometer. The "windows" in the spectrum between these absorption peaks allow for easier long-distance transmission. The resulting bandwidths hover around the frequencies of 35, 94, 140, and 220 gigahertz.

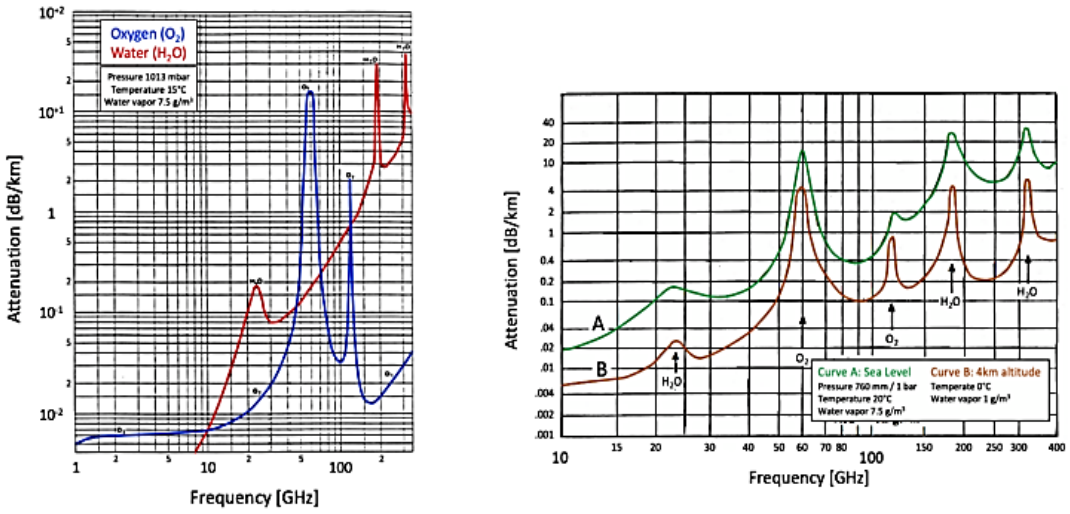

Figure 2.1: (a) Atmospheric and molecular absorption in different frequency bands [35], (b) Rain attenuation in different frequency bands [1].

2.1.2.1 mmWave Propagation, Channel Models and Opportunities:

Distance, obstruction (shadow fading), varying penetration losses, and atmospheric losses are just some of the many known influences on mmWave propagation. In addition, the RF waves are further reflected and diffracted by the presence of intervening and surrounding objects (or clutter). Since mmWaves are so short, it can be difficult to collect enough precise information about the environment and background noise to properly model the situation. Consequently, designing smart cell HO selection schemes and propagation systems for 5G requires a lot of work across the 5G protocol structures. In the same vein this presents a lot of research opportunities ranging from the

physical layer (smart beam form), mac lay (smart relaying mechanism) and at system level (smart HO scheme).

2.1.3 5G Mobile Network Standardization

The standardisation of enabling technologies in 5G mobile networks is currently underway. In 2017, 3GPP Rel-15 launched 5G (New Radio) NR for mmWave. It addressed RAN standardisation. This was required in order to operate mmWaves with LTE in a 5G framework in a cost-effective way [60]. The specification defined the basic operating points of 5G NR with sub-6 GHz bands used for LTE and previous generation technologies. It also included higher frequencies between 24 and 52.6 GHz [61]. Rel-15 further defined the framework for beam-based operations at the PHY layer and at System Level HO. For both cases the mmWave was to function for mobile broadband. The 3GPP committee later released work on the next generation 5G core network, it expanded to begin standardisation of "standalone" 5G. The work on improving use cases goes with Rel-16 completed in 2020[61] and Rel-17 study underway.

The cellular access for 5G networks has not been a seamless process because as it brings tremendous changes to the protocol structure. For instance, in the physical layer functions [3] beam forming is emphasised and has an upward impact on network functions like HOs and relays. The Cell Discovery process is one of the most important network functions that has been impacted and impacts the HO performance. Unlike previous generations, which broadcasted synchronisation signals into the cell,[34] the mmWave has adopted directional communication. This is done to increase range while decreasing interference and co-channelling issues [47]. The directional nature of mmWave communications, on the other hand, necessitates spatial alignment of the transmitter and receiver to begin message signal exchange. BSs and users ought to scan the available antenna configurations till they find one that allows them to communicate. If this phase is not properly managed, it can introduce significant HO latency and even the initial access thereby impairing delay-sensitive applications [44]. Thus, smart HO schemes are required to handle the directional cell discovery procedures, maintain a reasonable discovery delay, keep communication seamless and continuous for mobile users.

2.1.3.1 Cell Discovery Challenges and Opportunities

3GPP has adopted three hierarchical cell discovery forms, one of which involves a UE receiving beam measurement from multiple candidates of gNodeBs. The UE then notifies the serving gNodeB of the good target gNodeBs. Here, the UE antenna is pseudo-omnidirectional, whereas the gNodeB antennas are coarse wide beams, such as SSB beams.

An alternative procedure involves a sequence where the gNodeB chooses a single reported transmit beam as part of the new protocol and sends out numerous narrower candidates beams within that beam. The UE then gathers beam measurements from each transmission angle on the candidate/target gNodeB. Additionally, the UE notifies the serving gNodeB that it has successfully received a transmit beam or beams from the target gNodeB.

In the final adopted procedure, the gNodeB picks one of the transmit UE beams that the UE has reported and repeat that beam numerous times. The UE iterates through a number of narrower candidates beams in search of the best received beam from the gNodeB. Improvements to Small Cell performance, beam refinement based on SINR, and overhead/latency in beam management are just a few of the new features added in Rel-16.

2.1.3.2 5G Millimeter Wave Operation Architecture Adoption

In the initial release, 5G used multiple connections to support mmWave systems. However, with time, the cell sites that use them will function autonomously. Nevertheless, for some time, operators will stick to simultaneously deploying mmWave sites with other band assets. For instance, Enhanced Dual Connectivity (EN-DC) has been the pioneering technology that introduced mmWave to consumers. This allowed the initial release of the 5G NR 3GPP standards to be implemented on schedule. Specifications for the 5G Next Generation Core Network were absent till after the 2017 3GPP release. To that effect, there are six main options (and several minor ones) for bridging 5G and LTE networks for commercialization purposes.

One of them is the early specifications proposed "non-Standalone" configuration. Here, a 5G mmWave BS (gNodeB) connect to an LTE cell site (eNodeB) and use the enhanced LTE core network (EPC). This is the technology behind the worldwide rollout of 5G mmWave networks. When the Next Generation Core network's standardisation was finished, Option 2 included a 5G mmWave network (RAN and core). Option 3 involves user’s carrier aggregation being covered by

a macrocell that also contains one or more non-collocated small cells. This option works well if the mobile company wants to meet the needs of both LTE and 5G NR users while they coexist. However, the latency requirements are not met because the distance between the cells is usually too big [65]. In addition, the mechanism of Choice 3 is also dual connectivity. However, the concept was never widely adopted for LTE.

Option 4 incorporates elements of Option 3 but does not provide direct connectivity between the EPC and SN. Locations that don't share a physical space or use the same underlying technology can benefit from this (such as LTE and 5G NR). The implementation was pricey because the efficiency of lower-level protocol coordination, necessitates close interworking and minimal latency.

The other options involve NR Carrier Aggregation. To utilise both mmWave and LTE resources at once, NR carrier aggregation (NR CA) was implemented. It paves the way for operators to "refarm" (repurpose) some of their low- and mid-band spectrum assets for LTE. Due to the development of Dynamic Spectrum Sharing (DSS), NR CA expands the options available to operators with limited spectrum. Particularly in bands that want to use Cognitive Radio in LTE and 5G resource management i.e., by allowing them to strategically reshare the resource between primary and secondary users. Given the benefits of NR-CA, operators are expected to gradually increase their use of repurposed spectrum for NR as the number of 5G devices in use increases [47]. There are two keyways in which dual connectivity differs from carrier aggregation. To begin with, unlike dual connectivity, aggregation occurs at the media access control (MAC) layer. Secondary cells (SCells) can be rapidly activated and deactivated based on need, thanks to the MAC layer's coordinated response. [56]. Further potential remedies for issues with spectrum management in NR CA can be found in the form of cognitive radio networks (CRNs). Particularly via DSA four related functions, spectrum sensing, spectrum decision, spectrum sharing, and spectrum mobility, are introduced in [50, 51] to aid in effective spectrum refarming for NR CA. Carrier Aggregation and Dual Connectivity (DC and CA Channel Assignment per Band) are designed to offer a diversity of UL connectivity to make up for differences in coverage between bands. For PDSCH and PUSCH, this can be done with either method, but DC still relies on the consistency of the control and common channel bands. CA will be much more effective at

expanding coverage if it relies less on UL, which is only possible if the necessary coordination is established. This means there are currently two separate 5G mobile network paradigms in use.

2.2 5G Handover

As a user moves across cell regions, the communication must transition seamlessly from cell to cell. This process is known as handover (HO). Improving HO procedures is critical for mmWave links, particularly for 5G network as discussed in the preceding sections. In comparison to legacy technologies, mobile users in 5G mmWave network are sensitive to obstacles and are subjected to a much stronger shadowing effect. Furthermore, human bodies can disrupt well established connection by simple rotation of their mobile terminal Specifically, movements can cause directional antennas to be misaligned. Numerous papers on this topic [11], [12], have since been done though much more work remains to be done. Some of the HO solutions involve optimizing HO triggers, several techniques have thus far been considered. Some of the solutions involve HetNets having the option of reverting to legacy technologies where mmWave connectivity is not guaranteed. However, one big challenge usually ignored is that the LTE and mmWave data rate disparities are severe. Considering data rate differences can be in gigabits/sec, the HO scheme made must first opt for mmWave connectivity to avert of data rate disparities challenges that may incur huge changes in latency. Figure 2.2 depicts the data rate differences between LTE and mmWave for the same duration. As can be noticed in Figure 2.2, the disparities between LTE and mmWave data rates is more than 100%. The availability and use more BSs can lower the dependence on LTE fallback links and improve the average data rate. However, with complex power discovery methods, and more than usual BSs involved in the discovery of a single user, classic methods end up wasting resources. To overcome the challenge, 5G schemes need smart HO schemes to circumvent the effects of obstacles and sustain network connectivity without dedicating most of the resources to few users. In fact, in scenarios where obstacles fully overshadow a user, several users usually get caught in a fruitless cell discovery process. Therefore, 5G needs HO schemes that are intelligent enough to determine when to abandon the search and switch to LTE and vice versa when the data rates are too low. We therefore investigate on these issues in the sections below.

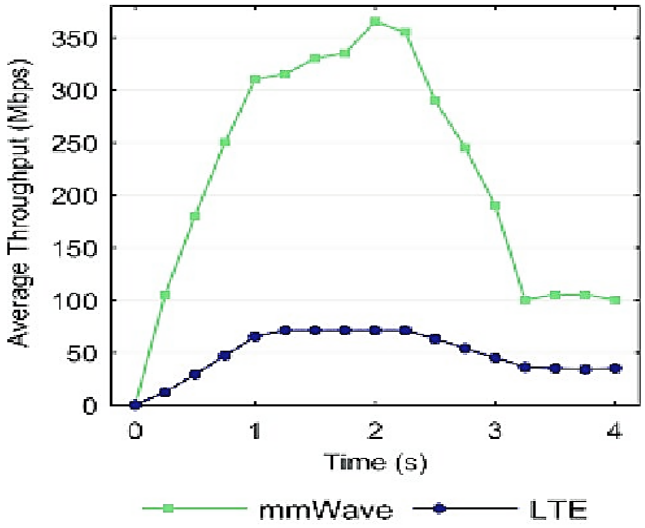

Figure 2.2: Average throughput difference between mmWave and LTE frequency bands [58].

2.2.1 Smart HO Drivers role in 5G and the Environment

The fundamental requirement of a consumer is affordability and dependable connectivity. The improvement of Smart 5G HO algorithms will give customers practical benefits in real-time and help telecom companies meet demand and pricing requirements. However, in order for 5G to succeed, mobility management will need to shift from the centralized approach frequently used in older mobile networks to a more distributed approach with decentralized coordination. The main elements that will make 5G HO appealing are those that lead to:

- The Reduction the energy and spectrum cost of 5G transmission.
- Helps in meeting the demand and flexible enough to use 4G and 5G mobile systems.
- Integrate energy and spectrum efficient Mobility Management intelligence.
- Adjust and conforms to smarter cell discovery algorithms.
- And increases the penetration of mmWave communication in autonomous mobilities /vehicles.

The 4G mobile network infrastructure developed over the last two decade makes it inflexible to adapt and integrate with the mmWave infrastructure developed for 5G [71]. However, the HO schemes used in 4G are not viable enough to sustain connectivity in 5G mobile users. This can be attributed to the high topographic and mobility sensitivity experiences of mmWaves. On the other

hand [147] owing to large demand of BSs required per unit area in 5G, energy efficiency is also another huge topic that requires smart HO algorithms to ensure 5G communication is sustainable [138]. Furthermore, owing to greenhouse gases into the atmosphere and climatic changes, demand for smart HOs will increase to reduce energy wastage in communication and harness energy harvesting and spectrum HO schemes [102]. The fact that renewable energy sources needed to power many 5G systems fluctuate in their ability to produce enough energy due to their dependence on outside factors further increases the need for smart HO to reduce power consumption. The improvement in 5G mobility, particularly HO technology, can help in stabilizing the electricity demand in mobile communication and in limited renewal energy sources [130].

2.2.2 Smart Cognitive Radio HO for 5G

The initial plan included 5G access networks being supported by a heterogeneous architecture [67]. Using a combination of legacy and cutting-edge technologies 5G organizes microwave systems into macro cells, to offer global coverage, while a large number of mmWave small cells deployed under the macro cells' coverage offer additional capacity. One interesting architecture that 5G mobile network has introduced is the New Radio Carrier Aggregation (NR CA) paradigm. This paradigm is different from the 5G NR DC model integrating the LTE and mmWave protocol structure at PDCP layer and in the Enhanced Core Network system. The NR CA paradigm is a use the MAC layer protocol integration model of LTE and mmWave spectrum specifically for 5G. Unlike the system level integration, NR CA paradigm require smart spectrum HO in the lower layers of the protocol structure. Here is where 5G can greatly benefit from the smart Channel exchange techniques of Cognitive Radio. Thus, cognitive radio is suggested as a novel approach to address the spectrum deficit issues. For instance, spectrum refarming research makes use of dynamic spectrum access (DSA), a key enabling technology for NR-CA [110]. The benefit is that CR network (CRN) schemes can optimise bandwidth usage while maintaining other generational networks' established regulations for spectrum allocation. Due to their awareness of the spectrum, CR-enabled devices can opportunistically use licenced spectrum bands that are not in use.

CRN emerges as an appealing remedy for the spectrum shortage issue that can be used in conjunction with 5G's mmWave bands. A framework for spectrum management is described in [72, [173], in which a CRN system performs four interconnected tasks: spectrum sensing [174],

spectrum decision [145], spectrum sharing [156], and spectrum mobility [137]. In licensed spectrum bands, CR nodes look for spectrum holes during spectrum sensing [178]. The best communication channel is chosen by CR nodes using spectrum decision based on the results of spectrum sensing. Channel access among the CR nodes is coordinated by spectrum sharing [159]. One important component of CR spectrum mobility that now enforces user switching on spectrum use is called "Spectrum Handoff." It is the process of moving a data transmission that is already in progress from the current channel to another open channel. This inevitably adds more latency to user communication, which eventually has an impact on performance. Connection management processes control and modify protocol stack parameters according to the current state of the spectrum HO to make up for the inevitable handoff delay. The number of spectrum handoffs is the total number of handoffs that take place during a single data transmission session. Probability of link maintenance is another crucial factor. It is described as the likelihood that, when a user leaves a channel, the communication link will be successfully maintained. The Handoff latency is used to define the accompany latency. Effective data rate is the average amount of data that is successfully transferred between two communicating SUs during each session. This is defined as the delay caused by the spectrum handoff process. Link maintenance probability and handoff latency are two of these parameters that are crucial performance indicators for spectrum mobility in NR-CA. Additional parameters have been proposed [134,135] in spectrum refarming for 5G mmWave bands to improve mmWave spectrum performance.

2.3 Review of Relevant Research Works.

Triggering HOs and obtaining connectivity on time for a user are good features of a smart HO scheme. Moreso are those that grant good average throughput and predict correctly the channel dynamics to be experienced after HO. This is usually done by tracking and learning performance of a user's signal quality. To that effect, several smart algorithms that can track, learn and predict signal quality have been proposed for 5G HO models. Unfortunately, most HO technologies pioneered in the last decade, relied much on technologies that are almost obsolete in mmWave transmission. Some of the technologies include coordinated multipoint [181], intercellular interference coordination [152], and interference alignment [153]. These have limited gains in 5G systems owing to the fact that mmWave use directional transmission.

Additionally, despite the rich benefits that multipath and scattering have in mmWave transmission. Outage may be a larger bottleneck in optimising multipath techniques. This is particular observable when there is need to deliver uniform capacity. Thus, newer smart HO may need to look at some of these issues. We discuss HO algorithms that can be used in 5G to address some of the issues raised and beyond. In tandem with the features of smart HO, we divide the HO types into four Categories, and these include: user/context aware 5G HO algorithms, Cognitive Radio inspired HO algorithms, Statistical and Machine Learning 5G HO algorithms.

2.3.1 The Role of Cognitive Radio in 5G

The challenge in 5G LTE-mmWave HetNets is that many competing propagation limitations exist when mmWaves are used. This makes sub-6 GHz cellular links such as LTE faced with the need to carry much of the load in densely populated areas. The same situation is obtained among fast moving using the mmWave-LTE dual connectivity HetNet model [34]. In fact at the very least, sub-6 GHz cellular systems need to coordinate numerous mmWave handovers [35] while also provide fallback links. This increases the overhead and affects the performance. In addition, because of the abstraction provided by functional separation, resource management, service requests, and resource allocation can be performed on multiple bands at different levels of the protocol layer. To create reliable HOs, the mobile node and access BS communication must be managed for link reliability seek, and the QoS achievable for data transmission at every other possible BS must be reliable. Despite competing factors, CR spectrum sensing and DFS/DCS can assist in identifying spectrum sub-bands and mmWave beams with the lowest instantaneous traffic load in NR CA due to its reliability.

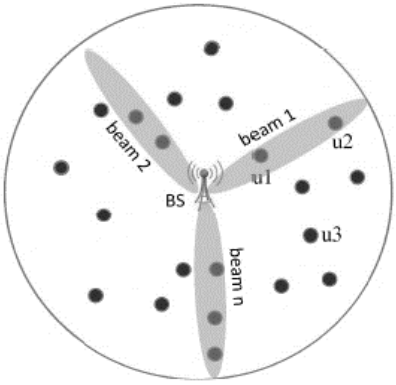

Figure 2.3: Random beam selection in a mmWave-CR-NOMA communication network [60].

For low level coordination between lower bands and mmWave BSs in HetNets, spectrum sensing and DFS/DCS play a crucial role in spectrum sensing. DFS can be deployed to use one of the free channels and set a direct path between the mmWave user aligned to BS beam. As depicted in Figure.2.3. for u1 or u2, and the user outside the beams, such as u3. CR schemes can make u2/u1 become a relay node for the serving BS to reach u3. One crucial example where CR can play a role is the mmWave-NOMA scheme implemented for downlink transmission in [36]. It enables one base station to communicate with multiple users. NOMA ensures that both users that fall within a specific beam, such as users u2 and u1, are simultaneously scheduled to use mmWaves BS. However, such a strategy will still incur significant system overhead, especially if there are numerous users in the cell. In the context of the CR-mmWave HetNet communications solution for 5G, a useful observation is that many users are not required to offer channel data. This user's effective channel gain is fed back to the BS as its scalar effective channel gain, as opposed to the entire channel vector. Then cooperative CR estimates of channel gain for u2 and u3 are used to extrapolate the channel the behaviour of u1. CR can avoid the disadvantage of HetNet HO, which requires a dedicated time slot for switching between different frequency bands, in this manner.

It is important to keep in mind that full duplexing is a concept that can be used in these situations and has been demonstrated to increase the spectral efficiency of the joint design of uplink and downlink [37]. For connecting with NLOS users who are close to the cell edge, like the one in u2 in Figure 2.3, this kind of cooperative NOMA is spectrally efficient. In a HetNet scenario, it also enables users like u3 who are completely outside of LOS to use LTE links as a relay channel to reach BS [38]. The following example illustrates how the results are advantageous. Assume that two users who are near the cell edge can be helped by a special relay.

Four time slots are needed for transmission when cooperative OMA is used. Four time slots are required in total: two for the BS to send the data from the two users to the relay and two more for the relay to send the data from the two users to one another. Only two time slots are needed for cooperative NOMA: one for the BS to broadcast NOMA to a specific relay, and the other for the relay to transmit NOMA to the two users. Because cooperative NOMA only requires two time slots as opposed to four, we can conclude that it has a higher spectral efficiency. [40] examines the advantages of buffer-aided relaying on NOMA with a dedicated relay. Recall that the concept

of a cooperative access scheme using dedicated relays is general and has been used in a variety of situations with varying numbers of transmitters and receivers [41].

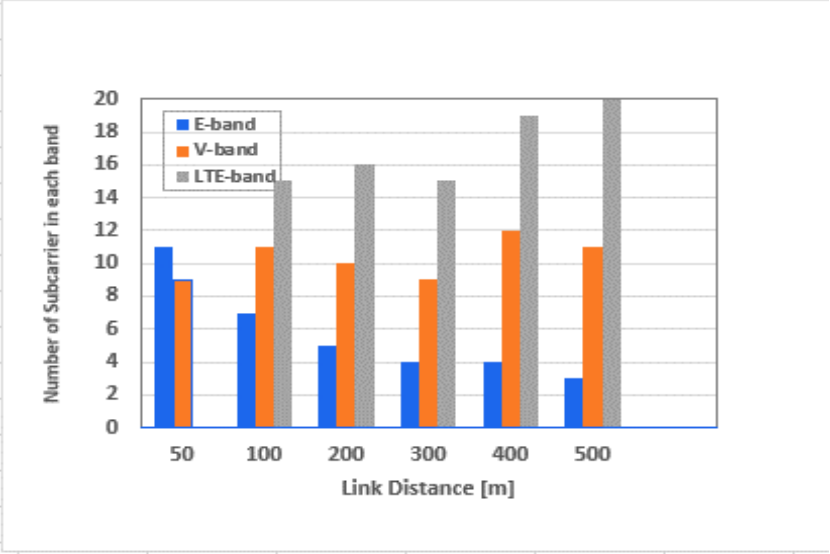

Figure 2.4: Variation of the available number of subcarriers in each band over distance [5].

Another CR application is in non-standalone New Radio (NR) for 5G. The 3rd Generation Partnership Project (3GPP) standardised it as the New Radio One (NR1). *NR1* exploits carrier frequencies in the mmWave spectrum, up to 52.6 GHz [38] for multi-connectivity options with sub-6 GHz radio overlay for coverage, (based on Long Term Evolution (LTE)). This duo connectivity is exploited at the MAC layer and is credited for a tight internetworking between LTE, and the mmWave links. Some of its applications are studies done in [32] to avoid overloading one band and maximize mmWave spectrum transmission range. The authors use dynamic access of the *E*- and *V*-band over a given transmission distance. This proved to be more spectral and energy efficient. Figure 2.4. demonstrates some of the results obtained in terms of link (subcarrier) availability of various bands over certain transmission distances. It compares the availability of lower band links (used in Long Term Evolution (LTE)) with mmWave's *E*- and *V*-band over distance. It is seen that over given distances, to balance the load and sustain the desired rate, intelligent switching between *V*- and *E*-band is not sustainable. It may need the long-ranged LTE links may to minimize *E* and *V* band's numerous handoffs. CR techniques can be easily utilized to identify 'spectrum holes or free subcarriers among *E, V* and LTE bands. This evolves around cell planning and coverage optimization with pilot power or energy adjustment [23]. CR

models can be used to bring lower layer knowledge to system level [34] HetNet decisions such as HO in NR1. Technically, this knowledge includes automatic physical cell identifier assignment and radio resource configuration in HetNets [45]. The achievement of such magnitude can enable newly deployed cell base stations to automatically get configured, be thoroughly tested, and autonomously authenticated prior to entering the operational state.

There have been several studies focusing on the CR techniques in HetNets and other mmWave-CR rapports. Mostly CR methodologies are used in deriving appropriate parameters for specific 5G mmWave scenarios. To that effect, several studies on 5G mmWave mobile [39] options and challenges with CR solution have been studied analysed and approved for future mobile communications.

2.3.2 Link Maintenance and Mobility Handling

Predicting perfect outage and coverage for mmWave transmission is a challenge and near to impossible. This is attributed to the vulnerability of mmWaves to propagation conditions. The probabilistic modelling of the system is not also easy due to limited information, vagueness, and of channels and the link blockage. Sensitivity analysis manifestations using SINR analyses in [38] based on various outage measurements and their influences on the ultimate response have shown huge disparities in terms of accuracy. This if applied in 5G may result into poor targeting and selection of channel links.

Inadequate channel link status estimates may prematurely lead to multiple handoffs during a single data transmission session and worsen overall performance in an effort to maintain user connectivity. The aforementioned situation can be made better by CR nodes using a backup channel list (BCL) mechanism [43]. Mobile nodes determine channel availability and the period of time that channel link is available by listing potential target channels into BCL and periodically updating it using shared access data among peers. Because of this, as opposed to HetNets, intelligent target channel selection [44] is made using minimal/partial spectrum sensing as opposed to traditional complete sensing analysis. BCL is used to maintain connectivity in mmWave communication because there may always be a large number of idle users, such as in networks installed in sports stadiums or convention centres. In order to assist other users in risky mmWave environments and ultimately increase 5G coverage and reliability, these idle user links can be used as BCL. The HO selection process will be the only issue when there are many BCL available [141].

A two-stage link selection scheme will be started using the Cognitive BCL mechanism. Users can initially connect to 5G based on their unique QoS requirements rather than their channel quality. Links that can guarantee the performance of the user with strict QoS requirements are specifically found and grouped into a subset in the first stage of link selection. The second stage involves choosing a subset of links from the qualified BCL that yield the highest rate for the other user. Only when there is an opportunity should this be done. It is demonstrated in [142] that this BCL selection strategy performs better than other known schemes overall, not just in terms of maintaining user connectivity. Additionally, it reduces the likelihood of a mobile outage overall. One instance of its prowess is demonstrated in Figure. 2.5. A mmWave BS with BCL users can leverage BCL to reach a user via other connected user when the direct link fails, e.g., when the user abruptly goes out of LOS in mmWave. This is unlike the current set up where a user will need a dedicated BS for HO or relay.

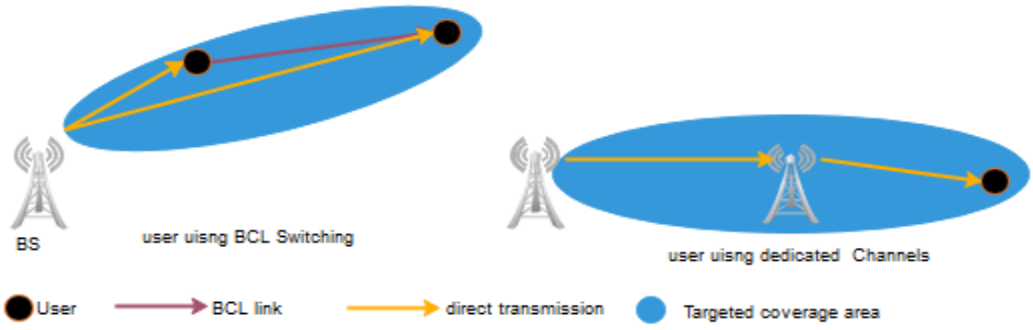

Figure 2.5: Resource allocation comparison for BCL mechanism in mmWaves and dedicated scheme [8].

2.3.3 Channel Sensing and Tracking

The mmWave channels can vastly vary over time due to mobility in mobile networks. Thus, the frequency for execution of channel estimation increases, differing from the conventional networks. This also increases Doppler spread to limit the channel coherence time too for instance. As suggested in [56], improved CSI acquisition and exchange techniques used in CRs can assist in channel learning at both transmitters and receivers to avoid such issues. The CR nodes reserve their own pilot sequences to ensure that code-domain orthogonality between participating channels is accurately estimated while acquiring CSI for UL reception and DL precoding. The different sets

of antennas in the coordination pool for channel reciprocity and estimation then exchange some of the CSI acquired from the other nodes. Thus, the HO schemes will not have to request all CSI from individual antennae. Instead, it can use the pooled CSI to make HO selection and decision.

2.3.4 Trade-off in Handoff and Beam Switch

Considering beam switch and handoff decisions are essential to mmWave link sustainability. Any decision made in favour of the other ought to be more energy-efficient given the considerable transmission energy wastage in RF chains. In this respect, power control through the efficient determination of threshold, e.g., the SINR is significant. Maximum SINR can define the allowable energy that can be consumed by user-BS for a given a resource block slot. Advanced spectrum sharing schemes have since been instituted in various mobile network management systems to maximize ultimate results and successfully control transmission power decisions. These include spectrum mobility, relaying, routing, and harvesting schemes. These schemes combine with other technical methods regarding channel state information, interference, and traffic data sharing to curb link failure in most applications**.** Therefore, understanding the propagation environment during handoff and beam switch is important.

Table 2.2. Taxonomy of CR-based techniques for smart infrastructure of mmWave network.

		HO	RELAY	BEAMFORMING	DCS/DSA
HetNet	Coverage planning	[18], [21],[32],[39],[49]	[31], [2], [1], [64]	[61], [52],[63],[64]	
	Self-Configuration	[25], [66],[61],[58]	[1], [62],[3],[4]		[61], [32]
	Load	[60],[61]	[61],[63].	[31], [32],[3],[44]	
	Power Control	[63],[64],[66]	[33], [2],[3],[4]	[22],[36],[38],[72]	[14], [29]
	Spectrum Control	[1],[2],[3],[4], [1], [2],[3],[4]	[1], [2]	[5]	[42], [43],[56]
Self-Healing	Reconfiguration Fault finding		[62], [21],[39],[43]		[59], [62],[55],[61]
Access Scheme	NOMA	[61], [2]		[24], [61],[63],[44]	[17],[11]
	OFDMA	[39],[67]	[77]	[64],[65],[69]	[17], [65],[68]

It is proved, for instance, that an advanced spectrum-sharing scheme, called bandwidth exchange in [41], is more energy-efficient than conventional cellular power control mechanisms based on SINR. This is because CR models consider factors such as transmission distance, cooperative and selective diversity gains among different participating CR users. Consequently, the most relevant research challenge is to exploit the full potential of the cognitive mmWave network. Moreover,

implanting cognitive entities into mmWave network architecture and map intelligence would enhance co-existence and neighbour awareness considering the impact of directional transmission that usually results in neighbour blindness. Thus, determining appropriate threshold is reminiscent of energy efficiency and power control decision in mmWave handoff /beam switch systems. The following Table in 2.2 lists some energy efficiency relaying and handoff and beam switch mechanisms that can potentially improve 5G application.

2.3.5 User Behaviour Aware HO Models in 5G

The resource allocation algorithm can be helped by additional user mobility data in deciding which, where, and how many mmWave cells should be set up in advance for upcoming handovers. 5G mmWave networks will be built on improved handover procedures. When compared to legacy technologies, the shadowing effect of mobile obstacles on signal propagation is much stronger. The connection can also be broken by human bodies, and the mobile terminal's rotation can lead to the directional antennas becoming out of alignment. There aren't many papers on this subject [11, 12], and there is still a tonne of research to be done. If the mm-wave service cannot be guaranteed, a number of factors, such as the potential for switching back to legacy technologies, can be taken into account to optimise handover triggers.

In [24], LOS probability functions are used to create an analytical model for network performance that is dependent on distance. The authors have shown using this model that dense mm-wave networks can significantly outperform conventional networks in terms of spectral efficiency. However, coverage gaps and transmission bottlenecks make a thorough mm-wave deployment impractical.

The pointing direction and travel distance in [50] are determined using the multi sectorial network planning (MSNP) nominal position. This is due to the understanding of the channel model and antenna capabilities. The grey polar grid serves as a representation of various antenna configuration patterns for which the mmWave BS can estimate the coverage. Notably, the path loss between the user and the BS must be completely known in order to select an appropriate beam, but path loss can only be precisely estimated using channel models [30] or anchor-based prediction systems [31].

The author proposed a number of efficient context-based cell discovery algorithms in [7] and [9]. They were also lengthened to accommodate reflected paths and obstacles. Similar to this, the

authors of [27] present an evolution of the 3GPP user association HO scheme and mention that one of the most important mmWave challenges is preventing connection losses. Mobility load balancing (MLB) and Mobility robustness optimization (MRO) were identified by 3GPP as two of the main use cases for LTE. Each is discussed separately in [34] and [35]. SON operation Cell overload is the fault of MLB. When a cell is overloaded, the load is divided among the other cells [5]. Improving system performance and end-user quality is the main goal of distributing custom traffic over system radio resources. A procedure for communicating resource status is supported by a load assessment, which is used to distribute the implementation of this function. Requests for resource status, responses, failures, and updates are just a few of the messages sent over the X2 interface that contain information that is helpful for this SON function. By setting the Cell Individual Offset (CIO) parameter, MLB was implemented. All UEs within this cell must use the offset specified in the CIO for the service and neighbouring cells in order to adhere to the transmission specifications of the A3 event.

A HO built around the optimization of user mobility parameters is called MRO. The network's HO performance is improved by choosing the best HO point based on user speed [6]. The reduction of HO-related failures is the optimization's main goal. those brought on, specifically, by handovers that took place too soon, too late, to the incorrect cell, or by Ping-Pong handovers. The main issue with the handover is that the MLB and MRO are configured to oppose Cell Individual Offsets (CIO), which is taken into account when executing the A3 event. This leads to defences for the improperly carried out handover, which causes the user to lose network access.

The author of [52] uses user-specific cell clustering to increase a 5G ultra-dense cellular network station's mobility robustness. When considering communications between vehicles and infrastructure, it is clear that connection loss and throughput degradation must be addressed. Depending on its mobility state, each device in this system is supported by a unique set of cooperative cells. In other words, the network adjusts the cell cluster size to suit the needs of each unique user. Dual connectivity and the separation of the data and control planes are used to achieve this. However, this method may have scaling problems due to user irrationality and mmWaves in 5G. [34] suggests a novel user-centric association rule-based HO as a solution to this ambiguity. By combining HO and power control optimization, the rule maximises EE. Along with the spatial randomness of user movement, the proposed mobility management system is also capable of

handling wireless channels that are temporally correlated. The suggested approach uses a tractable power control strategy, a HO time window, and mean-field games (MFG) in conjunction with stochastic geometry. In comparison to a method that uses a constant HO interval and transmit power, the proposed method establishes a baseline with a two-fold increase in long-term average EE at a typical active BS.

2.3.6 Statistical Based HO Schemes, Design strategies and Limitations

The author in [54] models the handover (HO) process for the user equipment using a Markov-based framework. It accomplishes this by using Markov statements to define various UE positions along a potential trajectory. As the UE moves along its trajectory, the path loss of the signal and, as a result, the SINR, change. The UE can choose to remain in one state, move to the next state, or return to the first state at any time. The behaviour is modelled as a Markov chain using transfer probabilities. A few Markov chain transitions starting from a specific initial state can determine the network's handover rate. Examine the impact of the aforementioned elements on the handover rate. The scheme modelled the handover process as Markov chain states using the Markov chain idea. The test case consists of a single macro base station (M-BS) and a femto base station (F-BS) separated by a predetermined distance. The user equipment is presumptively capable of entering the femto-cell boundary from the macro-cell boundary via multiple straight paths at various angles. The Markov-based framework put forth in [20] to model the user state during the handover process allows for the derivation of an ideal context-dependent handover criterion. An innovative Markov chain-based analytical framework that considers how the UE state changes throughout the handover process was used to determine the average user capacity. [3] proposes an alternative strategy that is comparable to the Markov chain. In this scenario, the mobility behaviour profile of a vehicle is created and updated using knowledge of its movement pattern. A discrete-time Markov chain built using the profile is used to model the transfer of the vehicle. Transfer is done to the cell where there is a high likelihood that the car will show up. The procedure eliminates the needless neighbour base station (BS) scanning and handover execution processes. The vehicle is also prevented from performing a redundant handover to an intermediate cell that it merely passes thanks to this procedure.

The author offers a Markov Chain Process (MCP)-based method for identifying the ideal HM, CRE, and TTT values after weighing the benefits and drawbacks of each of the three parameters in [55]. The simulation results that follow indicate that the best combinations for our scenario are 1 dB, 6 dB, and 60 ms. The first contribution of this paper is the mapping of the HetNets handover process into MCP, which allows for more simulation and prediction by allowing for the calculation and probabilistic analysis of all handover phases. Additionally, it develops a mathematical technique for modelling how HM, CRE, and TTT interact in HetNets, enabling the coordination of these three crucial elements in the optimization of the 5G system.

In The iterative Viterbi scheme, also known as the Viterbi algorithm, was first proposed in [1] and is one of the most widely used CSI-based symbol detection methods. The Viterbi algorithm is a powerful symbol detector that can achieve the minimal probability of error in recovering transmitted symbols for channels with a Markovian input-output stochastic relationship, which many real-world channels experience [2]. The receiver must have full CSI because the Viterbi algorithm requires that it comprehend the precise statistical relationship between the channel's input and output. JMLS modifies the Vertibi algorithm in [4] to account for ongoing and unexpected system changes. Over the course of its lifetime, JMLS may experience a variety of deterioration mechanisms and switch between them. In order to address the aforementioned constraint, the JMLS model is recommended in this paper as a potential strategy. Multiple degradation mechanisms, even within a single component, make it difficult to model the wear and tear of equipment. The temporal developments of deterioration indicators may differ significantly as a result of these various mechanisms. Taking into account such phenomena is crucial for both a complete assessment of the equipment's condition and a precise estimation of mmWave.

2.3.7 Machine Learning Based HO Schemes, Design and strategy Limitations

Development and application of artificial intelligence (AI) technologies have accelerated in recent years. Spectrum re-use can be improved, for instance, by implementing various cell types in 5G, with the aim of offloading traffic load based on coverage, location, environment characteristics, and user dynamics. HetNet models and AI have been combined in an effort to reduce 5G's O/CAPEX significantly. Given that the complexity of HetNets presents significant design challenges and obstacles, integrating AI-based networks and the "Self-Organizing Networks" (SON) framework with HetNets can be an effective future strategy for developing intelligent

mobile networks. Deep Learning (DL) and Reinforcement Learning (RL) are two artificial intelligences that have achieved significant success in numerous fields [96]. DL excels at perception, whereas RL excels at making decisions. Combining the advantages of DL and RL, DRL [98] provides a solution to the perceptual decision-making problem of complex systems. DRL can effectively address the problem of state space and action space continuity. It implements an end-to-end learning mode, employs original data as input and execution action, and significantly improves the algorithm's convergence and efficiency. DRL is widely used in HetNets-related applications such as traffic control [140], load prediction [130], automatic configuration [110-112], and network control [109]. As artificial intelligence technology slowly permeates more industries, networks of unmanned autonomy are gradually becoming a reality. A major player in the mobile network industry, HO, is increasing its autonomy and intelligence. The 5G HetNet makes use of a variety of AI applications, as shown in Figure 2.6.

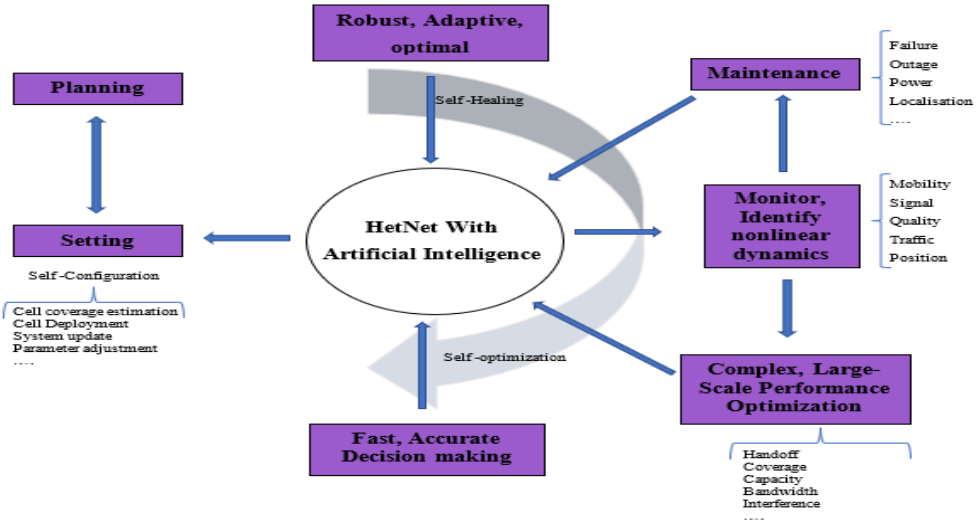

Figure 2.6: Illustration of AI-based techniques for self-organization on HetNets.

Recent interest in RL is due to the fact that it emphasises agents learning from their environment to map their behaviour and seeks the most accurate or ideal action decisions by optimising the value function. Deep Q-Network (DQN) algorithm proposed in [165] inaugurated a new era in DRL. The DQN algorithm utilises the target network and experience replay memory to increase the training process's stability, as well as the deep neural network's powerful function-fitting ability to circumvent the Q table's voluminous storage needs. The Q value for each action is the output of

the end-to-end learning approach of DQN, which only accepts original data as input. The DQN algorithm has been highly effective for discrete action, but it is difficult to implement for continuous action with high dimensions. The number of actions increases exponentially with increasing degrees of freedom if a continuously changing action is divided indefinitely many times. This results in the latitude catastrophe problem and presents significant training challenges. In addition, simply discretizing the action eliminates crucial domain structure information. The Actor-Critic (AC) algorithm [146], which is frequently used in continuous action spaces, is capable of addressing continuous action problems. The AC algorithm's network architecture includes the actor network and critic network.

The actor network generates the action's probability value. The critic network evaluates the output action. Consequently, the network parameters are continuously optimized, leading to the discovery of the optimal action strategy; however, convergence is challenging due to the algorithm's random strategy. In [127], the author proposed the Deep Deterministic Policy Gradient (DDPG) algorithm as a solution to the DRL problem in continuous state action space. DDPG is a model-free algorithm that combines the advantages of the DQN algorithm, a target network, and an experience replay memory. In addition, to ensure that DDPG can be applied to the continuous action space field, the algorithm based on the Deterministic Policy Gradient is used to convert the network output result into a specific action value. The DDPG can be utilised to effortlessly and convergently manage larger network structures and solve complex problems. [161] proposed a framework that mimics human behaviour for autonomous car-following planning. In some systems, networks and unmanned vehicles can learn from their environment through error. It has been discovered a planning model for unmanned vehicle links, and promising experimental results have been obtained. These studies demonstrate that DDPG can be used to study network behaviour, aid in the development of traffic flow models, and contribute to the creation of planned HO algorithms that mimic human behaviour.

By modifying the handover parameters to logically alter the cell size and antenna settings to shape the radio coverage, coverage optimization and load balancing can be accomplished automatically. Using coordinated adjustments to antenna tilt and pilot power among related cells, automatic coverage optimization also seeks to maximise system capacity and ensure adequate overlap between neighbouring cells.

B) Mobility optimization entails automatically adjusting handover and cell reselection thresholds in order to prevent unnecessary handovers and provide the optimal handover timing. It also aims to maximise load balancing, which can automatically force special users located on the edge of a congested cell to switch to or hand off to some less congested neighbour cells by adjusting the thresholds of the parameters. Effective neighbourhood maintenance is essential for mobility optimization; with AI, users can automatically refresh and reconfigure the list of neighbourhoods that require the fewest cells for roaming.

Link quality estimation in HetNets must always be performed with a high degree of accuracy to enable secure, robust transmissions. The use of a bio-inspired estimator based on the neural network paradigm can improve the accuracy of link-quality estimation [18].

c)Vertical Handoff (VHO) in HetNets: [22] is essential for ensuring that users receive seamless mobile service when moving between cells when using different RAN link layer technologies. Current VHO algorithms disregard the synthetic consideration of all currently available candidate networks, focusing instead on when to trigger and which connection QoS to enhance. By evaluating the complex conditions as a whole, AI-based techniques can aid in choosing the optimal parameter values. For example, [23] provides an adaptive parameter adjustment algorithm based on neural networks and addresses the VHO problem for improving user QoS and system performance in HetNets. Consult [24] and [25] for information regarding relevant studies.

2.3.8 Game Theory

Game theory is a powerful modelling tool utilized in resource optimization for interactive decision-making. It makes it possible to design flexible, low-overhead distributed systems. Users and BSs interact in a cooperative or non-cooperative manner in order to maximize their utility and achieve equilibrium. This utility can be any parameter based on the player's requirements [183]. A cooperative game enables players to negotiate while cooperating to achieve mutual benefits. In non-cooperative games, players maximize their own utility without forming any alliances. In game theory, one of the most prevalent assumptions is rationality, which implies that all participants act in their own best interests [154].

In 5G wireless networks, however, different players can have different objectives. One such example in a 2-player game were one’s objective to maximize its energy efficiency (EE) is deemed irrational by a player whose objective is to maximize its spectral efficiency (SE) [139]. Another

case was mmWave links are power limited and LTE ones are bandwidth limited [91]. Optimizing one set of networks may be regarded as irrational by the other. To optimize both we can use mean field Game Theory.

Mean-field games (MFG) and stochastic geometry are employed by the author to propose an approach that is implemented using a HO time window and a controllable power policy. The proposed method increases the long-term average EE at a typical active BS by a factor of two when compared to a baseline with a constant HO interval and transmit power. The mean-field game theory examines how individuals or interests make decisions in populations with weak interactions and extremely large sizes [91]. Over the recent past mean-field games have gained popularity in various research areas. This is attributed to its increasing significance in numerous scientific fields. The interaction between the players in mean-field games is of the mean-field type, i.e., they are coupled via their empirical measure. In [96],[97], the author expands the theory of mean-field games to the jump case networks. Particularly, this is to take accountability of abrupt changes and, different behaviours. Such capabilities can be extended to learn deterioration patterns in LTE and mmWave networks at the same time. Jump processes are indispensable in 5G due to abrupt and random behaviour of mmWaves. Specifically, when simulating sudden occurrences in mmWave in the real world. Some of the application for mean field is in financial modelling (option pricing and risk management), networks (electricity and banking), and statistics are a few examples (for modelling and analysing spatial data).

In [108] examined the mean-field game in a discrete state space whose dynamics are controlled by a Markov chain in continuous time. The representative agent control problem has been studied in a linear quadratic setting. The Hamilton Jacobi Bellman equation was driven by a dynamic programming strategy; as a result, the optimal strategy was achieved. The principal result is the demonstration that the optimal strategies for the mean-field game system correspond to 1/N-Nash equilibrium for the approximate system of N agents.

In [139], the author extends mean-field game to non-linear pure jump Markov processes that control agents in Euclidean space. This necessitates the use of nonlinear partial differential equations rather than linear transformations in Euclidean spaces. A generalization of the Koopman operator has been developed as a side effect. When examining the control problem in this context, the cost function is not required to be in linear quadratic form. It was demonstrated that the optimal

control obtained was of the Lipschitz type. Additionally, a fixed-point argument for approximating the Nash equilibrium is provided. In addition, we demonstrate that when a nonlinear pure jump Markov process is applied, the convergence rate will be of a particular order.

2.3.9 Autonomous Path Selection, Opportunities, and strategy Limitations

The industry's new research focus has gradually shifted to autonomous vehicles due to their high level of autonomy and adaptability [3]. Autonomous vehicles can work together with non-autonomous vehicles to increase productivity and road safety in addition to performing tasks autonomously in dangerous road environments. They are used in many different contexts, such as military operations, material transportation, and marine exploration. A key piece of technology for improving autonomous vehicles' autonomy is autonomous path planning. Based on specific navigational rules and user experience, autonomous path planning for autonomous vehicles must optimize the route within the safe navigation area. It has the ability to independently plan an ideal path from the known starting point to the desired destination while safely avoiding obstacles. But for that, you need a dependable, quick network. It will be essential to understand how reliable the network is along the way to the destination. Complex and variable navigational environments are a common occurrence for autonomous vehicles. Adopting a consistent and efficient method for regulating the trajectory of the car during the trip to its destination is necessary to ensure the safety of the road [4]. MmWaves have the ability to support a fast network. The research on autonomous vehicles mainly focuses on semi-autonomous task execution, autonomous route planning, navigation control, and autonomous collision avoidance.

Autonomous path planning is essential to ship automation and practical application because it serves as the cornerstone and premise of autonomous navigation [5]. In real-world navigation, vehicles frequently exchange data with one another, making it necessary to use accurate information and sane techniques to direct autonomous vehicles to avoid collisions and arrive at their destination. Planning routes should therefore take into account how to avoid unreliable networks. The technique for path planning for autonomous vehicles can guide the vehicle to take the best course of action and avoid obstacles. In order to implement the function of avoiding obstacles in the immediate area, it is also possible to divide it into the navigable area and the obstacle area, based on the obstacle information. To address the issue of autonomous path planning

for unmanned ships, many academics have conducted research and experiments. However, conventional path planning techniques frequently call for fairly thorough environmental knowledge as background knowledge, and it can be very challenging to gather such knowledge in an uncharted marine environment. Furthermore, the traditional algorithm requires a lot of computation, which makes it challenging to implement the real-time behaviour decision-making of ships and leads to a lot of inaccurate path planning information.

2.4 Summary

This chapter reviews the importance of 5G mobile networks and smart HOs. It then surveys on HO challenges in general and 5G mobile networks, the rational of deploying various handovers and motivation of the research. It goes on to review HO algorithms, schemes and their problems in 5G with mmWave networks, on four fronts and these include: Cognitive inspired HOs, user /network aware HO schemes, statistical modelling HO schemes and Machine learning HO schemes. Further, the chapter closes looking and game theory application in 5G and the importance of autonomous route selection in autonomous cars and how 5G can help improve the objectives of its research and highlights the contributions.

Chapter 3

Research Design and Methodology

A literature overview of various HO approaches applied to our 5G HO models for wireless networks was offered in Chapter 2. The diversity, contributions, and shortcomings in 5G HO protocol design were identified and discussed. While the primary goal of all mobile network HOs is to maintain continuous connectivity between mobile users and BSs, previous HO methods have difficulties or rather limitations when it comes to 5G mobile network applications. This is attributed to the advent, susceptibility, and sensitivity of mmWave mobile networks in user mobility and topographic dyanmic conditions. As a result, we chose a comparison performance-based approach with well-thought-out HO schemes. Chapter 3 frames the research design and technique used to achieve our thesis research objectives . In section 3.1, we present an overview of the research approach chosen and assesses its applicability in our investigation. Section 3.2, specifies the collection of research methods used in this study. Section 3.3 discusses the benefits and drawbacks of network simulations that are often utilised in this study field. Section 3.4 recommends using a simulator in this study. Section 3.5, elaborates about the simulation

environment and setup provided. This is then preceded by discussions about simulation validation process we used for this research in Section 3.6. Finally, Section 3.7 contains a chapter summary..

3.1 The Research Methodology

The research framework to choose is determined by the nature of the HO investigation. The goal of this thesis was to provide novel HO models and protocols for 5G mobile wireless networks while taking into account mmWave propagation reliability. In other words, the goal of this research was to create a HO model that solves a known research 5G problem. To that end, the design science research technique was used in this study design science research methodology (DSRM). Figure 3.1 displays the major DSRM stages and their uses in this study.

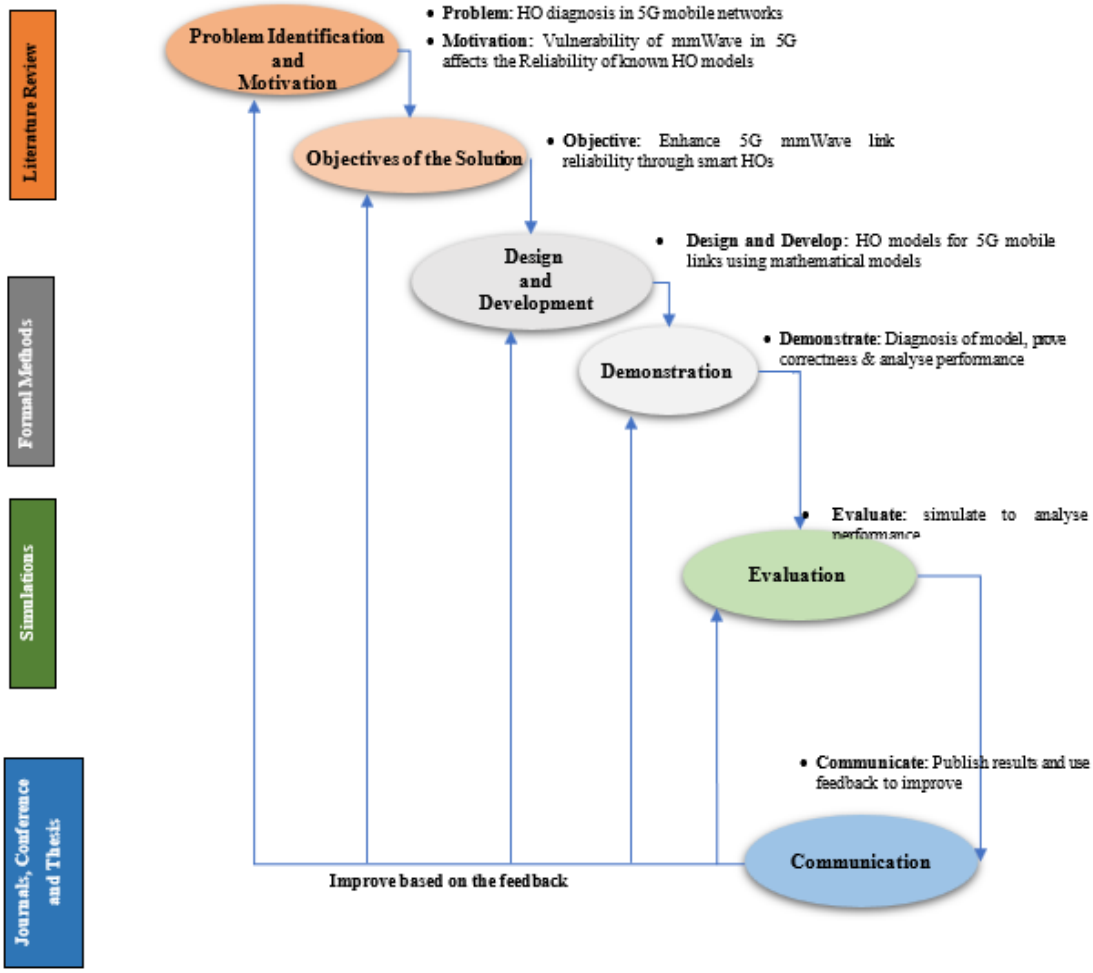

Figure 3.1: Enhanced Waterfall Design Method using Design Science Research Process (DSRP) principles.

The challenges addressed in this thesis motivated the selection of this methodology. From DSRP, we deduce its generic framework to formulate an enhance waterfall Design Cycles to fully explore our research objectives. Beginning with literature review, we extensively analyse and evaluate constructs, models, methods and theories used [22]. This culminates in six concrete step that include: Problem discovery, and its motivation, objectives, solution design and improvement, demonstration, evaluation, and communication [22].

3.1.1 The Identification and Motivation of Challenges.

The study looked at HO models/schemes challenges in sustaining the 5G connectivity between mobile users and mobile networks. Moreover, we considered on discovering, analysing and selection of target mmWave nodes prior to link failure in 5G mobile wireless networks. Considering the use of mmWaves in mobile networks has gained popularity, prioritizing there reliability is one challenge that 5G and beyond mobile networks ought to embark on. It is worth noting further that mobile networks unlike fixed networks are error prone. This makes HO mechanism even more critical. One of the most effective defects in 5G is that while mmWaves can provide higher data, they have short transmission range, are highly sensitive to user and topographic dynamics and directional in their transmission. This makes it hard for 5G to switch between BS in mmWave system over long-distance connectivity.

Providing smarter HO schemes enables selection of reliable mmWave links as HOs sustain continuous connectivity despite unreliable mmWave links. Thus, the primary motivation behind this research is to provide smart HO models for 5G mobile wireless networks. The HO models should consider the user' requirements, mmWave vulnerability and its communication capability. Based on these factors, a comprehensive literature review showed deployment of unreliable HO protocols in 5G. Moreover, it revealed that the existing models were restricted in optimising 4G microwave system such as the LTE networks.

3.1.2 Objectives of the Solution

The fundamental objective in our research was to enhance the reliability of 5G mobile wireless networks by providing smart HO schemes. The specific objectives are:

- Proposing new HO models that take into consideration the characteristics of mmWave mobile wireless networks.

- To investigate the capability of current HO models, diagnose their efficiency in 5G networks and their limitations and the compare with our proposed HO schemes in a 5G mobile environment.
- Based on user behavior patterns, network, statistical and machine learning data models, design, simulate and evaluate various HO solutions. The proposed schemes provide a wide range of solutions to combat the identified research gaps in 5G mobile networks.

3.1.3 Design and Development

This research introduces novel HO schemes for 5G mobile networks. These models consider the user and mobile network requirements in terms of.

- data rate
- latency
- energy efficiency,
- Spectral efficiency
- Link and HO failure rate, and dwell time of user connectivity per BS

Development phases defines four HO models proposed:

- First, we developed a user pattern/behaviour-based HO model. This one-use user behaviour pattern to identify the reliability of the target link/node in an instance where the serving node becomes faulty or out of range due to user movement, obstruction or signal instability. The adopted model considers user patterns and distance from target and serving cell to guarantee link reliability after HO. To that effect, the proposed model overcomes several limitations faced with Time to trigger HOs including too late, too early or wasteful HO experienced by traditional MRO models. Notably, this model becomes unreliable in instances where abrupt mmWave link unreliability/NLOS manifest from moving users. The scheme's possible links increases exponentially along increase in mobile users obstructing the line of transmission. Even when the proposed model does not guarantee convergence in most cases it is more reliable than traditional HO models except that it aways assumes a linear behaviour in target links
- Secondly, we developed an Artificial Learning HO model that employs statistical modelling to enhance HO techniques. This model overcomes system overhead using meta

data, reducing the number of training messages required by Deep Learning. It is proposed to predict intermittent mmWave link behaviour in a spectral and energy efficient way that need multiple testing rounds to be detected.

In the third scenario, we use mean field Game Theory and Jump Markov Linear System to predict the behaviour of mmWave links in HO. The mean-field game theory is the study of strategic decision making (in this case HO decision) in very large/diverse population of weakly interacting users/ individuals. Jump processes on the other hand are a very important tool in many areas of applications. Specifically, when modelling abrupt events appearing in mmWave link behaviour. Thus, the main purpose of this HO is to extend the theory of mean-field games to predict jumpy mmWave cases behaviour in both abrupt and gradual changes to the link state. Here we study the mean-field game in a finite state space of user dynamics and mmWave network changes where the dynamics of the indistinguishable BS agents is governed by a controlled continuous time Markov chain.

In the final case, we employ mean field and DRL integrated models to design a user route network availability aware HO scheme. This is very applicable to autonomous cars where they need to be always connected but are not aware of the reliability of the network on the route chosen destination

3.1.4 Demonstration

We illustrated how the suggested learning models could be employed to design various HO schemes for 5G. These principally diagnose and predict different levels of reliability of target mmWave links/cells. To that effect, the parameter considered included the user speed, system overheard, data rate, latency, spectral and energy efficiency etc. under different network topologies and mobile networks. In line with the above, the HO schemes were developed, and their viability in5G were proved.

3.1.5 Evaluation

During this stage, the HO's performance was evaluated based on a number of critical parameters. These include the communication overhead, data rate, latency, spectrum efficiency, and energy usage. In addition, the proposed methods' efficacy in various settings and population composition was evaluated by a comprehensive set of simulations run in the ns 3 simulator.

3.1.6 Communication

As a form of getting feedback, Journal articles and conferences were submitted to high-impact, research journals. Fellow Researchers could share their knowledge with regard to our findings and academics. Further constructive criticism was provided where necessary. The significance of this study was also emphasized in talks given at numerous academic conferences and seminars. With our research culminating into this dissertation We expanded results of 5G studies, explained the significant of HOs and offer recommendations for future study.

3.2 Research Design

The study design lays out the steps that will be taken to address the research questions. In particular, this part provides examples of the research methodologies that were used, explaining why they were chosen and providing further detail on how they were put to use. The purpose of this investigation was to examine the nature and difficulties of 5G mobile wireless networks and HO protocols. So as to enhance the dependability of 5G mobile networks, we proposed and researched the performance of proposed HO models. Different HO methods that have been frequently employed in wired and wireless networks were revealed by the literature review. We analysed these methods critically and spoke about their benefits and drawbacks. The comparative method has been frequently used to pinpoint inefficient mobile networks and has been identified as one of the most common diagnostic approaches. While this method shows great promise, it has few practical applications in today's mobile wireless networks. Based on our research, it is clear that typical HO models have limitations that prevent them from being used effectively in 5G mobile networks. By sifting through the existing literature, we were able to spot the holes in our knowledge and pave the path for the creation of novel HO models that take into account the needs of mobile wireless networks. The strategies employed in order to realise the study's aims are discussed in the next two parts. In the studies uncovered by the literature study, several techniques have been applied.

3.2.1 Simulation

Designing, implementing, and assessing computer and communication networks have all made heavy use of simulation in addition to theory and experiment [120]. Interest lies mostly in the simulation research approach since it allows the researchers to zero in on the study idea and learn

it thoroughly [124]. Systems and their functioning can be modelled using this approach [125]. To do this, you first create a model that stands in for the real system, and then run tests on the model to learn about the system's behaviour and rate its effectiveness in various settings [125].

Comparing simulation to other research methodologies, it is clear that simulation excels [126]. The main reason for this is that researchers can save both time and money by using simulation software to simulate an actual system rather than having to set up a prototype. Simulation also makes it possible to test out new strategies without the high cost and effort of implementing them in the real world. In addition, simulation allows us to see how changing settings impacts the system's efficiency. In addition, simulation simplifies research into intricate systems.

When studying mobile wireless networks, simulation is by far the most used approach [127]. The performance of protocols in mobile wireless networks has been the subject of extensive research and analysis [128]. Complex systems like these networks make it tough to test and deploy new strategies [129]. Research techniques like these are useful for modelling real-world systems and measuring how changing inputs affect outputs [130]. The simulation approach also allows for the modelling of extremely extensive wireless networks [124]. It also allows for the study of a wide variety of characteristics that may have an effect on network performance, such as wireless connectivity and node mobility.

3.2.2 Selection of Network Simulator

A network simulator help recreate the operation of a network in a virtual environment without the need to build or deploy the real thing. Using a wide variety of tools for network modelling and simulation, it has been possible to run simulations of mobile wireless network architectures [132]. The OMNeT++ [133], NS-2 [134], NS-3 [135], and NetSim [138] programmes are some of the most well-known examples of network simulators [132]. These simulators represent a system by modelling it as a succession of discrete events, and as such, they are examples of what is known as "discrete event simulation" [130]. Because each simulator comes with its own set of advantages and disadvantages, it is essential to consider a variety of aspects before deciding which model to implement in the course of your investigation [139]. The table that follows provides a summary of the primary features shared by these favoured simulators.

Table 3.1: Network simulator comparison [139].

Simulator	License	Latest Release	Language	Scalable	GUI	Mobility
OMNeT++	Open-source	2019	C++	Yes	Yes	Yes
NS-2	Open-source	2011	C++/OTcl	Limited	No	Yes
NS-3	Open source	2019	C++/Python	Yes	Limited	Yes
NetSim	Commercial	2019	Java	Limited	Yes	Yes

The following sections breaks down the study challenges of HO in 5G mobile networks, and the leading network simulators that have been utilised in this research. Furthermore, we will single out the top network simulators that have been employed in previous studies of this issue. Because of their prevalence in the academic and research communities and their general popularity, open-source simulators are our primary emphasis. Among the simulators that fit this description are NS-2, NS-3, and OMNeT++.

3.2.3 The NS-3 Simulator

In 2006, an open-source discrete-event network simulator, the NS-3 replaced the NS-2 simulator. The programming languages C++ and Python are used in tandem to create NS-3. Models in NS-3 are more accurate because they are written in C++. In comparison to NS-2, it offers significant enhancements in terms of scalability, modularity, and code readability.

3.2.4 Justification of the choice of Network Simulator

NS-3 excels in several ways over competing simulation environments [143, 145]. In the realm of network simulation tools, NS-3 is highly regarded [143]. It offers a robust graphical user interface and animation, just like commercial simulators, but it is provided for free to academics. Additionally, it is user-friendly and has robust third-party model framework compatibility.

Mobile wireless networks can take advantage of the Ns 3 framework's strong integration support because it provides a wide range of protocols and models for use in those networks [142]. In

addition to supporting the netSim GUI, it is also compatible with Google Maps for tracking a user's location and mobility in real time. In addition, it has Reinforcement Learning capabilities thanks to its integration with OpenAI Gym. Research uses of the reinforcement learning (RL) tools OpenAI Gym are numerous. For research into networking protocols and communication technologies, ns-3 is the gold standard among academic and commercial institutions. ns3-gym is a framework that combines OpenAI Gym and ns-3 with the goal of promoting the application of RL to networking research.

Table 3.2: Usability, architectural and performance comparisons of OMNeT++, NS-2, and NS-3 [143].

	OMNeT++	**NS-2**	**NS-3**
User support	Excellent	Discontinued	In progress
Learning time	Short	Long	Short
Integrability	Excellent	Limited	Excellent
Reusability	Excellent	Good	Excellent
Testability	Good	Limited	Excellent
Flexibility	Excellent	Limited	Excellent
Complexity	Low	High	Low
Memory usage	Average	High	Low
CPU usage	Moderate	High	High
Computation time	Moderate	Lowest	Moderate

3.2.5 More on NS3 Simulation Environment

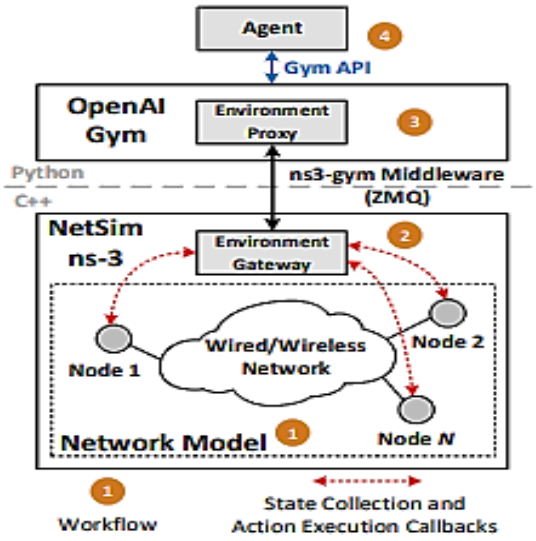

Figure 3.2: Architecture of a ns 3 and Open GYM simulator [146].

Fig. 3.2 depicts the architecture of ns3-gym, which comprises of the ns-3 network simulator and the OpenAI Gym framework. The former is employed during environment implementation, while the latter standardises its user interface. This work's main contribution is an interface between OpenAI Gym and ns-3 that enables concurrent operation of HO frameworks in 5G. Here, we provide a comprehensive explanation of the ns3-gym infrastructure.

It is important to note that ns-3 is integral to our overall architecture. It is the method of choice when creating a virtual world in which a reinforcement learning agent can operate. A scenario in simulation includes a representation of the network and predetermined alterations to the simulated environment. The precise models of communication components and channels shown in Figure 3.2 allow for the creation of even very complicated network models, which can then be studied under a wide range of traffic and mobility scenarios. Changes in conditions, such as the beginning and ending of traffic sources, are triggered during a simulation by events that the researcher predetermines.

The primary goal of the Gym framework is to standardise the means through which a networked environment's state can be accessed and actions can be performed. It is important to keep in mind that the simulation scenario itself serves as the sole source for defining the environment, making the Python code completely portable. Therefore, the simulator can easily swap out the agents'

implementation while maintaining the reproducibility of the environment's conditions. Connecting the ns-3 network simulator to the OpenAI Gym framework is the job of the ns3-gym middleware. It handles communication between the Gym agent and the simulation environment, including the transmission of status (i.e. observations) and control (i.e. actions). The software has two components: Environment Gateway and Environment Proxy. The simulator's gateway collects information from the environment and converts it into organised numerical data. It then translates the actions received from the simulator into calls to functions with the correct arguments. The proxy obtains information about the environment and makes it available to an agent using the Python Gym API. The researcher is responsible for defining the semantics of the transferred state and actions in ns3-gym middleware, as they are transferred as numbers.

3.3 Implementation and Validation of Simulation Models

Fig. 3.2 depicts the architecture of ns3-gym, which comprises of the ns-3 network simulator and the OpenAI Gym framework. The former is employed during environment implementation, while the latter standardises its user interface. This work's main contribution relies on the concept interface between OpenAI Gym and ns-3 that enables concurrent operation of HO frameworks in 5G. Thus, we provide a comprehensive explanation of the ns3-gym infrastructure.
It is important to note that ns-3 is an integral to the overall network architecture. It is the method of choice in fact when creating a virtual world in which a reinforcement learning agent using open AI gym can thrive in mobile communication. The simulation usually includes a representation of the network and predetermined alterations to the simulated environment. The precise models of communication components and channels shown in Figure 3.2 allow for the creation of even very complicated network models, which can then be studied under a wide range of traffic and mobility scenarios. Changes in conditions, such as the beginning and ending of traffic sources, are triggered during a simulation by events that the researcher predetermines.
The primary goal of the Gym framework is to standardise the means through which a networked environment's state can be accessed and actions can be performed. It is important to keep in mind that the simulation scenario itself serves as the sole source for defining the environment, making the Python code completely portable. Therefore, the simulator can easily swap out the agents' implementation while maintaining the reproducibility of the environment's conditions.

A crucial step in ensuring the validity of simulation findings is ensuring that the models used in the simulation have been properly verified and validated. The wrong settings can lead to misleading findings [145], despite the fact that n3 and open AI are reliable technologies. The ns3-gym is a set of tools comprised of two separate modules, one in C++ and the other in Python. To facilitate communication between ns-3 and OpenAI Gym, it utilises extensions to those systems. While common ns-3 models are used to construct the simulation environments, popular ML frameworks like Tensorflow and Keras can be used to create the agents.

In order to specify the observation space and the action space, the functions GetObservationSpace and GetActionSpace are called. They are only invoked once, at the start of the environment's life cycle. The following functions are called by the framework at various points throughout execution to gather the current status of the environment:

First, there's GetObservation, which can be used to retrieve the current values of any observed variables or parameters at any network node across all network protocol layers.

Step 2: Call GetReward to tally the benefits of the previous action.

The third method, GetGameOver, is used to verify an in-game end state.

Fourth, GetExtraInfo retrieves more data about the present environment.

Note that our framework's step can be triggered either by a predetermined time-interval, such as every 100 ms (time-based step), or by a specific event, such as packet loss (event-based step).

When analysing HO performance in a mmWave cellular network, it is crucial to take into account every single component. This is especially true because to the intricacy of interactions between the underlying LTE and mmWave physical propagation phenomena and the whole protocol stack. Since there have been no large-scale deployments of NR mmWave cellular networks as of yet, system-level simulators provide ideal candidates for end-to-end performance testing. Using the 3GPP channel model for frequencies above 6 GHz [11] and a 3GPP-like cellular protocol stack [12], researchers from New York University and the University of Padova have developed a mmWave cellular network module for the widely used ns-3 simulator [9].

In this dissertation, we use open AI Gym to induce a learning process and expand the capabilities of the model presented in [9], which has been tested for SINR-based HO. This opens the door for the employment of user-aware, statistical, and Machine Learning HO models to be deployed at each gNB in a 5G network. We also include the option of replicating 3GPP-compliant network-

pattern-aware HO systems. Our framework mimics the radiation pattern of a patch antenna element and offers the precise antenna gain in each direction in order to provide an accurate tool for formulating HOs and evaluating the end-to-end performance in mmWave mobile scenarios. Additionally, we describe simulation results using the realistic user mobility model, which takes into account end-to-end transport protocols (i.e. UDP) and a multi-site cellular network deployment. In particular, we describe the impact on system throughput and delay of changing the HO model used to govern the gNB, as well as the user and gNB densities and the 3GPP deployment scenario. Our results demonstrate that even in the most challenging channel environments, throughput and latency may be improved by selecting gNBs with better Signal to Interference and Noise Ratio (SINR) and longer connectivity.

3.3.1 Development Cycle of Simulation Models

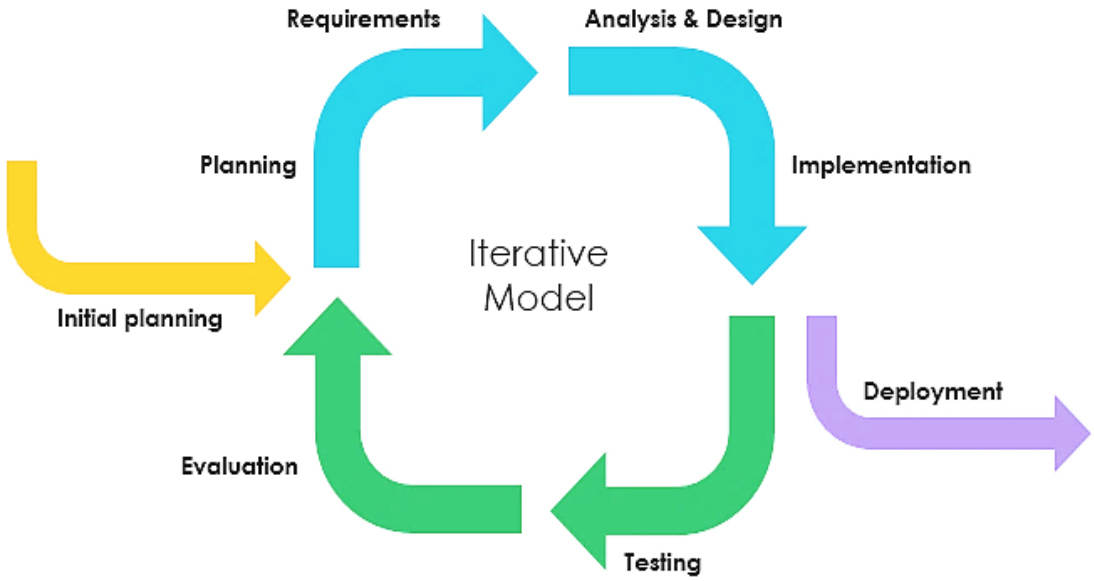

Figure 3.3: Model development process [147].

This research followed the model development process described above and shown in Figure 3.3. First, a conceptual model that represents the system was developed. Formal methods were utilised to validate the conceptual design. Notably, the correctness of protocols was tested. In addition, the performance of the proposed protocols was evaluated using standard methods. The procedure of model development that was outlined earlier and illustrated in Figure 3.5 was adhered to

throughout this research. To begin, a conceptual model of the system that accurately depicts it was built. Validation of the conceptual design was accomplished through the application of formal methodologies. Notably, the appropriateness of protocols was validated through testing. In addition, the performance of the suggested procedures was assessed by employing the approaches that are generally accepted.

In the end, the operational validation was implemented to guarantee that the results and outputs of the model were accurate. In specifically, the simulation tests were planned based on the requirements of 5G; they were created; they were carried out; and they were analysed. In addition to that, the results that were acquired were compared to the outcomes of previous procedures conducted under the identical conditions. In addition to this, confidence intervals were measured so that the range of accuracy for model validation could be described. For the purpose of calculating the confidence interval, the degree of confidence known as 95% is the one that is utilised in practise the most frequently. After that, we determined whether or not the sample values are contained within this interval; if they were, we deemed our model to be accurate.

3.4 Chapter Summary

The procedure of model development that was outlined earlier and illustrated in Figure 3.5 was adhered to throughout this research. To begin, a conceptual model of the system that accurately depicts it was built. Validation of the conceptual design was accomplished through the application of formal methodologies. Notably, the appropriateness of protocols was validated through testing. In addition, the performance of the suggested procedures was assessed by employing the approaches that are generally accepted. In this chapter, the methodology that is being used in this research, known as DSRM, was described. It also discussed whether or not this methodology was appropriate for use in the conduct of this investigation. In addition, the chapter provided an overview of the research design and described the research methodologies that were utilised, including formal methods and simulation. The majority of relevant previous studies have made extensive use of these methodologies. The ns3 simulator was chosen to be used for modelling, simulation, and analysis of our proposed protocols. Following this, a brand new head end (HO) model for mobile networks is presented in Chapters 4-7. The diagnostic model that has been proposed takes into account the inherent qualities of mobile networks, such as the fluidity of their topologies and the locations of their users.

Chapter 4

User-Adaptive Based Mobility Pattern Aware Handoff for 5G Mobile Networks

The mmWave communication is the main capacity booster for 5G mobile network. Yet, as reviewed in chapter 2, mmWave transmission is highly susceptible to user movements and topographic dynamics. To minimise the impact on link sustainability for 5G mobile users, the LTE-mmWave dual connectivity (DC) architecture was adopted for 5G. Its split control plane and data plane functionality mutually benefits the operations of LTE and mmWaves link transmission. The DC model uses the less attenuated LTE microwave links to coordinate mmWave cells while they provide on demand high-capacity rates. The LTE links on the other hand, guarantee 5G mmWave mobile users, long ranged network association and fallback links when mmWave links fail. Furthermore, mmWave links incur minimal signalling cost as the control plane data is largely managed by its LTE counterpart. Minimizing signalling cost is vital for freeing enough bandwidth for user data transmission. One way involves minimizing user fallback to LTE bands and prolonging user association with mmWave links. Owing to their long-ranged transmission, LTE links have better link quality most of the time, however, the microwave bandwidth in LTE is

scarce. Moreover, the optimum LTE data rate is very minute compared to that of mmWaves. However, the challenges are while the mmWaves offer gigabit data rates, propagation usually offer patchy coverage and irregular cell pattern [89]. To that effect, in this chapter, we proposed an intelligent pattern-based Handoff (HO) strategy to prolong mmWave link association for mobile users in the 5G DC HetNet systems. The long association of mmWave links avails an opportunity for 5G mobile communication to provide uniformly high data rates and, improve the signalling cost owing to less users falling back to using LTE links. At the same time, it frees the much need bandwidth in LTE links for control plane management, low data rate applications and fallback HOs. During the HO selection process of target link(s), the proposed HO model looks at:

(i) The potential distance the user will be in LOS with the target BS after the HO.
(ii) How gradual or abrupt the deterioration pattern of the mmWave target link will be after HO given the NLOS and LOS susceptibility of mmWave transmission. The irregularities of cell pattern due to patchy coverage in mmWave cells usually distorts cell patterns. This inevitably distort network performance and guarantees over a given distance/time/direction.

The knowledge of how the mmWave target link will deteriorate after executing a HO is vital for estimating its reliability after HO. And must be carefully predicted to avert selecting links with intermittent connectivity: abrupt/drastic performance changes immediately after a HO execution. Unlike the microwave bands in LTE that can gradually deteriorate with distance, mmWave deterioration is highly stochastic: it is punctuated with random, sometimes abrupt, gradual and drastic performance deterioration or improvement. Knowing the deterioration pattern of a mmWave target link avails the HO scheme an opportunity to estimate how consistent/uniform the data rate will be on the target link in the long run after HO. However, this requires CSI collected to be accurate to which is not always the case owing to rapid changes that accompany mmWave communication. We propose to predict the mmWave deterioration patterns using the “Jump Markov Linear Systems” (JMLS) [66]. JMLS is known to take into account of abrupt changes in system dynamics when predicting future behaviour [56]. Since mmWave transmission is prone to abrupt changes. JMS models can predict the continuous behaviour dynamics within each operation mode by using difference or differential equations. And abrupt changes between different modes the JMLS model using Markov chains [76].

4.1 Background

Statistical Markovian-based models such as Hidden Markov Models (HMMs) [233] and hidden semi-Markov models in [124] formulate some of the HO models in mobile networks. On the list of continuous-state processes are the Lévy process [125], the general path model in [126], etc. Refer to literature review for a comprehensive review of deterioration models used in reliability analysis. The major challenge among them is that in most of these statistical models, including advanced ones in [34], they presume that once a link/system has deteriorated, the deteriorating performance link never recovers. Such concepts if used in mobile networks, the solution for mobile users is usually to institute a HO [56]. Although this can be true among LTE links (not too sensitive to topographic and user mobility dynamics), in mmWave transmission, links experience a great deal of behavior changes. They can switch between different deterioration factors. Particularly the mmWave links can be partially in NLOS /LOS, recover to full LOS or fall into full NLOS condition needing a HO to sustain connectivity. Wrong prediction of mmWave link behaviour based on perpetual deterioration pattern can lead to too early or late or unwanted HOs. This increases system overhead and at worst, it causes HO failure. In this chapter, we develop the multi-objected HO model that adopts Jump Markov modelling principle to predict mmWave link reliability over distance and probable performance "switches" between different connectivity instances. The "Jump Markov Linear Systems" predicts the mmWave target link's deterioration pattern and accounts for abrupt changes using difference/differential equations. Further, it accounts for abrupt changes between mmWave modes using Markov chains. To that effect, the preceeding sections are divided as follows: The network and mobility model are introduced in Section 4.2. In Section 4.3, the JMLS model for mobile networks is presented. Section 4.4 involves Performance Evaluation 4.5 concludes the chapter.

4.2 Network and Mobility Models

The section describes the network model first. It then details the mobility model for the proposed HO model. We describe the underlying user mobility and network assumptions under which the proposed HO model provides valuable HO decisions and results.

4.2.1 Network Model

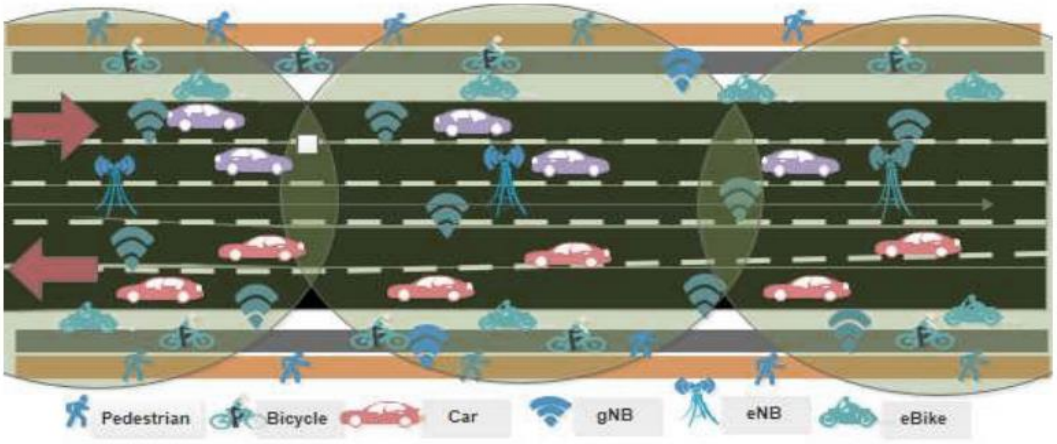

Figure 4.1: Multiuser type Mobility Model.

The proposed user mobility model involves a one-dimensional highway network. The user types include cars, cyclers and pedestrian. We use the LTE and mmWave cell deployment as shown in Figure 4.1 The transport protocol used is the user datagram protocol (UDP). The transmission range of each gNB in LOS is 200 meters. A Dual Connectivity non-standalone (DC) LTE-mmWave model is chosen. A 4G evolved packet core (EPC) network with LTE evolved node bases (eNB) controlling 5G's NR gNB at the PDCP layer is assumed. In this scenario, users do not connect to eNBs; rather, eNBs operate the control plane in respect to the core network. In accordance with configurations and models described in [2] and [3], all eNB and gNB can communicate seamlessly via X2 links. The gNBs manage user data packets via the data plane. Each LTE eNB is surrounded by ten mmWave gNBs and has a cell radius of 500m. Three LTE eNBs are positioned along a 3000 m stretch of highway for our setup as shown in Figure 2. And -5 dB SINR threshold is used for initial HO, users are initially free to connect to any gNB with the largest SINR difference. Technically, each central LTE unit manages the control table that governs the mmWave HO selection. And a constant bitrate of 150 Mbit/s content is sent to users from remote a server at the EPC. The channel conditions are modelled using the 3GPP model standards described under literature review and in [11].

4.2.2 Mobility Model

There are pedestrians, drivers of motor vehicles, cyclists, and users of electric bicycles. Within the two car lanes that are situated on either side of the road, users of automobiles move. On the other hand, riders of bicycles and electric bicycles travel along the edges of the road, one on each side. Figure 1 shows the movements they made in different directions. On either side of the pedestrian lane they are using, a pedestrian is free to walk in either direction. Three thousand metres is the total length of the road. Vehicles, such as cars and bicycles, bounce back in the opposite direction and arbitrarily choose a lane that goes in the opposing direction when they reach the end of the road. The car will automatically change its speed to match the one of the car in front of it once a new lane has been chosen. Each LTE eNodeB supports 200 full buffer users in total. Walking traffic makes up 75% of them, moving at a speed of 1.4 metres per second. On their bicycles and electric bicycles, they move at speeds of 7 and 8 metres per second, respectively. Drivers of the remaining one-third of the vehicles move between 10 and 14 metres per second. Drivers of vehicles that are more than three metres apart will slow down to between one and three metres per second in order to avoid a collision. Every three seconds, the vehicle's speed is either increased or decreased.

4.2.3 Network Resource Allocation Model

A) Outage Probability Estimation

The outage probability, P_π , given an observable set of signals,Y_k is marked by assuming Θ is a set of network parameter set for a specific access policy π[11]:

$$P_\pi(Y_k|\Theta) \triangleq P\left(\sum_l \sum_{s_k} b_l \log_2\left(1+\gamma_t(x)\right) \geq r^{m\zeta}(\hat{\gamma}_t)\right), \tag{4.1}$$

where γ_t and $\hat{\gamma}_t$ indicate actual and target SINR respectively. $r^{m\zeta}$, indicate the target user throughput for a given user state $s_t \in S$. b_l represents bandwidth for mmWave link l. We assume the gNBs transmit directionally equal power, P. The receiver sensitivity is denoted by, x_{kmin}. Each serving gNB given, P, ought to satisfy at least x_{kmin}. Moreover, given a threshold x_{k0}, where $x_{k0} > x_{kmin}$, any user-gNB link transmit power requirements must not surpasses P and must at the very least meet x_{k0} to establish a new connection or may end up losing a connection, i.e., a

truncation outage [16]. The maximum User-gNB transmission distance is $d = (P/x_{k0})^{\frac{1}{\propto^{k_L}}}$ to avoid losing connectivity. For x_{k0}, LOS and NLOS maximum distances are: $(P/x_{k0})^{\frac{1}{\propto^{k_L}}}$ and $(P/x_{k0})^{\frac{1}{\propto^{k_{NL}}}}$, respectively. Given received power x_t, is changing for a typical user at d from a target gNB, the receivable power projected is:

$$x_{t+1} = \max \sum_{x_t, u_t} \left\{ \frac{\hat{\gamma}}{\gamma^{min}} x_t - \frac{\alpha {u_t}^2}{\beta \hat{\gamma}^2} \right\}, \tag{4.2}$$

where $\{.\}^+ = \{max, 0\}$. x_t and u_t are current possible LoS and NLoS received power, respectively. $\hat{\gamma}$ is the targeted and γ^{min} is the immediate measured SINR needed to satisfy the $\mathbb{R}^m$. The parameters, α and β, denote received power and SINR scaling factors. And following user scenarios changing from NLOS to LOS or vice-versa, they account for power discrepancy in the projection. The data rate given *P* corresponding x_t, at t is defined as:

$$r^m = b \log_2 \left(1 + \frac{P|h^H \boldsymbol{p}|^2}{(1+ d^{\propto})} F_x\left(\left|\theta_k^l\right|\right)\right), \tag{4.3}$$

$$\varphi_k^l(^\circ) = \frac{1.4\ x\ 10^4}{f_c(GHz).\, v(km/h)}, \tag{4.4}$$

$$h^H p = \sum_{k=1}^{K} g_k(t) e^{-2\pi i f_d \cos\left(\varphi_k^l\right) t}, \tag{4.5}$$

where $\theta_k^l = \frac{2d \sin \varphi_k^l}{\lambda}$ represents the angle of arrival for beam *p*, v is velocity of a user moving at 50 Km/h or less. f_c, denote the carrier frequency. f_d, represents maximum Doppler shift while $\varphi_k^l(^\circ)$, denotes the main lobe angle of arrival. $|h^H \boldsymbol{p}|^2$ is channel gain [6] while $F_x\left(\left|\theta_k^l\right|\right)$ denotes the Fejér kernel function. For user speed approaching zero, $F_x\left(\left|\theta_k^l\right|\right) \to 1$, hence SINR approaches maximum over distance, *d*. Otherwise, SINR and F_x approaches 0 as v increases [4]. α, is path loss exponent [9] and their respective LOS and NLOS pathloss exponents are denoted by: $\propto^{k_L}$and $\propto^{kN_L}$, $\forall k \in \{1, \ldots, K\}$. $g_k(t)$, is the time-varying gain of the channel K clusters.

B) Data Rate and Energy Cost Estimation Problem

Presented with outage and power constraints, in user state s_t, at user distance d, from a BS the minimum rate, $\mathbb{R}^m(x_{k0} > x_{kmin})$, requirement problem can be defined as:

$$\max_{\Theta} \sum_{t} \sum_{S_t,l} \left(1 - \beta_l \left(P_\pi^{m|x_t} P_{NL_k} + P_\pi^{m|u_t} P_{NL_k}\right) r_l^m(y)\right) \geq \mathbb{R}^m, \tag{4.6}$$

where $P_\pi^{m|x_t}$ and $P_\pi^{m|u_t}$ are LOS or NLOS conditional outage probabilities for a user in the m^{th} state, respectively. r_l^m is the maximum attainable data rate for the $l-th$ link at d. β_l is a binary factor. It is 1 or 0 for either $P_\pi^{m|x_t}$ or $P_\pi^{m|u_t}$ depending on which one the user is active. The targeted received power, x_{t+1} needed at $d+1$ to sustain $\mathbb{R}^m$ condition in (4.6) given outage SINR constraints in (4.2) -(4.3) for typical user at d from a target link has a corresponding transmission energy cost, c_t, is influenced by two components: received and lost packet energy consumptions:

$$c_t = \beta \left\{ x_t \frac{c(t-w)}{\mathbb{R}^m} + e_0 * \zeta c(t-w) \right\}, \tag{4.7}$$

where β denotes the energy price unit energy, for transmitting $c(t-w)$, actual total number of packets received at *t* over window *w*. $c\,(t-w)/\mathbb{R}^m$ is the latency. x_t denotes the current received power at time *t*. e_0 is the unit energy per packet, $e_0 * \zeta c(t-w)$ denotes energy cost owing to lost packets during window *w*. Given constraints of receivable,x_t,and transmittable power, P, for optimum packet delivery latency, the problem to maximize link utility is formulated as:

$$\max \sum_{P} \sum_{x_{t,}} \{x_{t+1} \delta c(t-w) - \zeta P E_c\}, \tag{4.8}$$

where $x_{t+1}\delta c(t-w)$ denotes the optimum latency given x_{t+1} is received and packets $c(t-w)$ during window *w*. ζE_c is the latency discrepancy at time following received power drop from x_t to x_{t+1} as the user moves away (i.e., from transmission distance d_t with E_c to d_{t+1}) from the serving BS. Prior to HO execution, we learn to predict the long-term received power deterioration pattern $\{x_t, \dots, x_T\}$ of target links as user with different speeds move away from target BS. This is to avoid selecting target links whose performance is likely to abruptly dip due to sudden

blockage/NLOS effects despite showing signs of high initial performance. we utilize JMLS properties to predict the likely gradual or abrupt deterioration behaviour of target links $\{x_t, \dots, x_T\}$. The JMLS uses difference or differential equations to model continuous dynamics within each link operation state, and the abrupt deterioration patterns between the states are described by a Markov chain [117].

4.3 Jump Markov Modelling

JMLS model the mmWave link behaviour and reflect different states following which the mmWave link received power may deteriorate after HO, i.e., owing to varying user types and topographic dynamics. Letting x_t, be the receivable power as in (4.3), over a given state of a user at t, time. We first reformulate (4.1) – (4.9) into a JMLS problem form in (4.10) and then project the desired x_{t+1}, that the target gNB must meet before a HO is granted. Mathematically, we define the JMLS modelling equation as in (4.10):

$$\begin{cases} x_{t+1} = A(s_t)x_t + B(s_t)u_t + w_t \\ y_t = r^{min}(s_t)x_{t+1} + v_t, \\ \mathcal{M} = (\Theta, P(S), \pi, P_\pi) \end{cases} \tag{4.10}$$

where $x_{t+1} \in X$ denotes received power quantities in (4.3) given state, s_t, $u_t \in \mathcal{U}$ represents received power discrepancy estimates flocking any blockage/NLOS and is related to x_t by: $u_t = -Kx_t$; $A(s_t)$, where *K* is the control factor. $B(s_t)$ are SINR/power coefficient matrices in (4.3). $v_t \sim \mathcal{N}(0, \mathcal{Q}(s_t))$ and $w_t \sim \mathcal{N}(0, R(s_t))$ are transmission rate and state noise, respectively and are state-dependent. $I \cap P = \Theta$ is reflected by the governing parameter set $\Theta = \{A, B, R, r^{min}, \mathcal{Q}\}$,$s_t \in S$, is the system space reflecting the observed vehicular environment and at time slot t can be defined as:

$$s_t = \{v, r_t, T_t, d_t, \eta_t\}, \tag{4.11}$$

where: $v = [v_1, \dots, v_T]$ is a vector possible user velocity. $r_t = [r_1, \dots, r_t]$ denotes the vector of possible data rates satisfying x_{t+1} projections. $\boldsymbol{T_t} = [\boldsymbol{t_1^m}, \dots, \boldsymbol{t_N^m}]$ is the vector average service time for

x_{t+1}. $d = [d_t^m, \dots, d_T^m]$ is a vector of transmission distances. $\eta = [\eta_1, \dots, \eta_N]$ denotes a possible user direction in n^{th} with respect to serving gNB.

For a transition from s_t to s_{t+1}, the corresponding projected received power transition is denoted , $x_t|x_{t+1}$, the immediate reward, $r^{min}(s_t)$, for the observed signal,$y_t \in Y$, is defined as a function of energy cost efficiency: $r^{min}(s_t)$.The transition probability of a gNB link getting blocked or unblocked between states given transitions from s_t to s_{t+1} with corresponding receivable power transition x_{t+1} to x_t is given by:

$$P(S) \triangleq P\big(s_{t+1} = m_j | s_t = m_i\big), \tag{4.12}$$

4.3.1 Modelling Deterioration Pattern of mmWave links

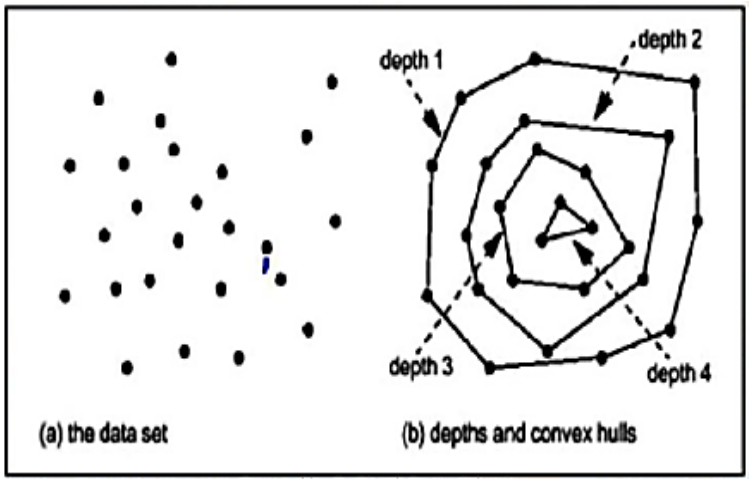

Figure 4.2: The user distances points with similar received power denoted by depth levels.

We assume that after receiving power x_t, at any depth in the user receives an immediate reward $r^{min}(s_t)$:

$$r^{min}(s_t) = \frac{r^{m}(s_t, a_t)}{P_t}, \tag{4.13}$$

However, the objective is to maximize the long-term reward. Since the objective of problem is to maximize gNB utility. If Y_T,X_T,A_T and S_T denote sequences of observable signals over $\{y_1, \dots, y_T\}$ corresponding to received power pattern $\{x_0, \dots\ x_T\}$, and SINRs ratios $\{a_1, \dots, a_T\}$ over $\{s_1, \dots, s_T\}$ states respectively within time space, T. The JMLS model learning scheme determines the maximum parameter sets Θ and possible receivable power, pattern X_T, satisfying SINR value, A_T,

over a finite set of transmission distances $\{d\} \in S_T$ or depths (see Figure 4.2) using the maximum likelihood function $P(X_T, S_T, Y_T|\Theta)$. This formulates a set of convex hull vertices. The JMLS variables for each depth are initially estimated according to maximum likelihood decision rule using Expectation-Maximization (EM) algorithm, [11]. EM uses Bayesian inference to automatically infer optimal missing maximum values of pattern X_T over S_T:

$$\arg\max_{\Theta} P(X_T, S_T, Y_T|\Theta), \tag{4.14}$$

EM estimates reflects the maximum or upper bound values of X_T expected to be observed to achieve maximum SINR value $A \in \mathcal{Q}$, having observed distorted CSI, Y_T, at each depth. EM use the complete-data log-likelihood function for unknown data to estimate unknown elements in X_T [14]:

$$\mathcal{Q}(\Theta|\Theta^k) = \mathbb{E}[\log P(X_T, S_T, Y_T|\Theta)|\, Y_T, \Theta^k], \tag{4.15}$$

where $\Theta^{(k)}$ is the current parameter set estimate at iteration k for a mmWave BS. It is followed by iterative steps, to obtain maximum values of X_T with optimal corresponding SINR ratio A_T over S.

$$\Theta^k = \arg\max_{x \in \Theta} \mathcal{Q}(\Theta|\Theta^k)|_{a(x)}, \tag{4.16}$$

For a given user-gNB distance the receivable power conditional distribution given the path of the user could be simply evaluated via a Rauch-Tunng-Striebel (RTS) smoother [12]. However, because there are so many potential received power values per depth as shown in Figure 4.2, EM estimates alone are not fully tractable in this case. To find the single most likely deterioration pattern given the limited CSI and channel observation, we adapt the Viterbi algorithm [13]. The Vertibi algorithm specifically looks for the pattern having the "best" partial cost at each distance step *d*, *depth* in Figure 4.1 corresponding to each state, *s* of distance points as the user moves away from the cell. We first note validity of distance-based estimation compared to phase and fading before detailing the Viterbi technique.

4.3.2 Distance Induced Approximations

Compared to the phases and fading coefficients of the channels, the rate of change of the users' distance and speed information is moderate [23],[34]. This means that it is more realistic to access

accurate distance and speed data only [145] and project the scanty CSI. Since the users' distances in Figure 4.2 are known, the base station will order the users according to the following criterion:

$$d_1, \ldots., d_k, \tag{4.17}$$

Then CDF of d_k can then be calculated from the probability of Expected, $E_i(X_T, S_T, Y_T|\Theta)$ values in EM within area, $\mathcal{A}(x)$ will receive power x, at a cost of, $\hat{c}i(s)$ i.e.,

$$P_{d_k}(X_T, S_T, Y_T|\Theta^k) = 1 - \sum_{i}^{k-1} P(E_i),$$

$$= 1 - \sum_{i}^{k-1} P_\psi \ , \tag{4.18}$$

where A(d) denotes the area with radius of d from BS, and E_i denotes the event that there are i users in $A(r)$ that will have SINR, ratio,$a \in A$ as per JMLS equations. Following Figure 4.3, we can get the possible distance CSI

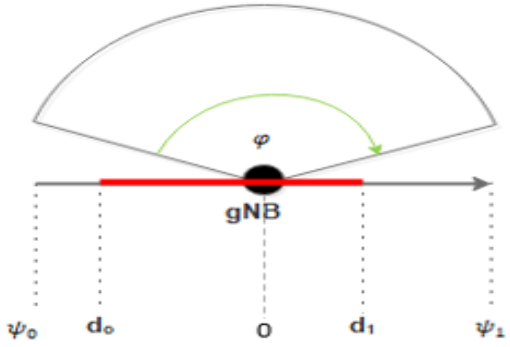

Figure 4.3. UE's maximum dwell ($\psi_o \leftrightarrow \psi_1$) and HO estimation points.

The average serving time, T, per gNB to a user in mode m is illustrated in Figure 4.3 and is defined as (4.23) – (4.25) [18].

$$T = \frac{P_\psi}{\lambda_v(\varphi)}, \tag{4.19}$$

$$\lambda_v(\varphi) = \frac{45}{128\pi E(1/v)} \int_{d_0}^{d_1} dt(1 - \psi_0^2 \ - d^2) \tag{4.20}$$

$$= x \int_o^\pi \sin\varphi \sqrt{1 - (\psi_0 \cos\varphi - d \sin\varphi)} \ d\varphi$$

$$P_{\psi} = 1 - \ P_{\pi}(Y_t|\Theta) \tag{4.21}$$

where P_{ψ} is the probability of target link availability. ψ_1 denotes user distance from serving gNB to where SINR equal the threshold, -5dB. d_0 and d_1 denote possible maximum dew time/distance points based on user velocity. $\lambda_v(\varphi)$ is the total HO rate. The switching time between beams sectors of the same gNB is in scales of nanoseconds, thus neglected [4]. $E(1/d)$ is the average deduced from the distance distribution in Figure 4.2 to give user speed distribution.

4.3.3 Viterbi Approximations

We use the Viterbi algorithm principle in [126],[145] to find the most likely deterioration pattern across different depth in Figure 4.17. For the Viterbi algorithm given many possible different deterioration values by EM for each depth in Figure 4.2. It finds the most likely possible recognisable receivable power cell pattern for a user moving away from the gNB. The values of the gNBs' patchy coverage estimated by EM, can be equated to points in the convex hull's vertices as shown in Figure 2. At each equidistant depth from gNB, the optimal decision rule is to find the least possible energy cost of using the target link up to that depth. In the sense of minimal probability of error is to minimise the EM maximum likelihood decision rule error. Thus, given observable channel, y^t, the estimated output in state, s^t, specifically at each depth denoted by s_d, is given by (4.22):

$$s^d(y^d) \triangleq \mathbb{E}[\log P(X_T, S_T, Y_T|\Theta)|\, Y_T, \Theta^k], \tag{4.22}$$

By defining the cost of transmission in state s^t, specifically at distance d :

$$c^t(s) \triangleq -\log P(x_t, s_t, |\Theta)|\, Y_T, \Theta^k, \quad s \in S^d \tag{4.23}$$

it follows from (4.23) that the log-likelihood function for total energy cost for all depth $d \in S^d$ can be written:

$$\log P(X_T, S_T, Y_T|\Theta) \triangleq \sum_{d=0}^{d} c^d(s), \tag{4.24}$$

and the prediction optimization problem for pattern $X_T,\ d \in S^d$ is:

$$\log P(X_T, S_T, Y_T|\Theta) \triangleq \underset{s\in S^d}{\operatorname{argmin}} \sum_{i=0}^{d} c^d\left(s_{i-d+1}^{i}\right). \tag{4.25}$$

Simply keeping all the best cost at each depth from each EM's previous iterations in memory will yield a policy that generates the average distribution over previous iterations. The author in [88] and [89] for example, successfully employed this strategy. However, it does not scale well in mobile networks because each BS could have a large number of alternative deterioration patterns with respect to each depth. This implies it will require a memory that scales linearly with the number of iterations. In fact, we anticipate that Viterbi will require more and more iterations to converge (as noted in [89]), necessitating the need to monitor an increasing number of cost policies, as the mmWave propagation environment complexity increases. The algorithm produces estimates sequentially on run-time. The normal Viterbi Algorithm [23] in fact works on the basis of accurate training data. This implies that the conditional PDF of the channel, i.e., the full CSI, must be explicitly known for the Viterbi to accurately predict the best possible deterioration pattern across different depth. As discussed in the introduction and under literature review, knowing the full CSI is extremely challenging due to rapidly changing mmWave channels. With EM inference methods, this means, the normal Viterbi algorithm will need a large training overhead. In the following section we adopt a ML-based Viterbi algorithm. It works with limited CSI to make viable decisions.

4.3.4 Machine Learning Based Viterbi Algorithm

Going by the principles of Deep Learning, assigning a number of neural networks learning the best cost and corresponding receivable power at each depth in Figure 4.2 could be too complex as the scale of the network and gNBs to be analysed increases. The complexity however could be avoided by formulating averaging rewards/cost policies at each depth. Formulating averaging policies minimizes memory cost at each state while maintaining a buffer state of best prediction from EM responses. This strategy could be equated to Self-Play method in [67], originally created for imperfect information learning with a limited training data. The author in [68] provides a summary of the self-play algorithm. Our Scheme technically learns an average policy instead of learning the policy indirectly through a collection of neural networks for previous EM approximation. it avoids looping over the entire data set in Figure 4.2. to find the optimal pattern. Thus we use a

representation of the policy created by a neural network to handle a potentially complex mmWave environment.

The method enables us to approximate cost with a single neural network unlike using several like for DRL. It further allows us to reuse EM data after convex vertices formation. Particularly, after re observing distorted data, the neural averaging policy is reused in a simple manner again after EM training for pattern prediction update till a certain level of prediction accuracy is lost. The method has the advantage over those that keep accumulating new training data. First, it is unnecessary to maintain a buffer whose size increases linearly with the number of EM iterations. The buffer is not required after formulating convex hull vertices for each depth. The data set in Figure 4.2 that does not form part of the convex hull vertices/points, particularly, is discarded. This makes our method to be empirically quicker and use less memory. This is done by simply reusing the EM training data until a certain level of accuracy is never attainable at HO.

Expected user receivable power hence corresponding SINR is defined in (4.26) [1]. Since viable the averaging policy is defined according to SINR can only be approximated at step d , we add the policy $\pi(a_k, x_k|Y_K)$, expression as the long-Term reward $r^{min}(s_n, a_n, x_n)$ at each step, d . The policy $\pi(a_k, x_k|Y_K)$ approximates the SINR [15] at each depth, with respect to k, vertices with distance d being the furthest distance for the depth from the gNB for user i, moving at velocity v:

$$r^{min}(s_n, a_n, x_n) = \frac{P|h^H \boldsymbol{p}|^2}{\left(\sum_{i \neq k} P|h_d^H \boldsymbol{p}|^2 + d_i^{\propto}\right)} F_x\left(\left|\theta_k^l\right|\right)$$

$$\pi(a_k, x_k|Y_K) = \sum_{k=1}^{K} \frac{P|h^H \boldsymbol{p}|^2}{(\sum_{i \neq k}^{K} P|h_d^H \boldsymbol{p}|^2 + d_i^{\propto})} F_x\left(\left|\theta_k^l\right|\right), \quad (4.26)$$

where $F_x\left(\left|\theta_k^l\right|\right)$ is influenced by user behaviour at time step $t < T$ as noted in (4.3).

Given policy $\pi(a_k, x_k|Y_K)$, the major challenge of computing the best cost, ci(s) from the observed vertices is simplified. A parametric estimate of $P_{S_{i-d+1}|y_i}(\mathrm{x}, \mathrm{s}|y_{i,}\Theta)$, denoted $\hat{P}(\mathrm{x}, \mathrm{s}|y_{i,}\Theta^i)$, is reliably obtained from training data using standard classification DNNs with a SoftMax output layer. The ML-based log-likelihood computation is depicted in Algorithm 4.1.

Algorithm 4.1.

1: Initialize a reservoir buffer *EM* for *JMLS* learning of average policy

2: Initialize parameter estimates $\hat{\theta}_o^s$ from θ_o^s in EM

3: **for** *EM-step* $k = 1, \ldots, K$ **do**

4: 1. NLOS distribution: Generate u_k with $\hat{\pi}_{\theta^{k-1}}$where $u_k = -Kx_k$ in (4.10)

2. Best EM: Train $\hat{\pi}_{\theta^k}$ estimates against u_k, with DNN.

6: Collect samples state-action using $\hat{\pi}_{\theta^k}$ and recalibrate EM estimates

7: Average policy: Update $\hat{\pi}_{\theta^k}$ by adjusting $\hat{\theta}_k^s$ (through gradient descent) to minimize cost:

$$\underset{s \in S^d}{\operatorname{argmin}} \sum_{i=0}^{d} c^d\left(s_{i-d+1}^i\right) = \mathbb{E}_{(s,a)_{\sim \mathrm{EM}}}\left[-\log(X_T, S_T, Y_T|\Theta)\right.$$

where $\hat{\pi}_{\theta^k}$ is the neural net policy with parameters $\hat{\theta}^s$, for

$$\Theta^k = \arg\max_{\mathrm{x} \in \Theta} Q(\Theta|\Theta^k)|_{a(x)}, \text{ in (4.15)}$$

8: **end for**

9: **Return** u_k, $\hat{\pi}_{\theta^k}$ (A, X, S)

4.3.5 ViterbiNet Algorithm

We adopt ViterbiNet to implement the above approximations, wherein the conditional distribution estimate $\hat{P}(\mathrm{x}, \mathrm{s}|y_{i,}\Theta^i)$,encapsulates the Markovian behaviours and is utilised to calculate the desired log-likelihoods. The ML-based ViterbiNet algorithm replaces the CSI-based portions of the Viterbi algorithm. The aspect of distance-based channel estimation policies is relatively easy to comprehend, allowing ViterbiNet to utilise a simple and standard DNN architecture. Due to the network's straightforward DNN structure, it is quickly trained with a small number of training samples. This suggests that the proposed architecture has the potential to adapt in real-time to channel variations with minimal overhead, possibly through the use of periodically embedded pilot sequences in the transmitted frame for online training. The strategy utilised in the design of ViterbiNet, retaining the prediction scheme while replacing its channel-model-based computations with dedicated ML methods, is applicable to the data-driven implementation communications in Algorithm 4.1. The Figure 4.4 summarises how the ViterbiNet has been modified to work. Here, the EM estimates based on the total cost are recalculated by DNN which then updates EM via (4.15) to improve its accuracy as per algorithm 4.1.

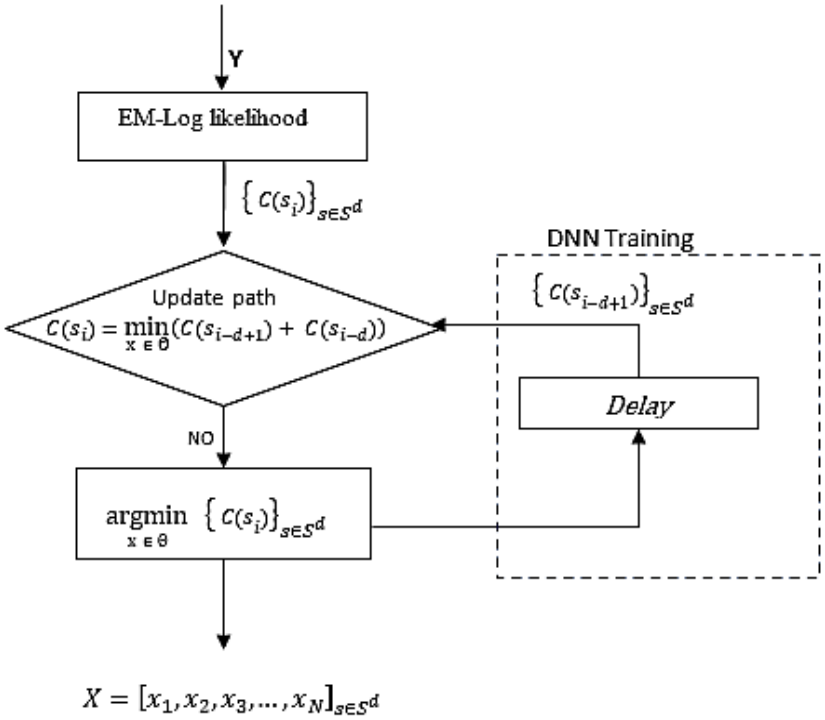

Figure 4.4. The Adopted DNN-based Viterbi pattern predictor [56][134].

4.3.6 ViterbiNet Optimization

This ViterbiNet method requires that a buffer be kept during retraining, whose size increases linearly as the number of iterations increases. This gives us reason to investigate a variant that is not only empirically faster but also consumes less memory. Now let's talk about EM and DNN together. DRL could easily be applied to the policy evaluation stage by using an equation to estimate the Q-function (4.15). On the other hand, developing a neural network that can closely resemble the cumulative Q-function is challenging. For conventional RL, we suggest a reparameterization with a single agent and less interactions with the environment. Relies on already obtained EM training data. In this situation DNN can calculate the EM-regularized Q function Q, but this requires the addition of Q-functions (k). Recalculating the EM values (4.15), however, could result in greater imprecision because there is no easy way to sum neural networks. We adopt a novel method in [2021]. Unlike one that requires a correct Q-function and two approximations method, we use a singular approximation approach to transform the EM approximation into DNN approximated.

For every sample step *d*, we replace the EM function at iteration *k*, namely $\mathcal{Q}(\Theta|\Theta^k)$ with a neural network whose parameters at *k* are retrained to minimise a cost function and improve the averaging policy value over K vertices. We assume that (s, x, a) represents the state because we want to readjust a, as a function of distance, d vertices in Figure .3 given policy $\pi(a_k, x_k|Y_K)$. We start by describing the following modified Bellman equation [45]:

$$\begin{cases} \tilde{Q}_{N+1}^{k}(x,a,s) = 0 & \text{for } N > K \\ \tilde{Q}_{N+1}^{k}(x,a,s) = r^{min}(s_n,a_n,x_n) + \tau \ln \hat{\pi}_{\theta^{n-1}}^{k}(x_{n-1},a_{n-1}|Y_n) + \\ \qquad \mathbb{E}_{(x',a')_{\sim EM}}\left[\tilde{Q}_{N+1}^{k}(x',a',s) - +\tau \ln \hat{\pi}_{\theta^n}^{k}(x_n,a_n|Y_n)\right] \end{cases} \tag{4.27}$$

here $x' \sim p_n(\cdot\,|x,a,u_{n-1}^{k})$ and $a' \sim \hat{\pi}_{\theta^n}^{k}(.\,|Y_n) a\,0 \sim \pi\, k\, n\,(\cdot\,|x\,0)$. The term $\tau \ln \hat{\pi}_{\theta^{n-1}}^{k}(x_{n-1},a_{n-1}|Y_n)$ penalizes the distance based reward policy for negatively deviating from the best cost in the previous iteration, $\hat{\pi}_{\theta^{n-1}}^{k}$, while the term $\tau \ln \hat{\pi}_{\theta^n}^{k}(x_n,a_n|Y_n)$ compensates for the upward adjustment of the averaged reward if the new estimate among the vertices for the convex hull offers better received power. Thus, over the long term, the target link that has a history of high stability with consistently high rewarded is preferred. Moreover, the algorithm update is with respect to Algorithm 4.1 is as follows: after initializing policy $\hat{\pi}_{\theta^k}$, repeat for $k \geq 0$ for all vertices:

$$\begin{cases} \text{NLOS update: } u_k = u_{k+1} \\ \quad \text{DNN Q} - \text{function update: } \tilde{Q}_N^{k+1} \quad as\ in\ (4.27), \\ \text{Reward Policy update: } \hat{\pi}_{\theta^n}^{k+1} = \text{softmax}\left(\frac{1}{\tau}\tilde{Q}_N^{k+1}\right) \end{cases} \tag{4.28}$$

Consequently, in spite of the innovative artificial log terms, this method does not bias the ViterbiNet algorithm convergence. It enjoys the same convergence guarantees but can make likelihood decisions faster based on penalties as explained in the following section.

4.3.7 Received Power Deterioration Estimation

At HO execution, each predicted received power pattern X_i, given discontinuous power updates on Convex Hull vertices in Figure 4.3 and outputs in (4.28) represents how a target link channel's received power on user i, is likely to behave over user-gNB distance d_i after HO. As earlier stated, to predict X_i over d_i, the best cost/reward corresponding receivable power $x_i \in$S, is deduced from each set of depth vertices X_i, making the Convex Hull observed by EM over different states,S_T. As earlier stated, the purpose of training is to predict the reliability of the mmWave link over a

reasonable transmission distance, d after HO. The extent of the prediction errors is however minimised by comparing how the individual target link expected deterioration pattern deviates from the averaging of all optimal values at each distance i.e., the global deterioration pattern compares against local expected deterioration pattern of target cells The degree of deviation determines the reliability of the local expected deterioration pattern of the target cell. In the next subsection, we show how online training is executed.

4.3.8 Online Training of the Viterbi Algorithm

We will then go over how the algorithm, can be used to train and monitor changes in online channel behaviour. Our suggested approach is meta-learning [24], a field of study that focuses on self-teaching ML algorithms. Using the observed channels, denoted by Y_T, across a number of states, denoted by the depth states in Figure 4.2, $S_T(j)$, where j ∈ N represents the depth index, the scheme calculates the receivable power, denoted by X_T. If the number of prediction errors does not go above the minimum penalties/cost in (4.27), the link is assumed that it will flawlessly be sustained in the presence of NLOS and sparse coverage [42]. This also implies the transmission data rate can be maintained with very minimal deviation from the prediction. The JMLS prediction pattern is deemed accurate and reliable if no HO request is initiated within the predicted coverage distance. To that effect, the need for additional training data is suspended. With only a few training samples, a ViterbiNet's simple DNN architecture enables effective and quick prediction. Further with the reward and penalty system in (4.27) when one out ways, the reliability of the target link can be assumed even before convergence. This shortens the HO decision process and suitable for online training system as shown in Figure 4.5. The Figure 4.5 essentially demonstrates the sequence estimates of EM function have observed part of the channel information under DNN, the correction error is estimated based on prediction failures, and the training data is reused based on the cost outcome/penalties in (4.27).

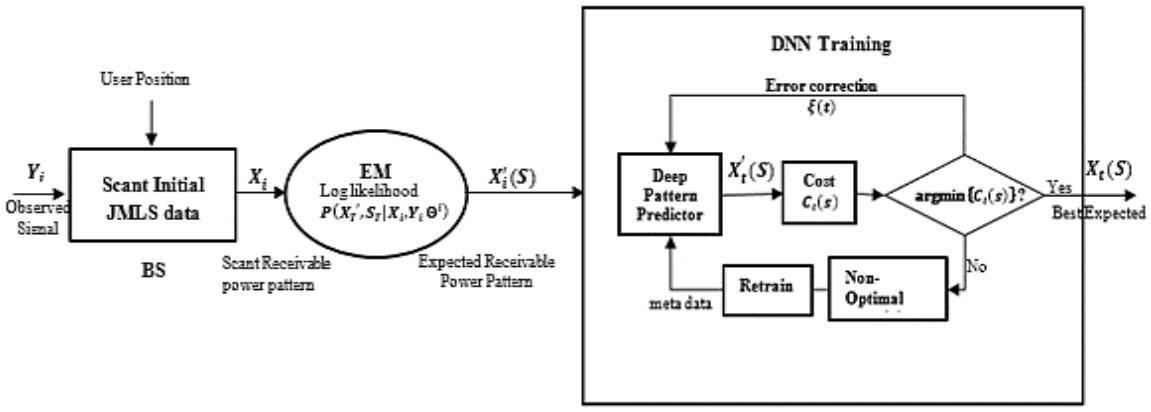

Figure 4.5: The Viterbi pattern predictor online Training.

An inherent disadvantage of decision-directed approaches is their susceptibility to decision errors. The meta-training samples pair $\hat{X}_T(j)$ $\hat{A}_T(j)$ and $\hat{Y}_T(j)$ used to make the HO decision, for example, do not accurately represent the channel behaviour resulting in $\hat{Y}_T$ if the JMLS-based HO fails to accurately predict the deterioration pattern $\hat{X}_T(j)$.This makes the suggested approach unreliable, especially in low signal-to-noise ratio (SNR) scenarios like non-line-of-sight (NLOS) scenarios, brokaged or reflected scenarios. Ultimately, this can result into unwanted HO. In such cases, the inaccurate training pattern may gradually degrade the HO prediction accuracy. The impact of poor decision-making can be lessened if allowable HO failure numbers are set in conjunction with the number of times a mmWave link tries to join LTE. In particular, to simplify retraining, the HO scheme can decide to reuse EM estimates as meta-training data and retrain the deterioration pattern predictions only when the number of HO from mmWave to LTE or mmWave HO exceeds the set threshold ξ_d(j). By deciding to reuse only accurate meta-training data, this approach allows for the management and mitigation of the effects of HO decision errors. The following Algorithm 4.2 provides a summary of the suggested online training system. Based on the anticipated nature of differences between HOs, the value of ξ_d(j), should be chosen. ξ_d(j), can be set to 0 to represent zero LTE HO when associating to a target gNB with gradual variations, thus guaranteeing the precision of the meta-training for retraining the HO ViterbiNet scheme. This setting may however limit the ability to track and predict moderate channel variations in partial NLOS or usage of reflected Signals, leading to needless HOs. If the scheme uses larger values of ξ_d(j), it may

encourage the use of LTE links. The proposed method can thus be used to regularise the extent to which LTE links can be used.

Algorithm 4.2: Online Training.

Input: User mobility model parameters, $\mathbb{P}$, v, DNN model weights φ, Parameters about DC communication: transmission power limits, bandwidth, channel gain, and NLOS and LOS path loss exponent., error threshold $\bar{\xi}$,Observed state of different user environment S; Set of observed signals $Y = [y_1, y_2, y_3, \dots, y_N] \in \mathbb{R}$,

Output: An optimized mmWave Deterioration path $X = [x_1, x_2, x_3, \dots, x_N]$ for target link

1. Initialize the deterioration path estimations with K randomly available training signal Y and states S
2. **for** $t = 1$ **do**
3. Draw y_t for JMLS parameter estimation Θ, where $(X_T, S_T, Y_T|\Theta)$
4. **Estimated Maximization (EM):**
5. $\underset{\Theta}{\operatorname{argmax}}\, P(X_T, S_T, Y_T|\Theta)$,
6. **Apply ViterbiNet with $\boldsymbol{\Theta}, \boldsymbol{\varphi}$ and $\boldsymbol{X_i}, \boldsymbol{Y_i}$ $\boldsymbol{\Theta^i}$**
7. Estimate $\boldsymbol{X_T}'$ over $\boldsymbol{S_T}$ states or distances, calculate number of error if given $\boldsymbol{Y_i}$ and $\bar{\xi}$
8. **If** $\xi < \bar{\xi}$ **then**
9. P$redict$ $\boldsymbol{X_T}'$ over new $\boldsymbol{S_T}$ for target links.
10. Calculate the cost of transmission, $\boldsymbol{C_t(s)}$ given $\boldsymbol{X_T}'$ over distance**,** $\boldsymbol{d \in S}$
11. **If** argmin $\{C_t(s)\}$ **then**
12. Retrain DNN weights $\boldsymbol{\varphi}$ using meta training $\boldsymbol{X_{T-1}}'$, $\boldsymbol{Y_i}$
13. Initialize replay memory using EM samples over $\boldsymbol{Y_i}$
14. Update mixture model Y_i and φ
15. **end if**
16. **end if**
17. **end for**
18. Output **X** for HO prediction over $d \in \boldsymbol{S}$

4.4 HandOver Decision and Execution

Given the deterioration patterns output, X_t. At HO, the scheme looks for potential target links whose SINR is above the threshold for the next step.

A) HandOver Decision

To find a single "best" pattern matching the expected pattern output in Algorithm 4.2 is challenging. To this end, we consider at each distance *d*, an allowable deviation of:

$$\mu_i(d) \triangleq \frac{P\left(X_{d,i,}\right),}{\sum_i^{k-1} P\left(X_{t,i,}\right),}$$

$$\triangleq 1 \Big/ 1 + exp\left(\sum_{i \neq j} c^d\left(s_{j-d+1}^x\right) + c^d(s_i^x)\right), \quad (4.29)$$

where $\mu_i(t) = P\left(X_{t,}\right)$

Under each user scenario, the overall reliability of a link and its variance are considered as the mixture of these mode-dependent estimates and are given by:

$$\hat{x}_i = \sum_{i=1}^{M} \mu_i(t)\hat{x}_{t,i}$$

$$\widehat{\sum}_i = \sum_{i=1}^{M} \mu_i(t)\widehat{\sum}_{t,i} + \sum_{i=1}^{M} \mu_i(t)\left(\hat{x}_{t,i} - \hat{x}_t\right)\left(\hat{x}_{t,i} - \hat{x}_t\right)^T, \quad (4.30)$$

where $\hat{x}_{t,i}$ and, $\widehat{\sum}_{t,i}$ are the mean value and the variance of x_t under the user type *i*

B) HandOver Execution

Given the current status of gNB, the SINR gradient can be estimated from pattern of received power, X_T, over S_T,states towards a predefined critical level i.e., near to the receiver sensitivity value. The time to trigger (TTT) is then defined as the time left for *pattern* X_T*, on BS to* reach for the first time this SINR failure threshold.

To assess how closely each target link's expected deterioration pattern matches the ideal X_T^* in C_T^*, we perform the Kaiser-Meyer-Olkin (KMO) test [17]. The KMO test compares the correlation between the overall/globally Viterbi-generated sequence, X_T^*values, and a local pattern expected on the particular gNB. Then, using the following formula (4.31) to estimate the overall KMO index value for the user mode in relation to the target cell:

$$KMO_{\hat{x}} = \frac{\sum_{x \neq \hat{x}} R_{x\hat{x}}^2}{\sum_{x \neq \hat{x}} R_{\widehat{x\hat{x}}}^2 + \sum_{x \neq \hat{x}} a_{x\hat{x}}^2}, \tag{4.31}$$

where $R = [r_{xd}]$ is the correlation matrix, $A = [a_{xd}]$ is the partial covariance matrix where a_{xd} is defined as

$$a_{x \neq \hat{x}.m} = \frac{r_{x\hat{x}} - r_{x.m} r_{\hat{x}.m}}{(1 - r_{xm}^2)\left(1 - r_{\hat{x}m}^2\right)}, \tag{4.32}$$

and,

$$r_{xd} = \frac{\sum_{t=0}^{T} (x_t - \hat{x}_t)\left(d_t - \hat{d}_t\right)}{\sqrt{\sum_{t=0}^{T} (x_t - \hat{x}_t)^2 \sum_{t=0}^{T} \left(d_t - \hat{d}_t\right)^2}}, \tag{4.33}$$

where $X_T^* = \{\hat{x}_0, \ldots, \hat{x}_d\}$ denotes the global optimal received power pattern over user-gNB distances $\{\hat{d}_0, \ldots, \hat{d}_d\}$ given the individual target cell's anticipated local received power value set $\{x_0, \ldots, x_d\}$ over corresponding average optimal distance set $\{d_0, \ldots, d_d\}$ $\hat{d}_t$ KMO test takes values between 0 and 1. The general rule for interpreting the measures is given in Table 4.1.

Table 4.1: Interpretation Of KMO Measure [23].

KMO	Interpretaion
0.9 and above	**Marvelous**
0.8 – 0.9	**Meritorious**
0.7 - 0.8	**Middling**
0.6 – 0.7	**Mediocre**
0.5 -0.6	**Miserable**
Under 0.5	**Unacceptable**

In this study, we select the target cells with KMO index above 0.751. If the KMO index value is less than 0.7, most likely the target link is not suitable for consideration despite high initial SINR. Therefore, given the KMO index list, the best cell (target) is selected according to highest SINR among highly KMO indexed cells with test values above 0.75. For instance, where the current serving cell SINR values and KMO target cell values, at HO phase much that of target the user will remain connected to the serving gNB and go back to KMO prediction phase. Otherwise, the scheme executes the HO process and then go back to prediction phase. It must be noted that KMO values less than 0.6 indicate high link failure chances hence not adequate to sustain connectivity over distance.

4.5 Performance Evaluation

Using system-level simulations, we looked into the performance. For analysis, data with a resolution of one Transmission Time Interval (TTI) (1 ms) were gathered from simulations lasting over 1000s of seconds. Table 4.2 provides a summary of the key parameters used. See [9] for a more thorough analysis of simulators. For the purposes of comparison, a soft-HO scheme [7] that creates HOs using only SINR/data rate served as a baseline.

TABLE 4.2:Simulation Parameters[7,9,53].

Parameter	**Value**
mmWave	28GHz
LTE Carrier frequency	2.1GHz
mmWave bandwidth	1GHz
3GPP Channel Scenario	Urban Micro, Urban Macro
MMWave max outage	-5dB
mmWave transmission Power	40dBm
mmWave max PHY Rate	3.2GHz
X2 link latency	1ms
S1 link latency	10ms
RLC buffer Size	5MB
S1 MME link latency	10ms
User speed	[1,50] m/s
UDP Source rate	200Mbits/sec
UDP packet Size	1028bytes
Default No. of BSs	80

4.5.1 Simulation Results Analysis

According to the estimated long-term rewards, of r_i^{min} , in (4.10), Figure 4.6 shows the convergence of the "best" partial cost over 240 iterations for each user type. Here, results show that expected deterioration sequences are generated more quickly because users moving at lower velocities have a faster converging rate. Figure 4.7 summarises how closely outage probabilities predicted by the Viterbi values match actual outage values. The indicates that the training is reliable and can be used to forecast the expected deterioration path prior to HO. Particularly because the estimated values before and after HO are typically similar.

Figure 4.6. shows the convergence of the “best” partial cost over 240 iterations the based-on reward estimates, $r_i^{min} \in \mathbb{R}^K$, for each user type. Here results show that the converging rate for users moving at lower velocities are faster hence expected deterioration sequences are generated faster. Figure 4.7 is a summaries how close outage probabilities estimated by the Viterbi values fair with actual outage values. In most instances, the values are similar hence the training is reliable and can be relied upon to estimate the expected deterioration path prior to HO.

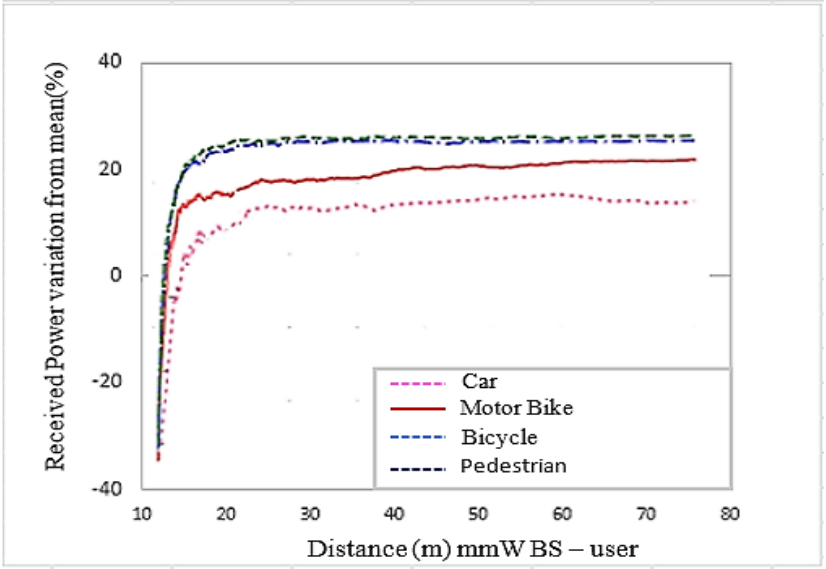

Figure 4.6: Convergence of the Viterbi-based EM algorithm.

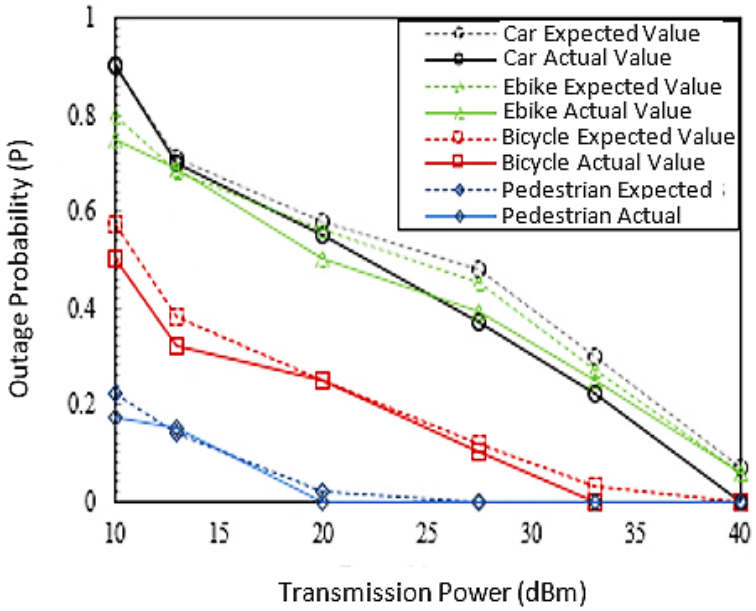

Figure 4.7: mmWave Outage probability to power pattern variation.

Figure 4.8 depicts how closely the predicted throughput and actual throughput correspond for pedestrians connecting to their nearest gNB. The analysis in Figure 4.7 regarding the precision of our JMLS prediction model is further supported by the similarity between the estimated and actual data rate in Figure 4.8 Contrary to this viewpoint, we can confidently rely on projected link deterioration patterns when making HO decisions. Using the accuracy of the deterioration pattern

or sequence at HO and far after HO, it is possible to estimate the link's dependability. For example, in Figure 4.8, future HOs and data rates will be more accurate if we can predict a target link with less abrupt data rate changes than the one used in the analysis. Our proposed model has been shown to process fewer HOs than schemes that rely solely on SINR variation [7,9]. Due to its accurate prediction of the anticipated deterioration pattern, our proposed HO scheme effectively selects links that may allow for a longer dwell time between HOs.

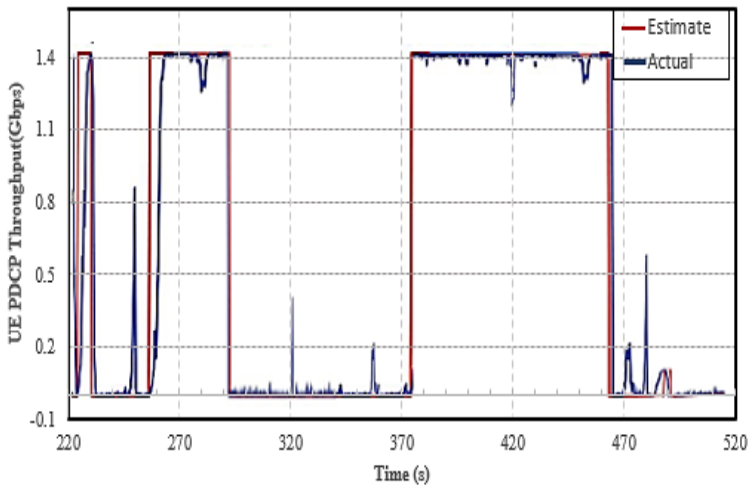

Figure 4.8: Prediction of expected data rate for a pedestrian transmitting via gNBs.

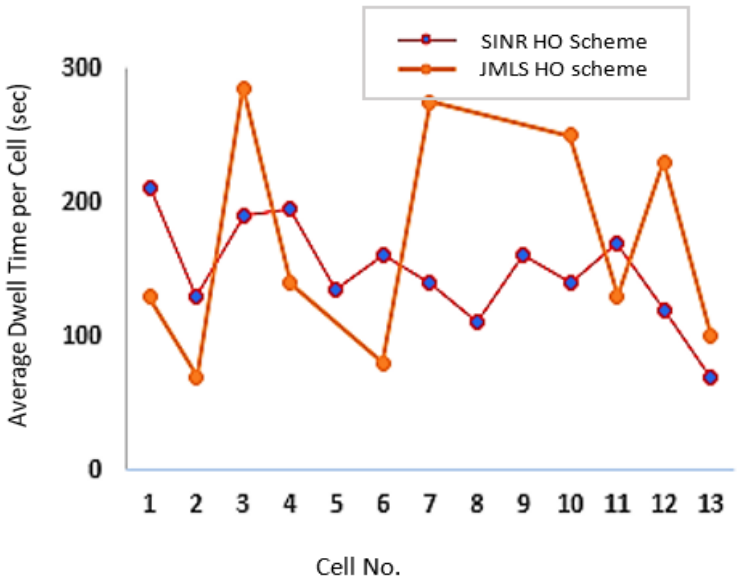

Figure 4.9: Dwell time of pedestrian on gNBs and number of HOs.

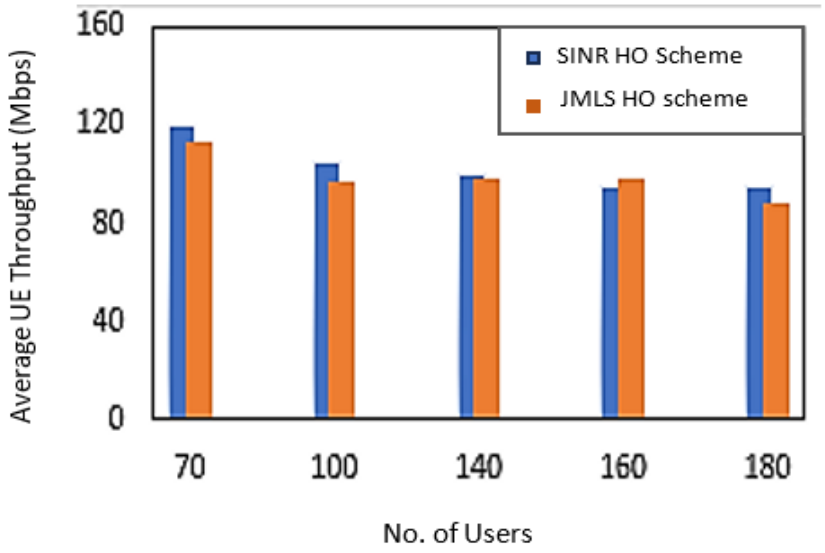

Figure 4.10: Comparison of average data rate for UE using the proposed model and the SINR/data rate Based model in [7,9].

Figure 4.9 shows how our plan can decrease HO that is not necessary because it lengthens average dwell time. Our suggested scheme generated ten HOs as opposed to the thirteen HOs produced by a scheme that relied only on SINR variation. It follows that our HO scheme lengthens dwell time relative to the traditional HO schemes in [7] and [9]. Figure 4.10 further affirms that even though dwell times are typically longer than average, user data rates are typically lower. This is because, as depicted in Figure 4.9, the proposed model has a longer dwell time between succeeding HOs. Riding on the results of Figure 4.9 to Figure 4.10, and considering the short range of mmWave transmission, a poor HO scheme with a large number of BSs would give a higher HO failure rate and poorer performance. Since our proposed HO scheme's longer dwell time entails mmWave BSs needed for HO per unit area is less, the number of mmWave BSs required to sustain connectivity for the entire network is also less. Particularly, considering the short range of mmWave Transmission, covering a given area would require a large amount of BS to avoid link failure for mobile users. Coupled with a poor HO algorithm, this would require multiple HOs hence a large system overhead and, large BS density to sustain continuous connectivity. A smart HO algorithm, however, reduces the area density requirement of mmWave BS. In addition, this also reduces the installation and operational cost of mmWave BSs in 5G mobile communication.

4.6 Conclusion

In order to address the issue of multiple co-existing and competing mmWave link deterioration factors in HO decision-making, this paper suggests using the JMLS. The JMLS model uses an additional discrete switching variable to take into account abrupt switches between NLoS and LoS scenarios due to the sensitivity, dynamics, and variability of propagation conditions, such as the co-existence of NLOS and LOS link deterioration. In order to better understand how users transition from one mode of link deterioration to another, the mmWave link behaviour is examined. This in retrospect offers a link deterioration analysis that is more accurate when applied to actual phenomena. Prior to making a HO decision, we model the analysis' findings and predict how target links will deteriorate for various user types. Results from the analysis and forecasting of deterioration patterns are very encouraging in terms of link reliability evaluation and estimated time to failure (dwell time). Despite the fact that power/SINR and data rates are thought to change linearly with distance, actual results (see Figure. 4.8) show a non-linear behaviour. Therefore, future research must take non-linear JMS analysis of mmWave link deterioration pattern into account when making HO decisions. This will make it easier to calculate data rate changes or spikes that the figure didn't predict. Furthermore, real data, as opposed to the simulated data used, should be used to validate the proposed model.

Chapter 5

Enhanced Deep Reinforcement Learning HandOver Model

Most HO schemes using machine learning or statistical modelling schemes require huge number of training data set to come up with reliable HO policies. Further even those that have integrated learning capabilities rarely take stock of the competing deterioration factors of mmWaves transmission. One example of such RL scheme in [36] where Mobility load balancing (MLB) and Mobility Robustness Optimization (MRO) HO schemes simultaneous use Reinforcement Learning to solve 5G mobile HO challenges. The scheme however does not consider mmWave deficiencies such as susceptibility to blockage and NLOs effects. It only simultaneously focuses on load balancing and user speed to trigger HOs. For others like [12], using Deep Reinforcement Learning, they require huge training data sets to formulate reliable HO policies. Furthermore, it is assumed that the base station has perfect knowledge of the users' effective channel gains. However, for a fast time varying network, the slow policy formulation in DRL might not be realistic. This is

especially true given how quickly the mmWave channel vector phases and fading coefficients change. Additionally, following directional isolation between links: the need for highly directional and adaptive transmissions, significant outages could result due to user mobility. Significant decisions on when and how to time and trigger HO is needed even amid limited CSI. In light of these challenges, the goal of this chapter is to forecast mmWave network behaviour and connectivity availability in target links using enhanced DRL learning schemes. Particularly to predict the link's reliability and formulate viable HO policies faster using DRL. We introduce an integrated DRL and Jump Markov HO model to reduce training time in DRL. The scheme allows DRL to capitalize on Jump Markov's simplicity and formulate novice policies using few training samples. Then using JMLS model data as meta data, DRL formulates stable policies faster without requiring large training sample. Thus, DRL leverages Jump Markov's ability to predict abrupt behaviours and learns to predict the extent to which abrupt changes may affect overall 5G performance. The scheme capitalises the effects of reflected links by selecting links based on high channel gain to compensate decrease in receivable power for the same value of SINR. From the results, our scheme had a better convergency rate and selected more energy efficient links.

The chapter structure is as follows. Section 5.3 gives a brief background about Deep Learning and RL HOs mechanisms and limitations. Section 5.2 introduces the Network and Mobility Model used. Section 5.3 discusses the DRL HO model and 5.4 presents simulation results obtained along with analysis. Section 5.5 discusses the findings. Finally, a brief conclusion in Section 5.6 ends the chapter.

5.1 Background

The development of 5G mobile network systems has given mmWave communication a lot of attention. This is attributed to the advantages of reducing the spectrum shortage and supporting high data rates, as discussed in chapter 2 [1]. In a cellular network, where HO techniques are implemented to optimize performance and ensure connectivity in a highly directional transmission environment, conventional HO schemes have fallen short. They typically result into high power consumption and prohibitive transmission costs solutions. To balance cost and performance, hybrid HO schemes composed of cascaded classic HOs have been proposed [5]. These schemes use fewer HO numbers than when using individual classic HO schemes. However, they suppose LTE link behaviour even when it is a mmWave HO executions in 5G. To that effect, combining

legacy HO solutions rarely yield desired results from mmWave usage. To counter some of the bias towards using LTE links, a high-dimensional multiuser (MU) 5G HO selection problem has been studied by [11]. Numerous algorithms, including maximum magnitude cell coverage selection and maximum signal-to-interference-ratio (SINR) beam selection, were recommended. For mmWave BS selection, the authors of [13] for instance, created an energy-max algorithm. The level of computational complexity offered by the optimization techniques was however extremely high. Furthermore, when trying to resolve a non-convex problem, it was merely caught in a local optimum. Moreover, the HO selection matrix and the execution matrix were primarily optimized separately by the existing algorithms. The joint design research was crucial to enhancing performance. However, it presented a more difficult task because it involved mixed-integer nonlinear programming (MINLP) problem.

Recent studies have suggested the use of Deep learning particularly, neural networks (NNs) for resource allocation [21], HO selection [18]– [20], and 5G design [16]. The concepts used an iterative algorithm to discover the connection. For instance, the iterative weighted minimum mean-square error (WMMSE) algorithm for selection was approximated by the authors of [17] using a NN with fully connected (FC) layers. The beam selection solution to a multiclass classification problem using support vector machine solution was investigated in [18]. The authors in [19] developed a dataset to test link selection strategies for the vehicle-to-infrastructure scenario. In order to optimise the link selection matrix, the angle-of-arrival (AoA) information was utilised. The design used a multi-layer perceptron (MLP) [20]. The challenge is the majority of these solutions, if not all of them, were offered by black-box NNs, which are by nature limited in their capacity for generalisation and interpretation. They usually call for lots of training data. So-called deep unfolding NNs were created in [22]– [26] to address these problems. These involve layer-wise unfolding of iterative optimization algorithms, just like the NN solution. Deep reinforcement learning (DRL) has also been used to successfully solve high-dimensional non-convex optimization problems [27]. The deep Q-Network (DQN) outperforms the traditional Q-learning when combined with the deep neural network (DNN). Mobile offloading [28], dynamic channel access [29], radio access [30], resource allocation [31], [32], and mobile-edge computing (MEC) [33] are a few examples of applications that heavily rely on the DQN. In [32], a solution that maximises network utility in heterogeneous cellular networks was developed using a multi-

agent DRL strategy. For the MEC network, the authors of [33] created an online offloading framework based on DRL. All these indicate the relevance of integrated solutions.

5.2 Problem Detailing, and Contributions

This section expounds the problem being solved and our specific contributions given the extent of the problem.

A) Problem Detailing

The problem is mmWaves have unpredictable random propagation behaviour. For 5G DC none stand alone, this makes the selection of LTE link more favourable. This random behaviour can cause both abrupt and gradual changes to network performance. Moreover, there are competing mmWave network performance deterioration factors such as foliage, human bodies etc. (as detailed in our literature survey) at play. These can cause drastic changes to mmWave performance in 5G over a very short time. For example, data rate can change from excellent at HO to extremely poor after HO or from LOS to NLOS within a split of time after HO and need unnecessarily more HOs. Considering no user plane data is transmitted during HO, too many HOs waste bandwidth [132] and of course energy e.g., in ping-pong condition. Further, classic HO schemes like the SINR/data-rate based HOs depend highly on the immediate observed data like SINR/data-rate [122], to determine the reliability of the link. Statistical learning HO schemes alike e.g. those using Hidden Markov [123], Fuzzy Logic or the more advanced schemes using JMLS [125] can make quite reliable HO link selection and decisions if the training data is largely repetitive. However, statistical methods are poor if the data for learning shows too much randomness as the one in mmWave transmission [106], i.e., gradual changes punctured with drastic and abrupt changes or vice-versa. In fact, when used in novel spaces, other than where they were not trained from or conditions under which they were not trained for, their performance is poor. For Deep Learning mechanized HO schemes just like most machine learning have proved to solve highly complex HO conditions even for novel spaces or transmission conditions they can readapt and form optimal policies in spite of being computational heavy. The challenge in HO selection performance with them generally is they take too long than classical and statistical HO schemes. And they need large training data set to formulate reliable policies. Further even though those that have both offline and online training module make good policies quicker when deployed, the offline training data set is not always available nor is it accurate to consider all link deterioration

conditions. To top is all, they are costly in terms of energy. Beside for fast changing training data, they can formulate unreliable policies.

B) Contributions

We formulate the problem of joint link selection reliability and energy efficient optimization for mmWave system with JMLS-DRL scheme. We develop an efficient framework of joint NN design. A DRL-based NN and a deep-unfolding NN are jointly used to optimize the HO link selection, reliability and energy efficiency in both LOS and NLOS scenarios. The specific contributions include:

- We integrate JMLS modeling and DRL process to simultaneously benefit from the fast formulation of policies in statistical Markov HO schemes like JMLS and Hidden Markov, and prediction accuracy of DRL. In a mutually exclusive way, DRL relies on JMLS for short term HO policies. DRL reuses JMLS data to train instead of requesting new training data from the environment. JMLS projections also rely on DRL for long term prediction.
- We formulate the mmWave link reliability HO selection problem as a Markova Decision Problem and a JMLS algorithm is developed to solve it. The state, action, and reward function are carefully designed for the BS agent using DRL principles. We make use of the channel link's characteristics to guarantee the mmWave constraints are satisfied and the HO selection mantra/repetition/wastage owing to erratic mmWave behaviour is avoided.
- We proposed a deep-unfolding NN where we use trainable parameters to optimize the reliability of mmWaves. Here, the iterative JMLS algorithm is unfolded into a layer-wise structure to jointly train the two NNs.

5.3 Network and Mobility Model

5.3.1 Network Model

In order to analyse viable cell pattern options for HO, four distinct user type behaviours are learnt in relation to mmWave BSs propagation patterns. Automobiles, pedestrians, cyclists, and e-bikers are the four different user categories instigating different user behaviour patterns. Thus, in addition to SINR values, we suggest using the likely received power and channel gain pattern to determine the optimal mmWave target cell or link. Every type of user introduces new dynamics to the transmitting operation of mmWave channel links that is worth capturing when making HO

decisions. Initially, the JMLS is utilised to model the expected cell pattern using Expected Maximization (EM) estimates. The EM link behaviour estimations are then used to optimise HO using DRL with EM data estimates acting as meta-training data within acceptable error margins. Unlike using first offline model for training where data is collected and trained separately, the evolution of the mmWave network behaviour is reflected almost immediately in the HO. The fluidity of mmWave behaviour makes it hard for offline models to make viable decisions. It is important to use online training so that changes to the policy can immediately be propagated to affect the experience and outcomes observed HO agent. On the other hand, starting with a purely random policy leads to slow learning in DRL. Particularly because the user must spend time to learn the rewards associated with exploration. Thus, in order to make training more efficient, we first perform JMLS-training using scant available data and simple extrapolation by following an ϵ-greedy exploration on the EM policy. After this pre-DRL training, DRL online training using DDPG with the EM experience replay refines the agents' HO policy in terms of energy efficiency. Figure 5.1 shows the use of JMLS modelling-training followed by DRL reinforcement learning.

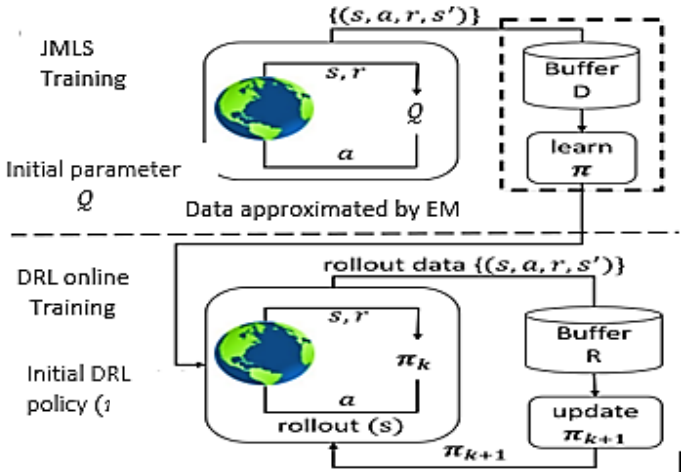

Figure 5.1: Online JMLS training, and DRL training our 5G setting.

5.3.2 Mobility Model

A Manhattan grid model is used to model the road network with streets and intersections (as shown in Figure 5.2 in an urban scenario. There are pedestrians, drivers of motor vehicles, cyclists, and users of electric bicycles. Within the two car lanes that are situated on either side of the road, users of automobiles move. On the other hand, riders of bicycles and electric bicycles travel along the edges of the road, one on each side. Figure 5.2 shows the movements made in different directions.

On either side of the highway pedestrian lane lanes are available, pedestrians are free to walk in either direction. Three thousand metres is the total length of the road. The user automatically changes its speed to match the one in front of it. Each LTE eNodeB supports 200 full buffer users in total. Walking traffic makes up 75% of them, moving at a speed of 1.4m/sec. Bicycles and electric bicycles, move at speeds of 7 and 8m/sec respectively. The remaining one-third of the are vehicles moving between 10 and 14 m/sec. Vehicles that are more than three metres apart will slow down to the speed of the leading car to avoid a collision. Every three seconds, the vehicle's speed is either updated.

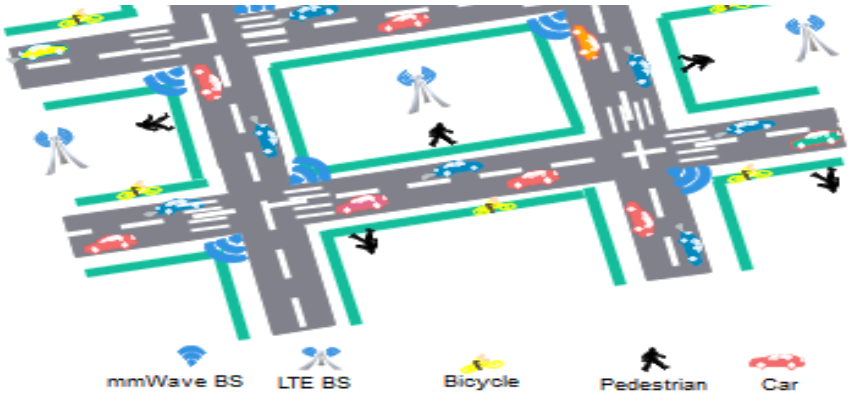

Figure 5.2: Multiuser type Mobility Model.

Given $\mathfrak{y}$ denotes user direction, the probability, $\mathbb{P}_{\mathfrak{y}r}$, of recovering the mmWave link after being blocked and remaining blocked, $\mathbb{P}_{\mathfrak{y}b}$ is [12]:

$$\mathbb{P}_{\mathfrak{y}r} = \frac{\mathfrak{y}}{K}\sum_{i=1}^{k}\frac{T^r}{T^r + T^b}, \tag{5.1}$$

$$\mathbb{P}_{\mathfrak{y}b} = \frac{\mathfrak{y}}{K}\sum_{i=1}^{k}\frac{T^b}{T^r + T^b}, \tag{5.2}$$

Where $K$is the total number of samples whist $k \in K$ is the number of possible blockings statuses, T^b and T^r are mean LOS and NLOS windows within a transmission range d given Kobservations. The rate of channel links switching from blocked to recovered/vice versa within d is $1/T^r$ and $1/T^b$ respectively. Otherwise $\mathfrak{y}$ is binary and is taken as 1 when a user is moving towards a target BS. $\mathfrak{y}$, is assumed to be 0 if a user is moving away and loses connection. The argument is link recovery chances are high if a user is moving toward the direction of the serving mmWave BS.

5.4 System Model

In this section, we first introduce the system model if the mmWave mobile network and then formulate the optimization problem mathematically.

5.4.1 Outage Probability and Data Rates

The outage probability, P_π , given an observable set of signals,Y_k is marked by assuming Θ is a set of network parameter set for a specific access policy π[11]:

$$P_\pi(Y_k|\Theta) \triangleq P\left(\sum_l \sum_{s_k} b_l \log_2\left(1+\gamma_t(x)\right) \geq r^{m\zeta}(\hat{\gamma}_t)\right), \tag{5.3}$$

where γ_t and $\hat{\gamma}_t$ indicate actual and target SINR respectively. $r^{m\zeta}$, indicate the target user throughput for a given user state $s_t \in S$. b_l represents bandwidth for mmWave link l. We assume the gNBs transmit directionally equal power, P. The receiver sensitivity is denoted by, x_{kmin}. Each serving gNB given, P, ought to satisfy at least x_{kmin}. Moreover, given a threshold x_{k0}, where $x_{k0} > x_{kmin}$, any user-gNB link transmit power requirements must not surpasses P and must at the very least meet x_{k0} to establish a new connection or may end up losing a connection, i.e., a truncation outage [16]. The maximum User-gNB transmission distance is $d = (P/x_{k0})^{\frac{1}{\propto^{k_L}}}$ to avoid losing connectivity. For x_{k0}, LOS and NLOS maximum distances are: $(P/x_{k0})^{\frac{1}{\propto^{k_L}}}$ and $(P/x_{k0})^{\frac{1}{\propto^{k_{NL}}}}$, respectively. Given received power x_t, is changing for a typical user at d from a target gNB, the receivable power projected is:

$$x_{t+1} = \max \sum_{x_t, u_t} \left\{ \frac{\hat{\gamma}}{\gamma^{min}} x_t - \frac{\alpha {u_t}^2}{\beta \hat{\gamma}^2} \right\}, \tag{5.4}$$

where $\{.\}^+ = \{max, 0\}$. x_t and u_t are current possible LoS and NLoS received power, respectively. $\hat{\gamma}$ is the targeted and γ^{min} is the immediate measured SINR needed to satisfy the $\mathbb{R}^m$. The parameters, α and β, are power and SINR scaling factors. Following user scenarios changing from NLOS to LOS, they account for power discrepancy in the projection. The data rate given P and corresponding, x_t, at t is defined as:

$$r^m = b \log_2 \left(1 + \frac{P|h^H \boldsymbol{p}|^2}{(1+ d^{\propto})} F_x\left(\left|\theta_k^l\right|\right)\right), \tag{5.5}$$

$$\varphi_k^l(°) = \frac{1.4\ x\ 10^4}{f_c(GHz).v(km/h)},$$

$$h^H p = \sum_{\hbar=1}^{K} g_{\hbar}(t) e^{-2\pi i f_d \cos\left(\varphi_{\hbar}^l\right)t},$$

where $\theta_k^l = \frac{2d \sin \varphi_k^l}{\lambda}$ represents the angle of arrival for beam p, v is velocity of a user moving at 50 Km/h or less. f_c, denote the carrier frequency. f_d, represents maximum Doppler shift while $\varphi_k^l(°)$, denotes the main lobe angle of arrival. $|h^H \boldsymbol{p}|^2$ is channel gain [6] while $F_x\left(\left|\theta_k^l\right|\right)$ denotes the Fejér kernel function. For user speed approaching zero, $F_x\left(\left|\theta_k^l\right|\right) \to 1$, hence SINR approaches maximum over distance, d. Otherwise, SINR and F_x approaches 0 as v increases [4]. α, is path loss exponent [9] and their respective LOS and NLOS pathloss exponents are denoted by: $\propto^{k_L}$and $\propto^{kN_L}$, $\forall k \in \{1, \dots, K\}$. $g_{\hbar}(t)$, is the time-varying gain of the channel K clusters.

5.4.2 Data Rate Optimization

The minimum rate, $\mathbb{R}^m(x_{k0} > x_{kmin})$, requirement problem in state,s_t, given outage and power constraints at distance d from a BS can be defined as:

$$\max_{\Theta} \sum_{t} \sum_{S_t, l} \left(1 - \beta_l \left(P_{\pi}^{m|x_t} P_{NL_k} + P_{\pi}^{m|u_t} P_{NL_k}\right) r_l^m(y)\right) \geq \mathbb{R}^m, \tag{5.6}$$

where $P_{\pi}^{m|x_t}$ and $P_{\pi}^{m|u_t}$ are LOS or NLOS conditional outage probabilities for a user in the m^{th} state, respectively. r_l^m is the maximum attainable data rate for the $l - th$ link at d. β_l is a binary factor. It is 1 or 0 for either $P_{\pi}^{m|x_t}$ or $P_{\pi}^{m|u_t}$ depending on which one the user is active. The targeted received power, x_{t+1} needed at $d + 1$ to sustain $\mathbb{R}^m$ condition in (4.6) given outage SINR constraints in (4.2) -(4.3) for typical user at d from a target link has a corresponding transmission energy cost, c_t, is influenced by two components: received and lost packet energy consumptions:

$$c_t = \beta \left\{ x_t \frac{c(t-w)}{\mathbb{R}^m} + e_0 * \zeta c(t-w) \right\}, \tag{5.7}$$

where β denotes the energy price unit energy, for transmitting $c(t-w)$, actual total number of packets received at *t* over window *w*. $c\ (t-w)/\mathbb{R}^m$ is the latency. x_t denotes the current received power at time *t*. e_0 is the unit energy per packet, $e_0 * \zeta c(t-w)$ denotes energy cost owing to lost packets during window *w*. Given constraints of receivable,x_t,and transmittable power, P, for optimum packet delivery latency, the problem to maximize link utility is formulated as:

$$\max \sum_{P} \sum_{x_t,} \{x_{t+1} \delta c(t-w) - \zeta P E_c\}, \tag{5.8}$$

where $x_{t+1}\delta c(t-w)$ denotes the optimum latency given x_{t+1} is received and packets $c(t-w)$ during window *w*. ζE_c is the latency discrepancy at time following received power drop from x_t to x_{t+1} as the user moves away (i.e., from transmission distance d_t with E_c to d_{t+1}) from the serving BS. Prior to HO execution, we learn to predict the long-term received power deterioration pattern $\{x_t, \dots, x_T\}$ of target links as user with different speeds move away from target BS. This is to avoid selecting target links whose performance is likely to abruptly dip due to sudden blockage/NLOS effects despite showing signs of high initial performance. we utilize JMLS properties to predict the likely gradual or abrupt deterioration behaviour of target links $\{x_t, \dots, x_T\}$. The JMLS uses difference or differential equations to model continuous dynamics within each link operation state, and the abrupt deterioration patterns between the states are described by a Markov chain [117].

5.4.3 Energy-efficient HO and power control

EE is downloaded bits per energy consumption in unit of bits/J. We consider EE_0(t) as the EE when the inactive BS allows the user association and EE_1(t) as the current EE after association. By the above EE definition, EE_1(t) and EE_0(t) are represented as:

$$EE_0(t) = {}^{\mathbb{R}^m}/_{P_c(t)}, \tag{5.9}$$

$$EE_1(t) = {}^{r^m}/_{(P_c(t) + P_o(t))}, \tag{5.10}$$

where $P_c > 0$ is the constant power consumption, $\mathbb{R}^m$ is the average rate given by the user's currently associated BS with transmit power P_o, and r^m is the rate after association. if the target dormant BS turns on and associates with the user with transmit power P_o. BS's HO energy

efficiency decision is carried out by comparing $EE_0(t)$ and $EE_1(t)$. Each inactive BS accepts HO when its $EE_1(t) > EE_0(t)$.

5.4.4 Distance-Weighted Degree-of-Connectivity

The degree of connectivity of a user is a measure of the number of its 1-hop neighbour BS to the serving BS. For a mobile user, in LOS, the closer edge of the transmission boundary of the serving BS is the user's 1-hop and is more likely the user will lose connectivity in near future. Therefore, we calculate the distance-weighted connectivity between the user and BS as a function of their Euclidean distance d_E, and transmission range, $d_x = \left(\frac{P}{x_{k0}}\right)^{\frac{1}{\propto k}}$ as alluded earlier. It is defined as:

$$\psi_d = \begin{cases} 1 & d_E \leq (0.6 * d_x) \\ 2.5\left(1 - \dfrac{d_E}{d_x}\right) & (0.6 * d_x) < d_E \leq d_x \\ 0 & 0.6 * d_x \end{cases} \tag{5.11}$$

The distance weighted degree of connectivity $\mathcal{D}$ of N user of type m moving in v range of velocities is:

$$\mathcal{D} = \sum_{v,m} \psi_d, \tag{5.12}$$

5.4.5 Problem Formulation

We focus on the joint design of the mmWave cell selection reliability, data rate optimization and energy optimization s to maximize the downlink sum-rate of the system while being energy efficient conscious in HOs. The sum-rate maximization problem is mathematically formulated as:

$$\max_{x,d} \sum_{s=1}^{S} b \log_2 \left(1 + \frac{P|h^H \boldsymbol{p}|^2}{(1 + d^{\propto})} F_x\left(\left|\theta_k^l\right|\right)\right), \tag{5.13a}$$

$$\text{s.t} \quad x_{t+1} \leq P, \tag{5.13b}$$

$$x_{t+1}\delta c(t - w) > \zeta P E_c, \tag{5.13c}$$

$$EE_1(t) > EE_0(t), \tag{5.13d}$$

The constraint (5.13b) guarantees that each projected receivable power needed to sustain the minimum targeted data rate exceeds the threshold of the receiver sensitivity but does not exceed

the transmission power. Otherwise, this may result in truncation outage. [135] ensures that the projected latency doesn't exceed that owing due to minimum data rate. Equation (5.13d) ensures that the HO decision is energy efficiently. These constraints guarantee that we select BSs to efficiently optimise to serve *K*, users. It is difficult to solve problem (5.13) since it is an MINLP, which is non-convex and involves discrete variable.

Though we can use an approximate algorithm to solve it, the scalability of the solution is very weak. In fact, owing to user mobility, data rates, latency, and the energy consumption are time-varying. We learn to predict the long-term cell pattern $\{x_t, \dots, x_T\}$ for HO decision. We propose a joint NN design to solve it in the following sections.

5.5 Framework of the Joint Design

In this section, a framework of joint NN design for Handover selection and Handover Execution is developed. The proposed framework includes two parts: joint NN design and EM estimator, which are elaborated as follows.

5.5.1 Mobile User HO based Pattern

The HO selection process for a user is recorded using the record matrix in, MxM with M denoting the full number of cells in the area of interest in Figure 5.1. The HO history matrix is denoted:$\boldsymbol{Q}$

$$\mathbf{Q} = \begin{bmatrix} q_{11} & \cdots & q_{1m} \\ \vdots & \ddots & \vdots \\ q_{m1} & \cdots & q_{mm} \end{bmatrix}. \tag{5.14}$$

The preliminary elements of $\boldsymbol{Q}$ are set to 0. Then as users join the network, and HO based on SNR (i.e., from one mmWave cell i to cell j, the value of q_{mm} is incremented by 1 and the network parameters in between HO are estimated by JMLS and cell reordered by JMLS. The diagonal elements of $\boldsymbol{Q}$ are always 0, any HO pattern that skips mmWave and engages LTE is awarded 1. The mmWave HO probability transition matrix in $\boldsymbol{Q}$, is matrix $\boldsymbol{P}$**:**

$$\mathbf{P} = \begin{bmatrix} p_{11} & \cdots & p_{1m} \\ \vdots & \ddots & \vdots \\ p_{m1} & \cdots & p_{mm} \end{bmatrix}, \tag{5.15}$$

where $p_{ij} = q_{ij}/\sum_i q_{ij}$ transition probability and satisfies the conditions; $\sum_i p_{ij} = 1$ and $0 \leqslant p_{ij} \leq 1$.for mmWave. Here, the value of p_{ij} is the probability that the user performs a mmWave HO from cell i to j; $i,j \in S$, where the system state space is given by **S**. The receivable power pattern is $\{X_0, X_1, \ldots, X_n\}$ at location of the users $\{d_0, d_1, \ldots, d_n\}$ at the consecutive points of HO observation $0,1, \ldots, n$. This can be constituted into a discrete-time Markov chain which can be modelled by JMLS and optimized by DDPG. It is possible to compute the probability mass function of the random variable of receivable power X_n; i.e., the probabilities at state $P(X_n = i|S_n)$ that the DTMC is in state i at HO step n. Given the n-step power transition probability matrix P^n can be estimated. The JMLS and DRL is select BSs that satisfy 5.13 and recalibrate (5.14) HO patterns.

5.5.2 JMLS System Definition

We first reformulate resource allocation problem in (5.13a to 5.13d) into a JMLS learning form with system state, action, and reward defining the deterioration pattern.

A) The JMLS Representation

We JMLS model the mmWave link behaviour and reflect different states following which the mmWave link received power may deteriorate after HO, i.e., owing to varying user types and topographic dynamics. Letting x_t, be the receivable power as in (5.2), over a given state of a user at t, time. We first reformulate (5.2) – (5.10) into a JMLS problem form in (5.16) and then project the desired x_{t+1}, that the target gNB must meet before a HO is granted. Mathematically, we define the JMLS modelling equation as in (5.16):

$$\begin{cases} x_{t+1} = A(s_t)x_t + B(s_t)u_t + w_t \\ y_t = r^{min}(s_t)x_{t+1} + v_t, \\ \mathcal{M} = (\Theta, P(S), \pi, P_\pi) \end{cases} \tag{5.16}$$

where $x_{t+1} \in X$ denotes received power quantities in (5.3) given state, s_t, $u_t \in \mathcal{U}$ represents received power discrepancy estimates flocking any blockage/NLOS and is related to x_t by: $u_t = -Kx_t$; $A(s_t)$, where K is the control factor. $B(s_t)$ are SINR/power coefficient matrices in (4.3). $v_t \sim \mathcal{N}(0, Q(s_t))$ and $w_t \sim \mathcal{N}(0, R(s_t))$ are transmission rate and state noise, respectively and are state-dependent. $I \cap P = \Theta$ is reflected by the governing parameter set $\Theta = \{A, B, R, r^{min}, Q\}$,$s_t \in S$, is the system space reflecting the observed vehicular environment and at time slot t can be defined as:

$$s_t = \{v, r_t, T_t, d_t, \eta_t\}, \tag{5.17}$$

where: $v = [v_1, \dots, v_T]$ is a vector possible user velocity. $r_t = [r_1, \dots, r_t]$ denotes the vector of possible data rates satisfying x_{t+1} projections. $\boldsymbol{T_t} = [\boldsymbol{t_1^m}, \dots, \boldsymbol{t_N^m}]$ is the vector average service time for x_{t+1}. $d = [d_t^m, \dots, d_T^m]$ is a vector of transmission distances. $\eta = [\eta_1, \dots, \eta_N]$ denotes a possible user direction in n^{th} with respect to serving gNB.

For a transition from s_t to s_{t+1}, the corresponding projected received power transition is denoted , $x_t|x_{t+1}$, the immediate reward, $r^{min}(s_t)$, for the observed signal,$y_t \in Y$, is defined as a function of energy cost efficiency: $r^{min}(s_t)$.The transition probability of a gNB link getting blocked or unblocked between states given transitions from s_t to s_{t+1} with corresponding receivable power transition x_{t+1} to x_t is given by:

$$P(S) \triangleq P\big(s_{t+1} = m_j | s_t = m_i\big), \tag{5.18}$$

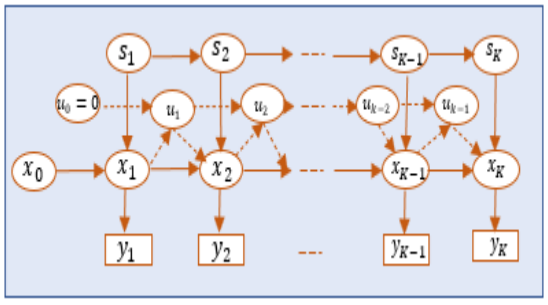

Figure 5.3. Dynamic Bayesian representation of JMLS

B) The JMLS Cell Pattern Estimates

let Y_T, represent a patterns of observable mmWave signals $\{y_1, \dots, y_T\}$. X_T denotes a set of the corresponding receivable power values $\{x_0, \dots x_T\}$ while $S_T = \{s_1, \dots, s_T\}$ denotes a set of states

between each HO in (5.14) up to triggering time T. Given $A_T = \{a_1, \dots, a_T\}$ is a set of SINR values in (5.16) desired over S_T, to meet the conditions in (5.13), the JMLS learning problem in each user type HO is to predict if the likely, X_T and given, Θ will meet the condition in (5.13). EM is used and uses the maximum likelihood function $P(X_T|\Theta, Y_T)$ given a finite observation in Y_T over S_T for all k_1,.., $k_2 \in T$ at distance $k_1 \leq k_2$. It automatically infers optimal value set of Θ for X_T[12] at each step k using Markov chains as seen in Figure4.3. The value function $\mathbb{Q}$ is defined by:

$$\mathbb{Q}(\Theta|\Theta^k) = \mathbb{E}[\log P(X_T, A_T, S_T|\Theta, , Y_T)|\, Y_T, \Theta^k], \tag{5.19}$$

Such that

$$\Theta^k = \arg\max_{x \in X} \mathbb{Q}(\Theta|\Theta^k)\ , \tag{5.20}$$

where $\Theta^{(k)}$, is the current parameter estimate at iteration k. Equation (5.22) prediction improve every step k. For a cell to qualify in updating matrix (5.14). The anticipated received power change, Δ, between x_k and x_{k+1} from s_k to s_{k+1} states must at least meet the receiver sensitivity threshold to avert outage condition in (5.1). Furthermore, the smaller the difference, the lower the change between x_k and x_{k+1}, defining deterioration pattern X_T. The update in (5.14) from $q_{mm}(s_k)$, to $q_{mm}(s_{k+1})$ can be chosen independent of time.

C) Fine Tunning

Given the distance weighted degree of connectivity $\mathcal{D}$, and formulating cells at each,$n-th$, HO, DRL is used to improve the cell selection criteria in (5.14). The cell with a best reward with respect to (5.13) is chosen among the possible ones based on $\mathcal{D}$ contributors. DRL however, requires full CSI information to formulate reliable policies. Obtaining full or accurate CSI to determine the pattern may be difficult owing to rapid changes in mmWave channels. At the same time, EM, estimates of target cells do not consider how partial NLOS and reflected signal can be relied on to avert link failure or failing back to LTE links [12]. To refine pattern X, without more CSI about Y_T, we use an online DRL scheme that uses EM estimations data over S, states as initial input experience for DRL training to determine user target data rates Y_T . As seen in Figure 5.4, and following a transition to $x_{t+1}$from x_t, the immediate reward, $r^{min}(s_t)$, for the observed signal,$y_t \in Y$, is defined according to constraint (5.13d):

$$r^{min}(s_t) = \frac{r^m(s_t, a_t)}{P_t}, \tag{5.23}$$

where $r^m(s_t, a_t)$ is a data rate greater than minimum rate $\mathbb{R}^m$. DRL however predicts the long-term benefits of choosing a particular mmWave cell over others. For simplicity, the long-Term reward over policy $\pi(a_k, x_k|Y_K)$ approximates the SINR [15], with respect to distance, d_n, and velocity v:

$$\pi(a_k, x_k|Y_K) = \sum_{k=1}^{K} \frac{P|h^H \boldsymbol{p}|^2}{\left(\sum_{i\neq k}^{K} P\left|h_i^H \boldsymbol{p}\right|^2 + d_i^{\propto}\right)} F_x\left(\left|\theta_k^l\right|\right), \tag{5.24}$$

where v_i is the velocity index of the user at the selected time step $t < T$ of learning.

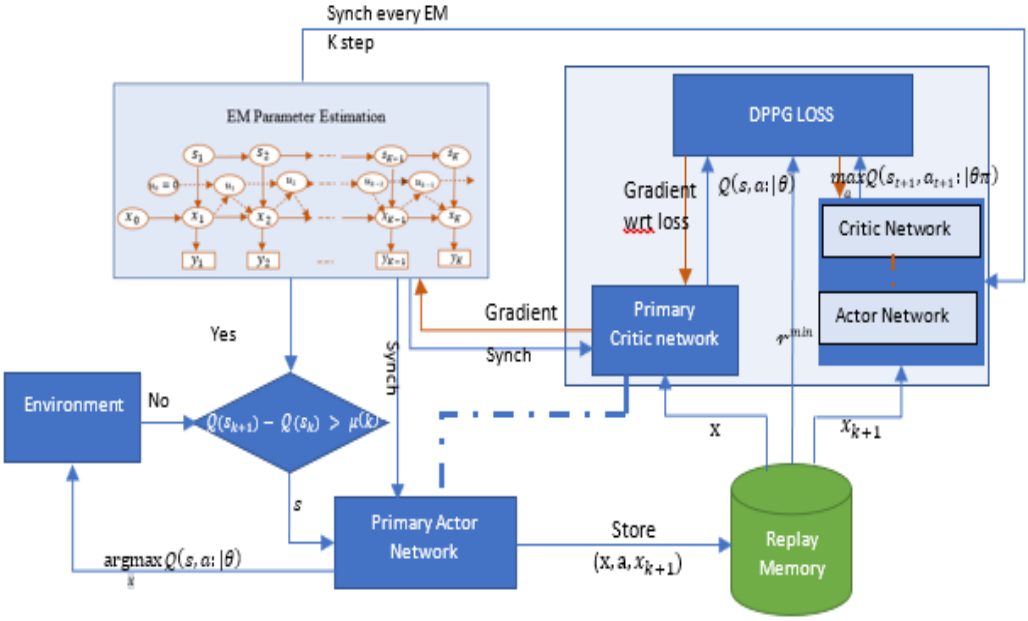

Figure 5.4: Deep Deterministic Policy Gradient (DDPG) algorithm structure.

Using the JMLS EM estimates as shown in Figure 5.3 , the DRL extends to the EM distribution implementation as follows in Figure 5.4. The JMLS EM-parameter buffer maintains the current EM parameters Θ^k. The DRL actors and critics map EM parameters to their current Q-network $Q(s, a;\ \theta)$ (see next section). These are synchronized from the EM parameter server before every re-learning step from the same data set. The learning agents furthermore maintain the target cell sequences in (4.15) initially updated by EM only. in their Q-network Q (s, a; θ). The learner's target network is updated from the EM parameter buffers, Θ^k, after every N gradient updates in the central parameter server. Note that N is a global set that counts the total number mmWave HO hence updates to the EM parameter buffer rather than counting the updates from individual/local

learner. DRL then updates deterioration gradients using the DQN gradient given in Equation 5.27. Here, gradient updates are not applied directly link predictions, but instead communicated used to update and adjust EM deterioration predictions. The EM parameter server then applies the updates that are accumulated from many learners.

5.5.3 Deep Reinforcement Learning in EM-Estimates

As can be seen in (5.20), EM estimates provide the highest achievable received power, or the upper bound of desirable received power in each state required to efficiently achieve high SINR, A_x, and consequently a high data rate, y. In order to efficiently obtain the same JMLS value, a_x, for a given state, given the optimal maximum receivable power, x*, the role of DRL is to determine the minimum or lower bound of receivable power, x*. The fact that the power at the receiver can change erratically with time, space, and frequency must be acknowledged and emphasized. If left unchecked or noted could lead to inaccurate reception at the receiver. To thus meet desired receivable power and receive the same amount of user data within a given QoS/SINR requirement in NLOS scenarios, the gNBs may need high transmit power levels. This is however energy inefficient and can lead to truncation outage. Even though the received power may be lower (as in NLOS), however, if the channel's gain is high at its peak, it will still be possible to receive the same or a similar amount of data while maintaining the same QoS and SINR. Therefore, prior to making a HO decision, it is crucial to understand both the pattern of the maximum and minimum receivable power. In order to determine the minimum desirable power given the maximum by EM estimation, DRL is therefore required. The least expected receivable power required to provide, A_x is here determined by DRL using EM data as initial experience (meta data). The DRL agent must then find the power, x, that provides the highest directly attainable reward in addition to the anticipated accumulated future reward of the resulting states, s, while only taking into account the SINR value, $a \in A_x$, possible for -x, in EM.

The EM $\mathbb{Q}$ value for the $(-x, a)$ −pair, is used as meta data by the agent to find the SINR that gives the smallest DRL Q value with a function value, V. The optimal value function V^* is obtained by solving x_{ko} for each given $-x_{ko}$ in Figure 5.3

$$V^*(x_{ko}) = \max_{\pi} \mathbb{E}\left\{ r^{min}(-x^*{}_t, a(s_t), x^*{}_t)|_{s_t,\pi(x_t|\theta\pi)} \right\}. \quad (5.25)$$

Technically, for a given optimum pattern, $-X^*$, in (5.25), the algorithm uses corresponding optimized parameter sets $\theta\pi$ and policy $\pi(s_t|\theta\pi)$ as input to DRL. The DRL scheme then determines the minimum desirable value, x_t, needed to achieve $a(s_t)$. It uses corresponding maximum value, $-x_t$, determined by EM in each state s_t, as initial experience and improves it by minimizing expected energy cost, $\mathcal{J}(x_t)$. The policy, $\pi(s_t|\theta\pi)$, is defined as:

$$\pi = \underset{\mathcal{J}}{argmin}\left\{\mathcal{Q}(a_t, x_t|\theta\pi) + \varepsilon \sum_{s_t \in S} P_\pi(x_t|-x_t, a_t)\mathcal{J}^*(x_t)\right\}, \tag{5.26}$$

where $P_\pi(x_t|-x_t, a_t) \rightarrow [0, 1]$denotes the probability of transition from $-x_t$ to x_t without change, $a \in A$, with least possible energy cost $\mathcal{J}^*(x_t)$, in s_t. The optimal policy π, derives the smallest possible value of $\mathcal{Q}(-x_k, a_k, x_k|\theta\pi)$, satisfying the following Bellman equation:

$$\mathcal{J}^*(x_t) \triangleq \min_x \mathbb{E}\left[r^{min}(a(s_t), x_t^*) + \sum_{x_t \in X} P_\pi(x_t|-x_t, a_t)\mathcal{J}(x_t^*)\right], \tag{5.27}$$

where s^* or distance are goal states where condition (5.23) is satisfied.

A) The Deep Deterministic Policy Gradient (DDPG)

Deep Deterministic Policy Gradient (DDPG) is employed to enhance the precision of the pattern. DDPG combines with DQN based on the EM algorithm to increase the network training's stability and efficiency. This facilitates the resolution of problems associated with continuous state and action space. Technically, DDPG uses DQN's experience replays memory and the target network to approximate the EM function values in neural networks and solve the non-convergence problem. Consequently, this algorithm is actor-critical model. To learn policies, it employs highly dimensional observation and action spaces. Agents utilize the primary network, the target network, and replay memory modules for this purpose.

Using a policy gradient method, primary networks match actions (SINR ratios in JMLS parameter sets) with expected received power. It is composed of primary-actor neural networks and primary-critic neural networks. In contrast, the target network sets target values for the optimal receivable power pattern, X, as estimated by EM. Given a condition, the replay memory stores, via the actor network, the tuple experience from EM Bayesian estimators and the environment (5.13d).

Experience tuples include the current and next states, the SINR ratio value following the state transition, and the reward for selecting the received power level in X T. For eq. (5.13d) scenarios, replay memory updates are sampled at random for training the primary critic network and establishing the target in the target network. Given EM parameter set θ and policy $\pi(s_t|\theta\pi)$, the cost policy gradient $\nabla_{\theta\pi}\mathcal{J}$, gives the values of $x_t \in X_T \forall\, y_t$ with a minimum change in $\nabla_{\theta a}\mathcal{Q}(a_t, x_t|\theta\pi)$ between $-x_t$ and x_t, and corresponding maximum change $\Delta\mathcal{r}^{min}(-x_t, a_t, x_t, s_t)$ for each value ,x_t, transitioning from $-x_t$ and is defined as:

$$\nabla_{\theta\pi}\mathcal{J} \approx \max_{\pi}\mathbb{E}\left[\Delta\mathcal{r}^{min}(a_t, s_t)|_{s_t,\pi(x_t|\theta\pi)}\nabla_{\theta a}\mathcal{Q}(a_t, x_t|\theta\pi)\right]. \tag{5.28}$$

The optimal value $\mathcal{J}^*(x_t)$ gives the highest possible expected future reward and lowest discrepancy from target values for each state. The policy gradient is explored by the primary actor neural network and the value function $\mathcal{Q}$ for the (x, a)-pair, is used by the agent to find the SINR ratio, a, and received power x that gives the lowest $\mathcal{Q}$ value and highest reward. Value iteration in DDGP terminates when ∀s ∈ S, $|\mathcal{J}_k(\text{x}) - \mathcal{J}_k(\text{-k})| \leq \varepsilon$ and termination is guaranteed for $\varepsilon > 0$. ε is similar to Greedy strategy with probability $1 - \varepsilon$[27]. Here ε decays as more iteration hence experience is gained. The primary critic network updates θa by minimizing loss function $Ls(\theta\pi)$ defined as:

$$Ls(\theta Q) = \mathbb{E}\big(\hat{y}_t - \mathcal{Q}(a_k, -x_k|\theta\pi)\big), \tag{5.29}$$

Where $\hat{y}_t$ is the target network value and can be obtained by:

$$\hat{y}_t = \mathcal{r}^{min}(a, x_t) + \varepsilon\mathcal{Q}^k(x_k,\pi^k(s_{k+1}|\theta_\pi^T)|\theta_a^T). \tag{5.30}$$

Here $\varepsilon\mathcal{Q}^k(x_k,\pi^k(s_{t+1}|\theta_\pi^T)|\theta_a^T)$, is obtained through the target network, i.e., the network with parameters $\theta\pi$, from EM with -X values and θa from X generated overtime for minimum desirable receivable power.The new values of targeted X pattern are updated by minimizing loss in (5.29). The gradient of $Ls(\theta Q)$ over $X_T$is calculated by its first derivative, which can be denoted as in [14]:

$$\nabla_{\theta\pi}Ls(\theta Q) = \mathbb{E}\ \Big(2\big(y_t - \mathcal{Q}(a, x_t|\theta\pi)\nabla_{\theta a}\mathcal{Q}(a, s_t|\theta\pi)\big)\Big), \tag{5.31}$$

According to (5.31), the parameter θ_Q of primary critic neural network can be updated. Specifically, at each training step, with a mini-batch experience$< s_t, a_t, R^{imm}, s_{t+1} >, t \in$

$\{1, \ldots, k\}$, randomly sampled from replay memory. For each point in X_K, the target network value is regarded as the previous and current version of EM parameters $\theta_\pi^T$and θ_Q^T. At each iteration, $\theta_\pi^T$and θ_Q^Tin (5.31) and (5.32) are updated with a weighted combination of the previous state. The prediction of target path takes the form of a distance-weighted combination of models:

$$\theta_\pi^T(\tilde{x}_k) = \omega\theta_\pi(-\tilde{x}_k) + (1-\omega)\theta_\pi^T(-\tilde{x}_k)$$

$$\theta_Q^T(\tilde{x}_k) = \omega\theta_Q(\tilde{x}_k) + (1-\omega)\theta_Q^T(\tilde{x}_k), \tag{5.32}$$

where $\omega \in [0, 1]$ is weight computed using Gaussian kernel parameterized by the transmission distance metric $d_k \in \tilde{s}_k$:

$$\omega_k = \exp\big(-0.5(x-\mu_k)^T d_k (x-\mu_k)\big). \tag{5.33}.$$

Using EM and replay updates, target neural networks generate target or ideal values for training and re-optimizing the deterioration pattern, X_T from $-X_T$. As a result, EM estimates from each iteration are used by DDPG as meta data. The initialization parameters and network structure of the target neural network are similar to those of the primary network. The target Actor and Critic network's parameters are gradually (Soft Replace) replaced with EM-estimated values during the training procedure. Here, to increase the stability of the training process, we use EM estimations as ideal initial values rather than directly and randomly training the parameters of the main Actor and Critic network. The EM experience tuples that make up X_T are stored in replay memory, and each value update of $x_t \in X_T$ also includes the update of the tuple $< -x_t, a_t, R^{imm}, x_t >$.

Figure 5.4 shows the structure of the proposed JMLS-DDPG algorithm. The DDPG algorithm determines the minimum receivable power values of a pattern using the EM parameter data set and maximum receivable power values ,-X, as initial inputs. The DDPG agents then output the minimum receive power values X necessary to maintain the same SINR ratio previously predicted and set by EM estimates for , S_K, taking into account that power effects on SINR can be reduced in locations with high channel gain. The SINR that is helpful to the agent in achieving the objective is rewarded positively, and if condition (5.13d) is not met, it is rewarded negatively. This replicates the reward of , x_k, in EM. The next minimum desirable receivable power's state information, the SINR ratio, the reward, and the current state information are all stored in the replay pool. By

randomly selecting sample data from the EM pool, the neural network gains experience and continuously modifies the SINR strategy. Gradient descent is then used to update and iterate the network parameters, further enhancing the stability of pattern, X, and the algorithm's accuracy. The search space for ideal minimum receivable power values is constrained by using EM experiences as initial training data input to DDPG. Any observed mmWave BS data rate that doesn't meet the corresponding receivable power is consequently immediately disregarded for training or consideration. As a result, the DRL training sample is smaller, which decreases the convergence time. Finally, by combining EM and DDPG predictions as the meta training sample, the enhanced DRL HO is obtained. The HO platform for HOs has finally incorporated the pattern model.

B) Online Update of Target Deterioration Path

DDPG divides the structure of the training network into the target network and the critic network (See Figure 5.4). The online network, consisting of an online (primary) Actor network and an online Critic network, is utilized to output the minimum expected received power in real-time, evaluate SINR ratio values, and update network parameters. Included are target networks such as the actor and critic networks. These networks are modified in accordance with EM values. On the other hand, the target Actor network system does not offer online training. The estimated pattern X_N for each user type is recalculated from new training samples only when the pattern prediction error based on EM estimates exceeds the minimally desired received power pattern. Therefore, the corresponding EM information used to generate the received power pattern X_t will be considered a reliable training sample for the target network in DDPG when the error given the energy efficiency is small enough that the channel gain compensates for the power loss in order to maintain the desired SINR. EM data is reencoded in order to create new training samples for the DRL and set new targets over S_t moving forward as meta-training. X_t represents the actual channel link deterioration behaviour from which Y_t is derived if the target mmWave network successfully follows the link deterioration pattern $\tilde{X}_t$. As a result, in Figure 5.4, it is not necessary to request new CSI from the environment but rather to continue using the corresponding pair $\tilde{S}_t$and Y_t of parameter set $\theta_\pi(s_t)$. With only a few new training samples, the model can be efficiently and quickly be retrained. The inherent disadvantage of decision-directed approaches, such as the Bayesian in EM, is their sensitivity to decision errors. For instance, if the link is unable to

successfully maintain connectivity, the meta-training samples $-\tilde{X}$ of $\tilde{X}$ over $\tilde{S}_t$ do not accurately represent the channel behaviour results in Y_t. This renders the proposed method unreliable, particularly in areas with low SINR where link deterioration pattern errors are common. In such situations, the inaccuracy of the training sequence could gradually diminish the precision of DDPG predictions. Nevertheless, the number of pattern errors or the effects of decision estimate errors can be used to determine when to generate meta-training when pattern errors in EM are low. For instance, in DDPG, we only retrain with new training samples when the total number of errors exceeds a certain threshold. By using only accurate meta training data, this method limits the impact of decision errors, ε. We cleverly concentrate on states with un-converged pattern values, i.e ., the (5.13b) condition not met, when using new training samples. The proposed Algorithm 5.1 provides a summary of our online training mechanism.

Algorithm 5.1: JMLS-DRL-Based Pattern Algorithm.

Input: User mobility model parameters, $\mathbb{P}_v$. v Parameters about DC communication: transmission power limits, bandwidth, channel gain, and NLOS and LOS path loss exponent.
Observed states S; Set of observed signals $Y = [y_1, y_2, y_3, \dots, y_N] \in \mathbb{R}$,
Output: mmWave Deterioration path $X = [x_1, x_2, x_3, \dots, x_N]$ for target link
1: Initialize the deterioration path estimations
2: **for** $t = 1$ **do**
3: Draw y_t for JMLS parameter estimation Θ, where $(X_T, S_T, Y_T|\Theta)$
4: **Estimate Maximization (EM):**
5: $Q(\Theta|\Theta^k) = \mathbb{E}[log\, P(X_T, S_T, Y_T|\Theta)|\, Y_T, \Theta^k]$,
6: $\Theta^k = \arg\max_{x \in X} Q(\Theta|\Theta^k)$
7: Define Pattern: $X = [x_1, x_2, x_3, \dots, x_N]$
8: **for** x_N **do**
9: **if** $Q(s_{k+1}) - Q(s_k) > \mu(k)^v$ **then**
10: **update** x_N **with DRL**
11: **else**
12: repeat step:(6) for all X
13: **end if**
14: **end for**
15: **Update EM Deterioration Path Estimations with DPPG**
16: Re-estimate $Q(s_{k+1})$ using primary network $Q(s, a|\theta_\pi)$
17: Initialize target network parameters with EM parameter set
18: Initialize replay memory using EM samples.
19: **for** each EM step **do**
20: Observe user state s_t and SINR ratio $a_t \in \theta_\pi$
21: Execute $a_t \in \theta_\pi$ and state x_t
22: Observe change in $r^{min}(a_t, s_t)$ and $Q(s, a|\theta_\pi)$
23: Update EM tuple < s_t,a_t, r^{min}, s_{t+1} > in replay memory.
24: Compute target value $\hat{y}_t$,update $Q(s, a|\theta_\pi)$ *and* minimizing loss
25: Update target neural networks
26: Update EM with θ_π and recompute step 6-12 for all X
27: **end for**
28: **end for**
29: **end for**

C) Global Path and Local Path Optimization Formulation

The local pattern is formulated based on local CSI from one mmWave BS. The local agent thus considers only the SINR ratio $a \in A_x$ and corresponding received power x values possible in the local environment over given states $\tilde{S}_t$. The long-term function for local deterioration pattern is:

$$\mathcal{Q}_{LP}(a_t, x_t|\theta\pi) \triangleq \mathbb{E}\left[\sum_{t=0}^{T} \delta^t\left\{r^{min}(a_t, x_t) + \varepsilon\mathcal{Q}(x_{t+1}, \pi\left(x_{t+1}|\theta\pi\right))\right\}\right], \quad (5.34)$$

where δ ∈ (0,1) is the discount factor and approaches 1 with more training sample. The global deterioration pattern is formulated based on a collective SINR ratio a_t and received power x_t , values from different mmWave BSs over $\tilde{S}_t$.The value function $\mathcal{Q}_{GP}$ is:

$$\mathcal{Q}_{GP}(a_t, x_t|\theta\pi) \triangleq \sum_{a\in A_{x_k}} P_\pi(a_t|x_t) * \frac{\alpha}{K}\left\{\mathcal{Q}(x_{t+1}a_t, x_t|\theta\pi)r^{min}(x_{t+1}, a_t, x_t)\right\}, \quad (5.35)$$

where $P_\pi(a_t|x_t)$ is the probability of receiving, x, given a in state s by EM. α is the learning rate over K samples in EM.

5.6 Hand-Off Considerations

With the help of the Kaiser-Meyer-Olkin (KMO) test [25], we can find out how much each local/individual mmWave target link's expected deterioration pattern deviates from its optimized global deterioration pattern. As shown in the comprehensive report table of [4], the global degradation pattern is obtained by gathering training samples from all mmWave BS with regard to user type/speed (CRT). Information about a BS's immediate surroundings is used to infer how a particular user type will experience degradation. This is comparable to the user data in a report table (RT) in [4]. For all Target BSs with a 3dB SINR or higher, the KMO indexing test is used to determine how closely a global deterioration pattern matches that of the target link at the time of the HO request. Correlation between the overall value of the KMO index and its definition of (5.36)

$$KMO_{\hat{x}} = \frac{\sum_{x\neq\hat{x}} R_{x\hat{x}}^2}{\sum_{x\neq\hat{x}} R_{\widehat{x\hat{x}}}^2 + \sum_{x\neq\hat{x}} a_{x\hat{x}}^2}, \quad (5.36)$$

where $R = [r_{xd}]$ is the correlation matrix, $A = [a_{xd}]$ is the partial covariance matrix where a_{xd} is defined as,

$$a_{x \neq \hat{x}.m} = \frac{r_{x\hat{x}} - r_{x.m} r_{\hat{x}.m}}{(1 - r_{xm}^2)(1 - r_{\hat{x}m}^2)}, \tag{5.37}$$

and

$$r_{x\hat{x}} = \frac{\sum_{t=0}^{T} (x_t - \hat{x}_t)(d_t - \hat{d}_t)}{\sqrt{\sum_{t=0}^{T} (x_t - \hat{x}_t)^2 \sum_{t=0}^{T} (d_t - \hat{d}_t)^2}}, \tag{5.38}$$

where $x_t \in X_T$, is the optimum lower bound target link value of received power at state s_t. $d_t \in s_t$ is the minimum expected user-BS link distance $\hat{x}_t$ and $\hat{d}_t$ are values for the global deterioration path. KMO test takes values between 0 and 1 and Table I summarizes index values. The general rule for interpreting measurements is in Table I. In this study, we select the target cells with KMO index of 0.751. If the KMO index value is less than 0.7, most likely the target link is not suitable for HO consideration though it might have the highest initial SINR. Additionally, during HO phase, If the serving BS still has a SINR value of 3dB, the user maintains the connection to the serving gNB. This avoids wasteful HOs. Otherwise, we execute the HO process and then go back to prediction phase.

Table 5.1: Interpretation of KMO Measure

KMO	Interpretaion
0.9 and above	**Marvelous**
0.8 – 0.9	**Meritorious**
0.7 - 0.8	**Middling**
0.6 – 0.7	**Mediocre**
0.5 -0.6	**Miserable**
Under 0.5	**Unacceptable**

5.6.1 Measurement Definition

To test whether the HO scheme is successful in reducing the number of unnecessary HOs, we counted how often the same HO was performed. When a user experiences multiple HOs, it is because the HO scheme has repeatedly reselected the same serving BS to which the user is already connected. Maintaining the connection rather than reselecting the same BS for HO is more efficient and thus wasteful. The total data rate of mmWave BSs with various HO schemes was also analyzed. We also compared the HO overhead of various schemes. If the overhead is high, then the HO scheme will be inefficient with the bandwidth it has available. Finally, we compared our

proposed scheme's performance to that of a different scheme called the DDPG only scheme. The DDPG-only method does not take context into account and does not employ meta-training techniques. In particular, it prefers to use randomly generated training samples to EM's-refined ones. In addition, we compared the results to those of a preexisting soft-HO DC model HO scheme. By taking an average of the SINR/data rate of all candidate target cells, the scheme chooses the optimal one.

5.6.2 Handoff Process

We consider that every traveler is equipped with a functional mobile device or is utilizing an in-car access connection (e.g., bus or train). The HO takes place between the mobile phones and the mmWave cells. The suggested scheme also describes the arbitrary actions of various user types. Therefore, before implementing the suggested handover scheme, basic user characteristics (speed) and BS locations should be communicated. Additionally, EM estimates about the gNB deterioration pattern are made from initial connection data and subsequent inference.

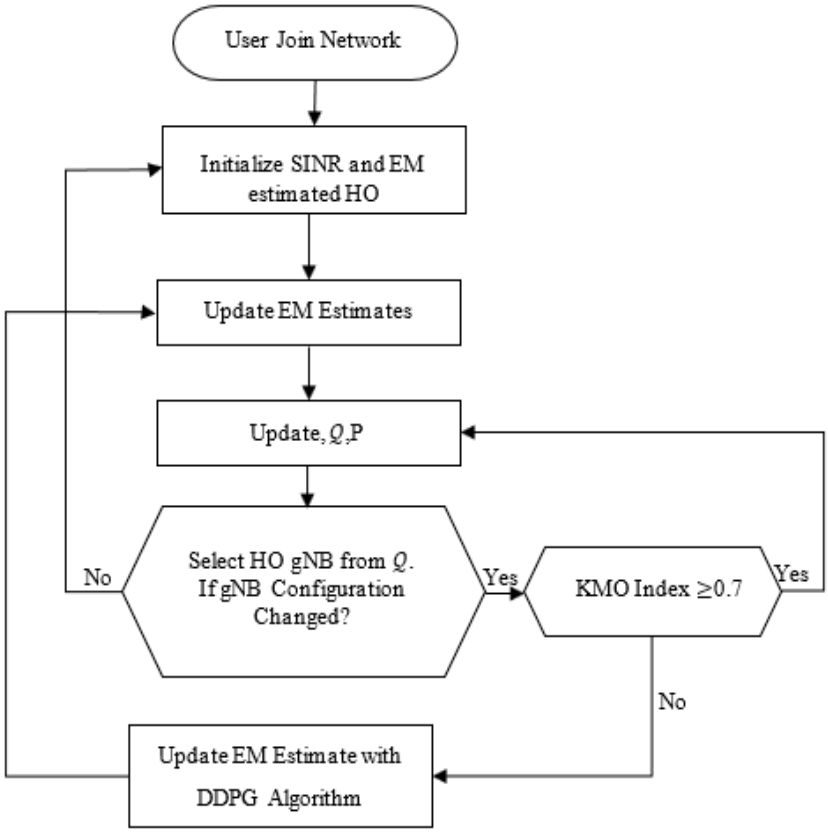

Figure 5.5: The Proposed HandOver Procedure.

Figure 5.5 shows a flowchart of the suggested handover plan. Here the first set of EM parameters are initialized as users join the network based on SINR. And a little later, they switch gNBs based on SINR threshold as matrix $\mathcal{Q}$, P, data in (5.13) and (514) are initialized. Ns-3 allows running more than one kind of HO scheme and each of them can be conditionalized. Thus before sufficient HO records have been gathered, a typical handover procedure is started based on SINR comparisons and initial EM estimates of JMLS values based on initial user connection. The current serving gNB cells is assumed by the proposed algorithm to be the first HO in (5.3), so any modifications to the cell configuration e.g., link failures will call for a new optimization by DDPG for $\mathcal{Q}$, P, using EM data. Considering, DDPG is deterministic in nature, the EM estimates give it a better estimation of formulating policies than offline. In fact, it shapes the boundaries over which DDPG can explore its data training policy. Once DDPG formulates reliable policies based on energy efficient power projections over a given state space it updates matrix $\mathcal{Q}$ based on EE /rewards. The confidence in the outcomes of DDPG selection criteria is based on KMO index. A DDPG selected cell will only update matrix $\mathcal{Q}$, P, using EM estimates if KMO index is higher than 0.7 on the selected target cell.

For two or more consecutive cells in the $\mathcal{Q}$, matrix with higher KMO index of 0.8 and above, the HO scheme can concurrently choose can concurrently choose them in one selection to facilitate two or more HO without redoing the HO selection. This reduces the selection time. It goes without saying that the target BS from one to the next needs to receive signals that are at least as strong as the absolute minimum at HO. We use a probability matrix P^n for a multi -stage handoff for this purpose. Cell q^{11} and q^{12} in $\mathcal{Q}$ with connection probability P^1 and P^2 for instance can sequentially be used without doing a reselection given their KMO indexes and initial data. At the Same time, the KMO index can be used to bypass the HO on cell in $\mathcal{Q}$ if the user is only passing through a cell with low KMO index that one selected based on DDPG results. In a city setting where the cells overlay one another, this technique works well. Without DRL, the optimization and self-reconfiguration processes are not possible. It is extremely rare for the handover to continue functioning while skipping two or more intermediate cells as the vehicle travels through a real environment. The probability matrix for handoffs occurring in three steps or more is therefore regarded based on the credibility of the data used (KMO index) beside the transitional probability.

When KMO index of the next target cell is below 0.6, the one-step LTE HO is used. This implies that learning is restarted and that the LTE cell to which the vehicle is most likely to move is selected for handover. The traditional SINR HO scheme is used for LTE link selection if the projected SINR values on the anticipate AT pattern values drop below the threshold before the expected distance is covered for another HO to be initiated. The HO system will restart a handover to the closest gNB with the strongest X pattern over distance, as determined by KMO, in order to prevent a ping-pong effect [134] and longer reliability on LTE links.

5.6.3 Numeric Results Analysis and Comparison

The following section will primarily concentrate on the numerical assessment of the JMLS-based pattern prediction HO model system's performance. In the Subsection, we analyse mmWave channels as time-invariant channels. We conduct a numerical comparison of JMLS with the conventional model-based SINR-based algorithm [34] and previously proposed speed-based HO models [50] to evaluate their performance. The robustness of our HO algorithm is evaluated against imprecise pattern training and predictions. This evaluation assesses the resilience of our HO scheme to mmWave channel interference and blockage. The parameter values in table 5.2 are used when required.

Table 5.2: The Simulation Parameter.

Parameter	Value
mmWave	28GHz
mmWave bandwidth	1GHz
3GPP Channel Scenario	Urban Micro, Urban Macro
MMWave max outage	-5dB
mmWave transmission Power	46dBm
mmWave max PHY Rate	3.2Gbps
X2 link latency	1ms
S1 link latency	10ms
RLC buffer Size	5MB
S1 MME link latency	10ms
User speed	[1,50] m/s
UDP Source rate	200Mbits/sec

We factor in a network region denoted by A. It has reference transmitter-receiver pair or pairs that are essentially constant, as well as potentially interfering transmitters drawn from a non-homogeneous Poisson Point Process (PPP) . The location of interference sources as seen from the perspective of the reference receiver is captured in a network region where the reference receiver is located. The obstacles in the human body are depicted as circles with a W-diameter. They come from a density PPP independent of and are observable to Y_i. If the diameter-W circle surrounding

some obstruction receives signal Y_j with received power, Xi from the transmitter, the reference BS is said to be obstructed to reach the reference receiver.

Considering the gain h_i owing to fading of the wireless link follow X_i behaviour at the reference receiver. We consider it as a normalized Gamma distributed random variable parameter. Thus, where X_i is not attainable because of blockage with W, diameter, we say the communication is a NLOS line of sight (NLOS) otherwise, it is in LOS (LOS). The path-loss exponent of the link receiving X_i, is α_i. It represents value(s) as α_L when received power pattern X_i corresponds to LOS values. Otherwise, it is denoted as α_N to denote NLOS scenarios between the transmitter and receiver. Additionally, we assume outage probability p_t for each received power, x_i at s_t given a constant transmit power P. Based on the above assumptions, we assume a received power pattern characterized by three parameters – SINR, transmitted power and blockages. At any time, the reference receivers are assumed to be pointed in the direction φ where the reference transmitter(s) are assumed to be located.

Assuming prior knowledge of the mmWave channel link l is based on the above assumptions, we will proceed with the following analysis in our proposed DRL schemes. In this numerical analysis, a fully connected network is implemented using three layers. The layers consist of a 1100 layer, a 10050 layer, and a 5016 layer (which is equal to the length of m). The intermediate activation functions used are sigmoid and ReLU. The SINR distribution can be closely approximated using the mixture model estimator [32], which utilises an EM-based fitting to represent JMLS as a Gaussian mixture. The training of the network involved 3000 samples for training, utilisation of the Adam optimizer [43], a learning rate of 0.01, up to 100 iterations, and a mini-batch size of 27 observations. The objective of this training is to reduce the cross-entropy loss. Compared to [12,115,121], it has been observed that the number of training samples used is equivalent to or less than the standard amount. The limited training examples allow to counter check how quick the accuracy policy can be generated given observable channel:

$$y_t = \mathcal{r}^{min}(s_t)x_{t+1} + \mathcal{w}_t, \tag{5.39}$$

A finite-memory causal scalar channel is initially considered: NLoS and blockage that incorporates additive white Gaussian noise (AWGN) are given by $\mathcal{w}_t$. In the context of the blocked channels

with AWGN, it is assumed that W[i] is a zero-mean unit variance AWGN that is independent of the state S_t. Additionally, the channel vector x (γ) ∈ X is utilised to represent an exponentially decaying received power profile, which is given by (x)τ, $e^{-\gamma(\tau-1)}$) for γ>0. The input-output relationship is defined by equation (5.39), where y_t is the output, X is the input, $h(\gamma) \in \mathcal{r}^{min}$ is a channel gain function , $S[t-\tau+1]$ is a sequence of states over distance, v_t, is the noise, and $\rho > 0$, ∈γ represents the signal-to-noise ratio (SNR) while γ is the SINR. It is important to note that while the impact of the blockage decreases with an increase in γ, the channel (7) maintains a memory length of $l = 4$ for all γ values analysed in this research.

The Error of SINR prediction is calculated numerically for various values of the parameter x, for each channel. A new DNN is trained for every individual value of received power, x. The prediction error values are computed for each SNR ρ by averaging over 20 channel vectors h(γ), where γ is varied in the range [0.1, 2] from one state to the other. The numerical computation of the Prediction Error Rate (PER) of the JMLS method, DRL and JMLS-DRL is performed for the purpose of comparison.

To examine the robustness of our JMLS-DRL to imprecise training, we evaluate its performance under the condition that the receiver is provided with a distorted version of $h(\gamma)$ and received power x_t changes over S_t states. This involves utilising a replica of $h(\gamma)$ and x_t that has been subjected to independent and identically distributed (i.i.d.) Gaussian noise with a mean of zero and a variance of σ^{2e} over s_t The Gaussian channel (7) employs $\sigma^{2e} = 0.1$. There are two scenarios that we are taking into consideration: The study compares two scenarios: Perfect CSI, where the JMLS has precise knowledge of h(γ) based on the channel model, and the other one trained using labelled samples generated with the same h(γ) as the test data.

In the second scenario, CSI uncertainty, received power for a particular state s_{t+1} is computed using the noisy version of h(γ), while the labelled data used for training is generated with the noisy version of h(γ) instead of the true one. The training process for CSI uncertainty involves dividing the 5000 samples into 10 subsets. Each subset is generated from a distinct channel that has a different realisation of the noise in h(γ). The information is randomly generated in an independent and identically distributed (i.i.d.) manner from all states in all scenarios presented in [5,54]. The test samples are produced from their respective channels, using the actual channel vector h(γ).

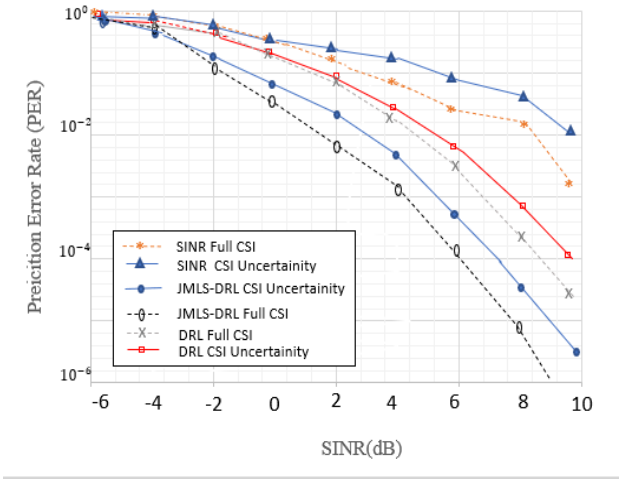

Figure. 5.6: PER versus SNR with mmWave channel.

Figure 5.6 displays the average numerically computed Prediction Error Rate (PER) values obtained from 50000 numerical simulations, as a function of $\rho \in [-6, 10]$ dB, for the channel with AWGN noise. Upon examination of Figures 5.6, it is evident that the JMLS -DRL algorithm's performance is better than conventional HO algorithm that utilises channel state information (CSI). In the case of Additive White Gaussian Noise (AWGN), where the output of the channel follows a Gaussian mixture distribution, a slight deviation is observed at high Signal-to-Noise Ratios (SNRs) due to the mismatch in the model caused by approximating the distribution of y_t as a Gaussian mixture. The utilisation of a Fejér kernel kernel estimator for on the SINR is anticipated to enhance the performance of JMLS-DRL in the given scenario. However, this necessitates a fundamental understanding of the input-output correlation, and this is done by the KMO model system. Figure 6 further demonstrate that JMLS outperforms the SINR and DRL scheme. While the SINR and DRL HO algorithm has been previously shown to approach the performance of the CSI-based JMLS algorithm when using microwaves with sufficient training, the inconsistencies of mmWave propagation in the training data provided in this study resulted in them having inferior performance.

In the case of the channel with AWGN, the JMLS and speed-based HO schemes both demonstrate a error rates of 4.7 x 10^{-3} at an SNR of 8 dB. Similarly, the SINR detector achieves a PER of 8.5 x 10^{-3} at the same SNR value. At a signal-to-noise ratio (SNR) of 28 dB, the JMLS algorithm attains a prediction error rate (PER) of 5.1 x 10^{-3}. In contrast, the SINR and Speed based HO

models achieve prediction error rate values of 6.6 x 10^{-3} and 8.3 x 10^{-3}, respectively. The findings indicate that JMLS-DRL model, which incorporates basic DNN architectures within the JMLS algorithm, necessitates notably less training in contrast to DRL-based HO schemes in [1]. It can be seen in Figure. 5.6 that JMLS-DRL scheme performs much better than the SINR algorithm in the presence of CSI uncertainty. In particular, JMLS-DRL model is able to achieve reasonably good PER performance under each of the channel conditions for which it is trained, in contrast to the conventional algorithm, whose performance is significantly diminished in the presence of imperfect CSI. While Speed based HO outperforms SINR with the same amount of uncertainty, and the performance disparity is more pronounced in the AWGN channel, SINR based HO is proven to be more resilient to erroneous CSI than the speed-based algorithm, as was also reported in [15].

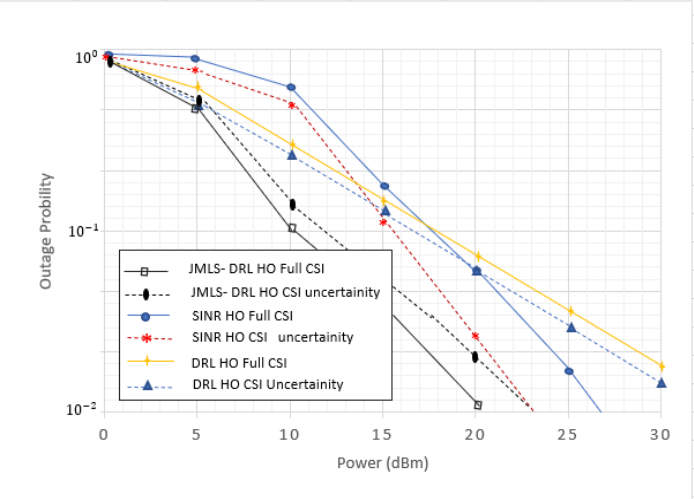

Figure 5.7: Outage Probability versus Power

Figure 5.7 shows the variation of power vs the outage probability over a sequence of states. It is clear from Figure 5.7 that the JMLS-DRL scheme achieves performance that is nearly identical to that of the algorithm with perfect CSI compared to other schemes. it is also important to note that at higher dBm, there is a slight outage Probability (OP) gap between CSI and uncertain CSI. This discrepancy suggests that training is more challenging at higher power values hence SINR, where there is less variability in the training samples than at lower received power SNR levels, and that bigger training sets are therefore needed to boost performance at high received power values.

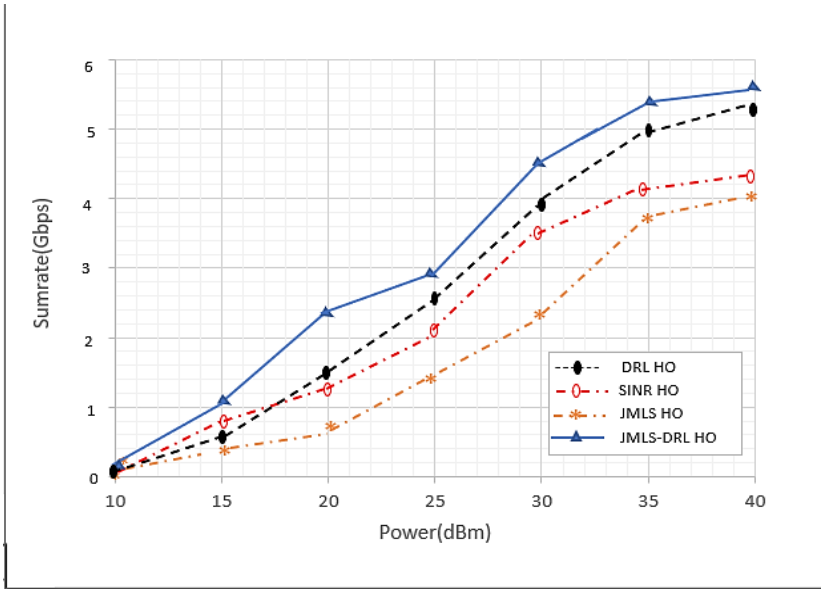

Figure 5.8: Sum Rate versus Power

To illustrate this, we show in Figure 5.8 the sum rates values attained by JMLS-DRL for the received channel power with AWGN at SNRs of 12–16 dB, or greater SNRs than those shown in Figure 5.6, when trained using training sets of sizes 2, 5, 15, and 25–103. The channel-model based algorithm with perfect CSI is contrasted against the numerically measured performance, averaged here over 106 Monte Carlo simulations to accurately capture sum rates values in the order of 105. By looking at Figure 5.8, we can see that DRL algorithm needs more training data to perform as well as our algorithm did when it was taught using samples from the same channel that it was tested, or with perfect CSI. It is also noted that our model with full CSI surpasses others using full CSI, with the exception of very short training sets, and that its performance is overall better. These findings show that JMLS-DRL based is able to attain high data rates compared to DRL training using samples from the same channel realisation due to the increased diversity created by training under CSI uncertainty. JMLSDRL's primary advantage is its capacity to precisely predict requiring simply knowledge of users and EM estimates, not much CSI is necessary. A further advantage is its enhanced resilience to unreliable CSI, as shown by the simulation research shown in Figs. 5.6-5.7. Another advantage is its capacity to operate consistently in environments where the DRL or SINR based algorithm are exceedingly challenging to implement, even when complete instantaneous CSI is available.

5.6.4 Simulation Results and Analysis

We used the DC LTE-mmWave model developed at NYU and the University of Padova in our simulation [1]. In the DC model, LTE base stations control mmWave base stations. The model thoughtfully takes into account the operation of a complete mmWave cellular network. Implemented are the 3GPP channel model for frequencies above 6 GHz, a cellular protocol stack comparable to that of 3GPP [1] and the ns-3 simulator. For its training, the JMLS-DRL algorithm makes use of the OpenAI Gym framework. Open AI Gym can be used to build RL systems and is compatible with the ns-3 simulator. It also works with ns-3 applications and other network-based learning environments to support the training of agents. By simulating the entire system, we were able to examine the end-to-end performance of the 5G mobile network. The analysis made use of data gathered from a simulation that lasted for over a thousand seconds at a TTI (Transmission Time Interval) resolution of one millisecond. Table 5.2 provides a concise summary of the most important parameters. A comprehensive review of simulators can be found in [7] –[11].

5.6.4.1 Simulation Result Analysis

Figure 5.9 and Figure 5.10 show the ratio of inefficient HOs to training episodes for the JMLS HO scheme and the JMLS-DDPG HO scheme, respectively. After the initial pattern has been defined by EM estimations for every other episode, DDPG uses EM estimated data as initial training data samples while it obtains new training samples from the environment. When estimates from EM data don't conform to EE conditions, only then does it request for more training samples. Our results show that, compared to using DDPGs-only HO scheme, the proposed scheme significantly reduces the number of wasted HOs in a shorter amount of time. For instance, while the DDPG-only scheme requires close to 400 episodes to reduce repeated HOs to minimal levels of less than five percent, it only takes 250 episodes using the proposed approach. This also means that it can effectively and strategically predicts the deterioration patterns of target links with fewer data training samples. The use of meta training techniques is justified in [4] who argues that there is little variation between training samples of received power obtained from the same channel state over time .

Figures 5.11 and 5.12 compare cumulative average reward behaviour across training episodes for various user types. Several inferences about the HO scheme performance were made on the results obtained. Firstly, DDPG-only HO scheme provides extremely imprecise early predictions as

evidenced by rewards behaviour of the deterioration pattern for various user types in the initial training episodes. This clarifies the results obtained Figure 5.9 which exhibited a high number of inefficient or redundant HOs that occur early when using DDPG-only HO compared to JMLS-DDPG scheme in Figure 5.10.

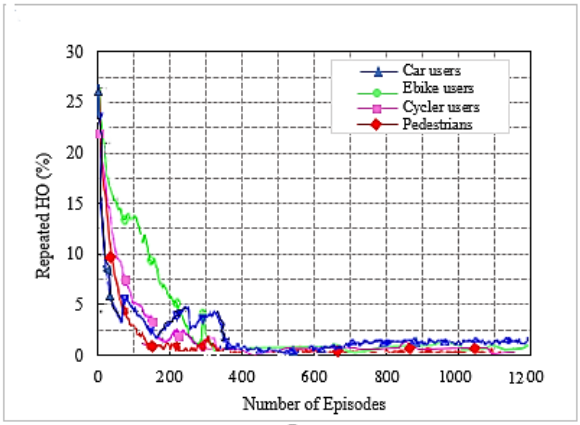

Figure 5.9: Number of wasteful HO vs Number of training Episodes for DDPG only HO scheme.

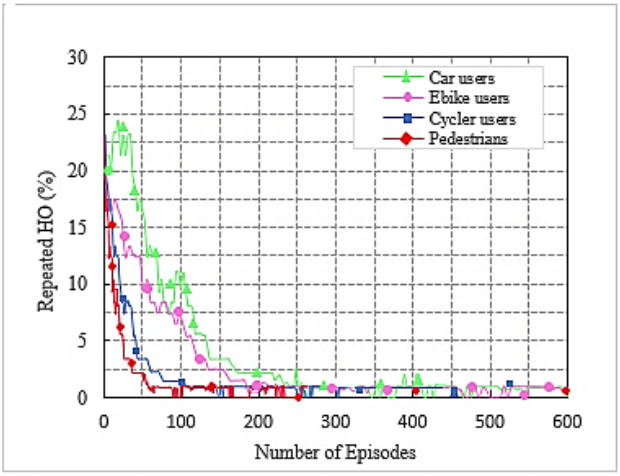

Figure 5.10: Number of wasteful HO vs Number of training Episodes for JMLS-DDPG only HO scheme.

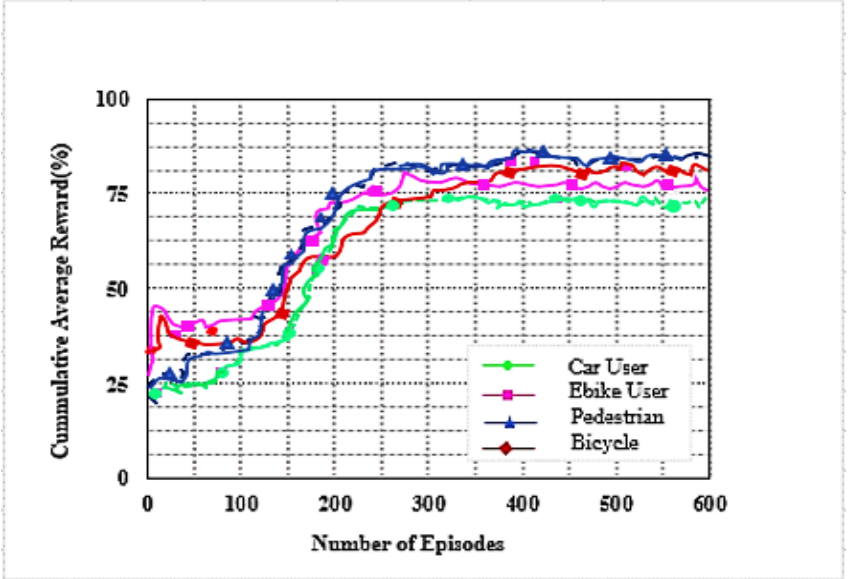

Figure 5.11: The cumulative reward vs Number of training Episodes for DDPG only HO scheme according to user type.

Secondly, as shown in Figure 5.12, the proposed JMLS-DRL-scheme converges with nearly the same and greater reward for all user types, whereas in the DDPG-only scheme convergence is independent for each user type, (see Figure 5.11). What this means is that after 200 episodes, the JMLS -DRL algorithm can adopt a universal deterioration pattern of user received power with respect to a specific target link. Conversely, the DDPG-only HO scheme necessitates a unique deterioration pattern for each user type. This implies it has to formulate a unique policy for each user type with regard to a specific target link. This prolongs its decision time and compromises its accuracny when HO decisions need to be made quickly. On the other hand, this makes our proposed scheme to formulate a single policy for all user types. Ultimately it makes out proposed scheme easy to foresee how the target link regardless of the user type will behave and quickly make HO decision irrespective of the user type. Based on this analysis, our scheme is better and faster than the alternative in selecting target links . Nevertheless, it must be emphasised that both schemes rely on the received power at a given state/distance from the serving BS not dropping below the a certain corresponding value i.e. receiver sensitivity value. Otherwise this can lead to pattern prediction inaccuracies and cause unwarranted HO.

In addition, we investigated the level of predictions accuracy at different training episodes. Two different predicted deterioration patterns after 200 and 500 episodes are shown in Figures 5.13 and

5.14, respectively. Comparing Figure 5.14 results after 500 episodes with Figure 5.13 after 200 episodes or observations, our proposed JMLS-DRL-enabled HO algorithm provides a more accurate estimation of likely received power for different user types after 500 episodes.

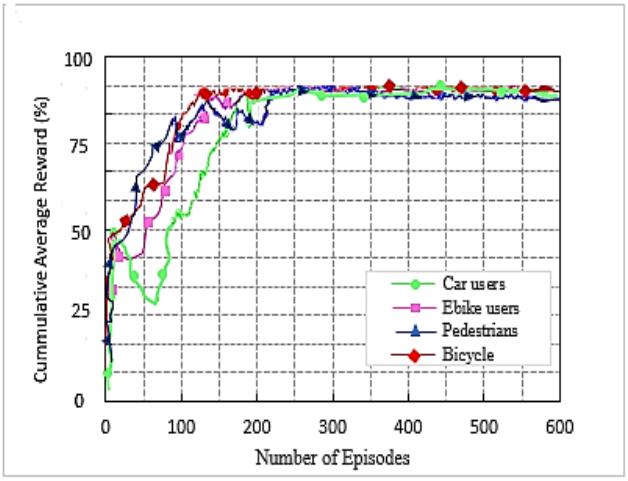

Figure 5.12: Compares the cumulative reward vs Number of training Episodes for our proposed JMLS-DDPG HO scheme according to user type.

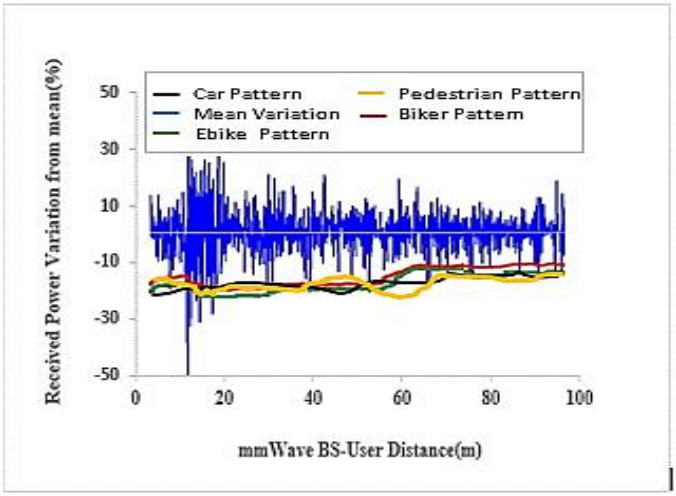

Figure 5.13: The actual average received power pattern variation at the UE in percentage about the mean value after 200 episodes in proposed JMLS-DDPG HO scheme.

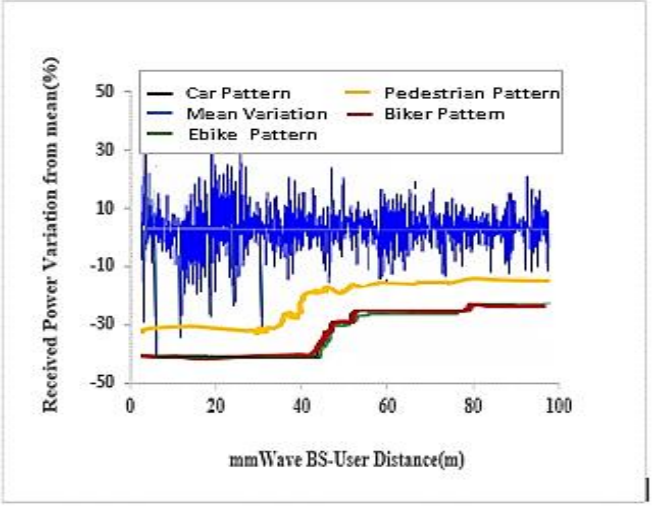

Figure 5.14: The Received power pattern variation in percentage for JMLS-DDPG HO scheme.

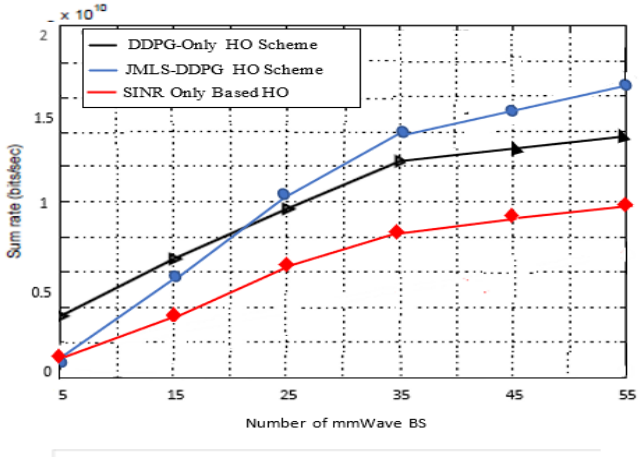

Figure 5.15: Compares sum rate for threes HO schemes vs Number of BSs.

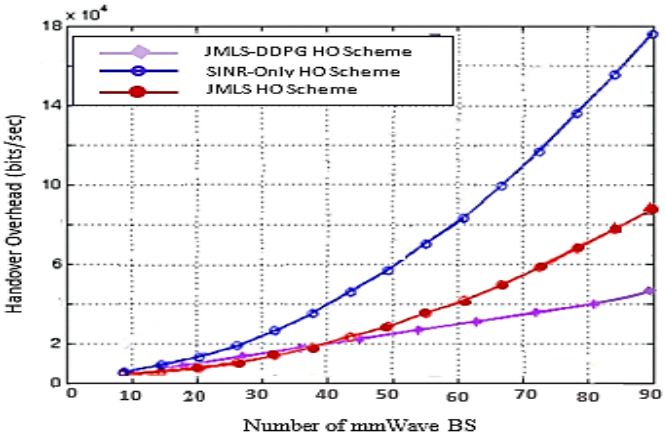

Figure 5.16: Compares overhead vs Number of mmWave BSs.

Figures 5.15 and 5.16 show the results of our comparison with a soft-HO DC-based scheme [4] that relies solely on SINR [7] and a DDPG-based scheme [6] in terms of sum rate and HO overheard respectively. Three different HO schemes are compared with respect to the sum rate versus the number of BSs in Figure 5.15. According to [4], a SINR-based scheme only takes into account the SINR of the serving cell or link and the target cell when making comparisons. Our proposed system makes use of both new and old CSI, whereas the other system gets updated with each new episode. It is clear that the proposed scheme selects/uses BSs efficiently. Between 35 and 40 BSs, the other two schemes appear to reach saturation. This is because our scheme requires few training samples and performs in-depth analysis of CSI data. When we recycle our training data, our algorithm has more time to investigate the characteristics of the connections between nodes. Furthermore, the proposed scheme cannot acquire additional knowledge about the desired link degradation pattern due to the absence of mmWave base stations. The lower sum date rate in the 5–15 mmWave BS range is indicative of this. More mmWave BSs means more data types are considered during each episode. The prediction of the target link path for the DDPG-only HO scheme is enhanced through the acquisition of new training samples in each episode, despite a very low amount of BS. However, the predictions' inaccuracy quickly becomes obvious due to the target link path's rapid variability.

The amount of overhead they produce is another factor to consider when comparing the effectiveness of various HO solutions. The three different HO proposals and their varying levels of induced overhead are shown in Figure 5.16. It is obvious that the SINR-based HO causes more handovers . This is because at each attachment to a new BS, a number of new measurement reports must be analysed in order to allocate new subcarrier resources. However, with the proposed HO scheme based on DRL, the HO overheads are reduced because the historical link data necessary for reliability is reused and exchanged in advance of the HO. Our proposed scheme benefits even more from this flexibility because of the ability to re-use meta data effectively. As a result, the proposed scheme outperforms both the DDPG-only and the SINR-HO schemes in terms of over head.

The results in Figure 5.15 and Figure 5.16 further justifies the integration of DRL and JMLS. It makes both of them more reliable in both the short and long term. JMLS due to its high inference levels when using EM scheme is poor in long term. Particularly because the predictions are

sometime exaggerated when training data is not enough, and meta data is the only source. However, its short-term predictions have proven to be more reliable as availed in chapter 4. On the other end and as shown in Figure 5.11, DDPG-only schemes have shown to want more training data to formulate reliable policies over the long term. To that effect, a combination of the two, as shown in Figure 5.12, creates a symbiotic improvement to learning and predicting the pattern of target links. Particularly, the addition of JMLS improves the DRL HO parameter predictions in the short term. This short-term prediction accuracy by JMLS is consolidated by the DRL's long term accuracy (compare Figure 11 and Figure 12 convergence rate and behavior).

5.7 Conclusion

This This chapter introduced a new HO scheme as a potential remedy for the problems caused by the unique mmWave propagation characteristics in a HetNet structure. The use of mmWave bands in conjunction with LTE bands in multi-user settings was taken into account, along with a resource allocation issue. We considered a scenario in which a downlink LTE-mmWave HetNet makes use of a strategy based on the analysis of mmWave link behaviour patterns to meet the challenges posed by HO. The resulting optimization answer considered the link's behaviour modelled with JMLS, DRL, and meta-training strategies. The best HO link was then chosen using KMO testing principles. Our HO scheme outperformed both the HO scheme based solely on DDPG and the HO scheme based solely on SINR in the simulation results. This demonstrated the usefulness of deterioration pattern analysis in choosing mmWave links for 5G networks. We can first conclude that, before carrying out HO, the characteristics of a long-term behaviour analysis for mmWave target links are assumed by our pattern-analysis HO scheme. Conventional HO schemes employ unreliable techniques, such as only considering the instantaneous behaviour of target links when deciding which one to use. On the other hand, this method takes a more methodical approach to choosing the optimal target link. It would be interesting for future works to consider the competing effects of pathloss, channel gain, and transmission power when determining the receivable deterioration pattern of the target link. You would find this to be the case if you were trying to identify the pattern. This is so because of the impact their variability has on the information transfer rate. Although highly directional beam antennas are needed at the PHY layer to achieve an acceptable link quality, it may be worthwhile to explore how to effectively manage or avoid the

unfavorable effects of both mobile and static blockages when selecting mmWave links for use in HO schemes. Research into expected patterns of behaviour for target links could focus on this area in the future. In conclusion, it would be fascinating to study backhaul configurations that can efficiently support the proposed HO scheme.

Chapter 6

The HO using Mean Field Game Theory with Jump Markov Learning

Most Reliable HO policies cannot be developed without a massive training data set, which is a prerequisite for most HO schemes using Machine learning schemes. What's more, even products with AI learning capabilities often lack the trustworthy information on individual needs needed to enhance specific connectivity requirements. Instead, all users' requirements for the network are combined into one category for policymaking purposes. For cooperative decision-making concerning resource optimization, game theory is yet another potent mathematical modelling tool. It paves the way for the creation of scalable, distributed systems that require less resources than traditional machine learning methods. Users and BSs engage in cooperative or non-cooperative interactions with one another to maximize their utility and reach equilibrium. This convenience can take the form of any playable parameter. Players in a cooperative game are able to negotiate

for their mutual benefit by working together to achieve a common goal. In contrast, in non-cooperative games, each player acts independently to maximize their own utility. The assumption of rationality is central to the study of games of strategy, which holds that each player will always make decisions that are in their own self-interest. However, in wireless networks beyond 5G, various actors may pursue divergent goals. For instance, in HO, a player whose goal is to maximize its spectral efficiency may view an opponent whose goal is to maximize its energy efficiency as being irrational (SE). The sequel expands upon this topic extensively. In contrast, Mean Field Game Theory (MFGT) is concerned with the study of differential games involving an infinite number of players whose interests and circumstances may vary widely. Due to its sensitivity to user mobility and topographic dynamics, mmWave is one example of a technology whose user population exhibits large-scale behavioral dynamics. To expect a HO scheme to reliably collect comprehensive player state data in most situations is unrealistic. Thankfully, there is no point in attempting such an impossible task because, according to mean field game theory, all you need to do is employ strategies based on the distribution of the other players. A simplification of this magnitude reduces The authors of [102, 103, 104, 105] popularized the term "DRL" in that year, and their work provides extensive documentation of the concept. The term "mean field" was inspired by mathematical physics mean field models, which study the interactions between large numbers of identical particles [122]. Mean field games get their name because here "particles" are "agents" or "players." Nash's certainty equivalence principle describes similar ideas that were developed independently and around the same time by [88, 89, 90, 91].

6.1 Background

Strategic decision making in massive, low-contact populations is the focus of mean-field game theory. Due to its growing importance in many scientific disciplines, mean-field games have been a hotspot for study in the last decade. Mean-field games are a special case of stochastic differential games in which the players' interactions are of the mean-field type, i.e., they are coupled via their empirical measure [189]. Mainly, this thesis aims to predict 'jumpy' behaviors in 5G mmWave by extending the theory of mean-field games. Since mmWave transmission is often unreliable due to its uneven coverage and erratic cell patterns, jump processing is essential for modelling and predicting sudden shifts in mmWave transmission. In particular, the HO scheme can decide

whether or not to initiate a HO based on the intensity of abrupt behaviour changes of a target link when modelling abrupt events appearing in mmWave propagation [191]. Such behaviour can be modelled using the Jump-Mode Switching (JMS) approach, which uses difference or differential equations to model the continuous dynamics within each operation mode and a Markov chain to describe the abrupt transitions between modes. Following this, mean field game theory can be used to transform the differential equations into differential games [156].

Mean field game theory takes a number of different approaches to studying differential games with an infinite number of possible outcomes and players. One approach is to study the limit of Nash equilibria in large-scale differential games as the number of players approaches infinity. Alternatively, one can try to solve differential games with finitely many players by attempting to guess the equations that Nash equilibria of infinitely many-player games should satisfy [193]. Until recently, not much was known about the first strategy. To formally derive an equation for the limit to Nash equilibria, Lions [108] explains how to do so using a nonlinear transport equation in the space of measures (the "master equation"). There is no known way to pass to the limit of the Nash system, and the question of whether or not a solution exists and is unique is an open problem for this equation. Recent advances have been made on both questions [129,151,162], and we explain some of the ideas in the second section of these notes. The MFG system, which is the primary focus of the first chapter of these notes, is based on the observation that the "characteristics" of the infinite dimensional transport equations solve, at least formally, a system coupling of a Hamilton-Jacobi equation with a Kolmogorov-Fokker-Planck equation. In fact, the feedback control provided by the solution of the mean field game system yields Nash equilibria in differential games involving a large (but finite) number of players. It was first brought to light by Huang, Caines, and Malham [89] and has since been developed in a number of other papers [95, etc.].

The JMLS [7] online learning algorithm is investigated in [13]. Non-convex problems, however, make it easy for the optimization to get stuck in a local maximum. JMLS is combined with RL and GT methods in [9] and [11], and particle-filter-based RL is used to predict perturbations and states within a random trajectory sample in [15]. In addition, the JMLS formulation is supplemented in [16] with game theory concepts to address a decentralised decision-making setting. All these ideas revolve around enhancing GT and DRL functionality. We combine game theoretic principles with JMLS ideas to speed up accurately predicted failures in 5G links and ensure HO stability.

The rest of the chapter is organized into sections as follows. Section 6.2 highlights the contributions, the proposed learning and network Model. Section 6.3 presents proposed learning model. Section 6.3 is the problem model and Section 6.4 shows the system complexity. Section 6.5 presents the performance evaluation and the simulation results obtained. Finally, a brief discussion in Section 6.6 concludes the chapter.

6.2 Contributions

- We propose a mean field GT-JMLS based HO model unlike trying to make optimal policies via a centralised system like in AI, we use a distributed HO decision making process. Helps assess the target link's ability to not just meet the desired QoS but sustain connectivity given a specific locality. Unlike global HO policies, in AI solutions, the mean field GT-JMLS based HO model, formulates local based HO policies. This makes learning faster, besides, considering, mmWaves in 5G are highly sensitive to environmental changes, the level of disruption and what might excite or disrupt different parts of the network or BS links may not be uniform. In such situations, making unique HO decisions is more effective.
- We propose a HO scheme that learns to predict not just the immediate network performance but also the deterioration pattern of a link after HO.
- We propose a scheme that trades in high power requirements for high channel gain links in meeting the desired SINR. This also improves energy efficiency. We particularly explore the effects of relying on received power and channel gain when selecting target links in HOs. The concept explores trade-offs between channel gain and power requirements in meeting the desired SINR in both NLOS and LOS scenarios. We note that mmWave links are power limited in NLOS scenarios and bandwidth-limited in LOS scenarios. Therefore, in power limited scenarios, we assess how high channel gain can be used to compensate power requirements and meet the desired SINR. This is vital for link connectivity survival in vulnerable situations e.g., NLoS scenarios. It is also vital minimizing HOs where the LoS obstruction is temporal. We study different user types and their respective effects on network performance. We study the impact of each user type behaviour on the network performance and target link deterioration pattern.

6.2.1 Proposed Learning Model

JMLS is used to make mmWave link behaviour projections based on limited observed signals. JMLS using EM estimates can carry out all possible deterioration patterns projections expected by network after a HO over different states given the observed set of signal. Doing this however is intractable in practice since the number of possible patterns increases exponentially along with time. Further, considering the vulnerability of mmWave propagation, different areas of the network will make mmWave network behaviour respond differently. In fact, [45] claims that the actual outage in mmWave propagation is environmental dependent. The idea of mean field game theory is to restrict the space a HO scheme can predict the likely target link performance deterioration pattern. Thus, our proposed HO scheme unlike, ML algorithms in AI uses a distributed HO decision process. The HO scheme doesn't need to get data from other areas to make HO decision MFGT has to find and prove that indeed the target link selected for HO will behave uniquely in a way that somewhat; meets the predicted behaviour, meets the efficacy desired in terms of energy and data. Projections made by EM are based on this initial distribution (the agent's best guess at the overall player[user] dynamics). From this initial distribution, each player (BS) can deduce their own unique HO strategy using the Hamilton-Jacobi-Bellman equation. After HO, the forward Kolmogorov equation can be used to predict the actual dynamics of the overall behaviour of how links on that BS will behave by plugging in the individual behaviors based on the distribution modelled by JMLS. The rational expectation hypothesis predicts that the HO initial distribution will match the final distribution that is predicted by JMLS modelling. This dynamic forward/backward procedure is central to the theory of the continuous-time Mean-Field Game.

6.2.2 Proposed Mobility Model

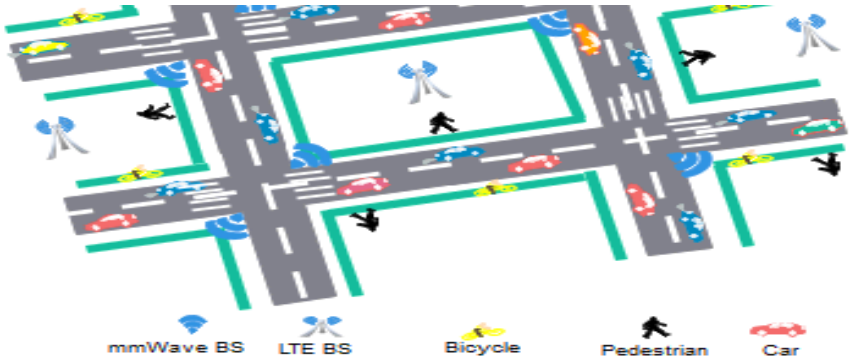

Figure 6.1: Multiuser type Mobility Model.

A Manhattan grid model [12] is used to model the road network with streets and intersections in an urban scenario as shown in Figure 6.1. The road network area is 100m x50m. We have three types of users: *pedestrian*, *cyclists*, and *cars*. A third of the total number of users are pedestrian with speeds of 1.4m/s. Another third are cyclists with velocities between 7 - 8 m/s. The other third are car users with average velocities of between 10-14m/s. Cars in the range of 3m or less to each other adjust their velocities by between 1-3m/s to avoid collision. Car velocities are updated every 3s to decrease/increase. Each street consists of the right and left lanes for each user-type. Given user directions i.e., $\mathfrak{y}$, the probability, $P_{\mathfrak{n}}$, of delay or slowing at the crossing is [13,15]:

$$\mathbb{P}_{\mathfrak{y}} = \frac{1}{k}\sum_{i=1}^{k} \frac{str(G_i) - 1}{link(G_i) + \left(P^i/T^i\,\delta_{int}v\right)}, \tag{6.1}$$

where k is the number of all possible crossings at the intersection, $str(G_i)$ is the number of crossings available after the i^{th} street is blocked or not allowed due to traffic rules, $link(G_i)$ is the number of street available to cross to after the $i-th$ number of streets is blocked. G_i represents the crossing after the $i-th$ combination of streets. T^i and P^i are the average delay time at crossing and the probability that the user will be in LOS with a serving BS, respectively. δ_{int} and v are user density and lane average velocity, respectively. Here, $\mathbb{P}_{\mathfrak{y}}$ lies in between (0,1]. It holds true that higher $\mathbb{P}_{\mathfrak{y}}$ values imply, low robustness at crossing, i.e., longer delays with link degradation.

6.3 Problem Model

6.3.1 Outage Probability

The data rate, r^m, given transmit power, P, is defined as [16]:

$$r^m = b \log_2 \left(1 + \frac{P|h^H \boldsymbol{p}|^2}{(1+ d^{\propto})} F_x\left(\left|\theta_k^l\right|\right)\right) \tag{6.2}$$

$$h^H p = \sum_{\hbar=1}^{K} g_{\hbar}(t) e^{-2\pi i f_d \cos\left(\varphi_{\hbar}^l\right)t}, \tag{6.3}$$

$$\varphi_k^l(°) = \frac{1.4\ x\ 10^4}{f_c(GHz).v(km/h)}, \tag{6.4}$$

where $\theta_k^l = \frac{2d \sin \varphi_k^l}{\lambda}$ is the normalized central angle of arrival of beam, p, with respect to a user beam given user velocity, v, under 50 Km/h and carrier frequency, f_c, as in (6.3). f_d is the maximum Doppler shift given the central $\varphi_k^l(°)$. $|h^H \boldsymbol{p}|^2$ is the channel gain [6]. $F_x\left(\left|\theta_k^l\right|\right)$ denotes the Fejér kernel value variation such that as SINR approaches maximum, as user speed approaches zero i.e., $F_x\left(\left|\theta_k^l\right|\right) \to 1$, particularly because the user beam aligns with the transmission beam. Otherwise, F_x approaches 0 as v increases [4] due to rapid variation. α is the path loss exponent [9] and their respective LOS and NLOS pathloss exponents are denoted by: $\propto^{k_L}$and $\propto^{kN_L}$, $\forall k \in \{1, \dots, K\}$ in LOS and NLOS, respectively. $g_{\hbar}(t)$ is the time-varying gain of the channel over K clusters.

6.3.2 Resource Allocation Problem

Assuming, Θ, is a set of optimization parameters from (6.2) to (6.4) with access policy, π. The outage probability, P_π, over observable signal set Y_k can be defined as [2] and [11]:

$$P_\pi(Y_k|\Theta) \triangleq P\left(\sum_l \sum_{s_k} b_l \log_2\left(1 + \hat{\gamma}_t(x)\right) \geq r_l^{m\zeta}\right), \tag{6.5}$$

where $\hat{\gamma}_t$ is the measured SINR and $r_l^{m\zeta}$ is the targeted data rate given channel state, $s_t \in S$. b_l is the bandwidth given link channel l on signal $y \epsilon Y_k$. We assume three sources of outage for a

mmWave link, and all mm-BSs directionally transmit equal maximum power P. First all users have a receiver sensitivity of x_{kmin} and a threshold x_{k0}, thus must sustain $x_{k0} > x_{kmin}$ to avoid outage. Secondly, any user-mm-BS link that requires transmit power that exceeds maximum P to meet target data rate will not be established, i.e., will experience truncation outage [16] at a distance, d. That is for LOS and NLOS scenarios, users at distances greater than, $(P/\, x_{k0})^{\frac{1}{\propto^{k_L}}}$ and $(P/\, x_{k0})^{\frac{1}{\propto^{k_{NL}}}}$, [4] from target mm-BSs are unable to communicate respectively. Thirdly outage may be due to not just meeting the desired rate, $r_l^{m\zeta}$, in (6.3**).** The minimum rate requirement problem of R_m under total power constraint and minimal outage requirement at t is defined as:

$$\mathbb{R}^m: \max_{\Theta} \sum_{t} \sum_{S_t, l} \left[\left(1 - \beta_l \left(P_\pi^{m|x_t} P_{NL_k} + P_\pi^{m|u_t} P_{NL_k}\right)\right) r_l^{m\zeta}(y) \right], \tag{6.6}$$

where $P_\pi^{m|x_t}$ and $P_\pi^{m|u_t}$ are LOS or NLOS conditional outage probabilities for a user in the m^{th} state of a user, respectively, given the mean outage probability, P_{NL_k}, at a particular transmission distance. $r_l^{m\zeta}$ is the optimal attainable rate for the $\boldsymbol{l^{th}}$ link at d. β_l is a binary factor. It is 1 or 0 for either the first or second component of (6.6) depending on whether the expected outage condition at d is given access policy π is in LOS or NLOS. The corresponding energy cost, E_c, consists of two parts: transmission energy consumption in LOS and energy consumption in NLOS as defined as [6.7,6.8]:

$$E_c = \beta \left\{ P_i \frac{c(t-w)}{\mathbb{R}^m} + e_0 * \zeta c(t-w) \right\}, \tag{6.7}$$

where P_i is the transmission power in link i, $c(t-w)$ is the total actual number of packets received at time t during window w at a distance $d_t \in s_t$ and e_0 is the unit energy per packet size in byte. $\zeta c(t-w)$ is the lost number of packets given the total expected packet over window w. β is the cost per unit energy. Thus, under the constraints of maximum transmission power, receiver sensitivity, gain and bandwidth constraints in NLOS and LOS scenarios [5], the problem is maximizing link utility with efficient and least power cost. To maximize long-term link utility of a target link, the long-term projected power, x_{t+1}, over for a minimum latency, $c(t-w)/\mathbb{R}^m$, is estimated as (6.8):

$$x_{t+1} = \max \sum_{x_t, u_t} \left\{ \frac{\gamma^{min}}{\hat{\gamma}} x_t - \frac{\alpha {u_t}^2}{\beta \hat{\gamma}^2} \right\}^+ , \tag{6.8}$$

where $\{.\}^+ = \{max, 0\}$. x_t and u_t are current LoS and it is likely discrepancy in NLoS received power over the same distance between user and BS, respectively. γ^{min} is the minimum required SINR to satisfy the desire QoS and $\hat{\gamma}$ is the measured SINR. The second term parameters, α and β, are power and SINR scaling factors, respectively, account discrepancies between NLOS and LOS scenarios for the same distance.

We utilize JMLS to model feasible optimal receivable power for long-term link utility of target links with optimal latency and minimal energy cost. JMLS is known for modelling abrupt and continuous behavior changes in failure prone systems [16] thus can incorporate abrupt effects of mmWave's NLOS and LOS switching dynamics.

6.4 System Model

6.4.1 JMLS System Definition

We first reformulate link optimization problems in (6.5) – (6.8) into JMLS learning problem:

A) The JMLS Representation

Let the JMLSs describe various values of (6.2) – (6.8) to be defined as [12,16,16]:

$$\begin{cases} x_{t+1} = A(s_t)x_t + B(s_t)u_t + g(s_t)w_t \\ y_t = r^{min}(s_t)x_{t+1} + Q(s_t)v_t, \\ \mathcal{M} = (\Theta, P(S), \pi) \end{cases} \tag{6.9}$$

where $x_t \in X$ is the current receivable power in LOS given state, s_t, with the initial value, x_0, at an initial distance, d_o; $u_t \in \mathcal{U}$ is the associated power discrepancy due to blockage/ NLOS effects; $A(s_t)$ and $B(s_t)$ are dynamic SINR/power coefficient matrices with respect to (4c), $Q(s_t)$ and $g(s_t)$ are dynamic weighted noise due to channel-gain measurement on SINR and EE(r^{min}), respectively. s_t denotes the state governing parameter set $\Theta = \{A, B, R, r^{min}, Q\}$ and belongs to Markov process at *t* as defined as:

$$s_t = \{r_t, T_t, d_t, \eta_t, v, b\}, \tag{6.10}$$

where:

$\boldsymbol{v} = [\boldsymbol{v_1}, \dots, \boldsymbol{v_T}]$	**vector of user velocity,**
$\boldsymbol{r_t} = [\boldsymbol{r_1}, \dots, \boldsymbol{r_t}]$	vector representing user data rate,
$\boldsymbol{T_t} = [\boldsymbol{t_1^m}, \dots, \boldsymbol{t_N^m}]$	vector of average service time,
$\boldsymbol{d} = [\boldsymbol{d_t^m}, \dots, \boldsymbol{d_T^m}]$	vector of transmission distance points that have matching SINR and power values,
$\boldsymbol{\eta} = [\boldsymbol{\eta_1}, \dots, \boldsymbol{\eta_N}]$	a vector of user direction in n^{th} sample,
$\boldsymbol{b} = [\boldsymbol{b_1}, \dots, \boldsymbol{b_T}]$	a vector of mmWave gNBs.

$\mathcal{r}^{min}(s_t)$ is the immediate reward following a HO to a link with receivable power, x_{t+1}, given observable signal $y_t \in Y$. $\mathcal{r}^{min}(s_t)$ is defined as a function of energy efficiency at the receiver:

$$\mathcal{r}^{min}(s_t) = \frac{r^m(s_t, P_{Los})}{P_{Los} + P_C}, \tag{6.11}$$

where $r^m(s_t, P_{Los})$ is the maximum expected data rate, $P_{Los} \leq P$ is the maximum transmission power at distance, $d \in s_t$ for a target mm-BS given likely user-receivable-power x_{t+1}. P_C is the circuit power consumption. The transition probability function between different states is given by:

$$P(S) \triangleq P\big(s_{t+1} = m_j | s_t = m_i\big), \tag{6.11}$$

The normalized transmission energy cost index, $\mathcal{J}(xt)$, over N sampled distances/states from BSs is defined by:

$$\mathcal{J}(x_t) = \boldsymbol{E}\big\{\sum_{j=1}^{N} \|x_t\|^2_{Q(s_t)} + \sum_{j=0}^{N-1} \|u_t\|^2_{R(s_t)}\big\}, \tag{6.12}$$

where $\|x_t\|^2_{Q(s_t)} > 0$ and $\|u_t\|^2_{R(s_t)} > 0$ represent weighted norm energy costs for actual received packets and lost packets (e.g., due to NLOS effects), respectively, given the expected number of packets over window, w. $J(x_t) \triangleq \varepsilon E_c$, where ε is normalization factor of the energy cost, E_c, in (6.7).

B) JMLS Training

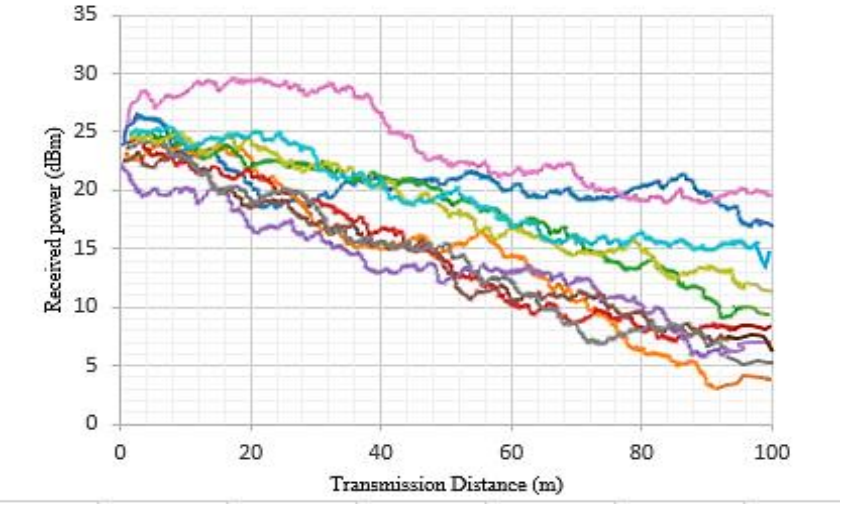

Figure 6.2: A general overview of notable received power deterioration pattern over distance from different user type when using EM projections.

Given Y_T, $X_T$and S_T denote patterns of observed channels $\{y_1, \dots, y_T\}$, corresponding receivable power pattern $\{x_0, \dots x_T\}$, and states $\{s_1, \dots, s_T\}$ until time T, respectively. For a finite set of observed signals, Y_T, over S_T, the model learning problem is to predict the best pattern, X_T, over parameter set Θ for a finite distance $\{d\} \in S_T$ that will satisfy the desired QoS, e.g., SINR taking into consideration latency and energy efficiency effects. For a given channel, y_t, the estimated received power is:

$$x_t(y_t) \triangleq \arg \max_{x_t \in X_T} P(y_t, s_t | x_t, \Theta), \tag{6.13}$$

where $P(y_t, s_t | x_t, \Theta)$ is the conditional distribution of y_t if x_t is observed in state s_t. Here given initial received power prior to HOs, the initial values of Θ given $x_t \in X_T$, are estimated using Expectation-Maximization (EM) where we estimate the likely observable signal given the receive power as x_{t+1}, i.e., $P(Y_T | \Theta, X_T)$ [12]. EM uses Bayesian inference to automatically infer initial unknown values of parameters, Θ, in (6.11) derived using complete-data log-likelihood function, $Q(\Theta | \Theta^k)$, [12] defined by**:**

$$Q(\Theta | \Theta^k) = \mathbb{E}[\log P(X_{T-1}, S_T, Y_T | \Theta) | X_k, \Theta^k], \tag{6.14}$$

where Θ^k is the current known parameters estimate at the, $\boldsymbol{k}^{\mathbf{th}}$, iteration. The values are optimized based on the updates of the actual SINR changes:

$$\Theta^k = \arg\max_{A\in\Theta} \mathcal{Q}(\Theta|\Theta^k), \tag{6.15}$$

and correspond to energy cost in (6.7) as

$$\log P(X_T, S_T, Y_T|X_k, \Theta^k) \triangleq \sum_{k=1}^{k} \|x_k\|^2_{\mathcal{Q}(s_k)} \tag{6.16}$$

Figure 6.2 shows the possible link power deterioration patterns that meet the SINR threshold as different user types move away from their serving BSs. Using EM estimates in eqn. (6.15) and (6.17) - (6.20), the learning space over which the target link power deterioration pattern is determined over distance is noted. The received power pattern over distance need to be enough to meet the desired SINR and avert outage in all scenarios. In scenarios where the LoS is not clear, we look at the extent to which channel gain influence the SINR value per state. We consider three possible values and choose the target link with the least energy cost hence best EE. The three factors considered are:

(1) the Energy cost in LOS given maximum received power is received for maximum SINR value $a \in A$:

$$\mathcal{J}(x_t, a) = \max_{(x,a)} \boldsymbol{E}\left\{\sum_{0\,j=1}^{N} \|x_t\|^2_{\mathcal{Q}(s_t)}\right\}; \tag{6.17}$$

(2) the cost in NLOS given channel gain, h, compensates limits in received power for max-SINR value $a \in A$ (power limited);

$$\mathcal{J}(h_t) = \max_{(h,a)} \boldsymbol{E}\left\{\sum_{j=0}^{N-1} \|u_t\|^2_{R(s_t)}\right\}. \tag{6.18}$$

(3) the minimal cost based on a balanced gain in h and x_t;

$$\mathcal{J}(h_t, x_t) = \min \boldsymbol{E}\{\alpha\mathcal{J}(x_t)\} + \min \boldsymbol{E}\{\beta\mathcal{J}(h_t)\}\,. \tag{6.19}$$

Considering $c(\mathcal{J}(h_t), \mathcal{J}(x_t, a), \mathcal{J}(h_t, x_t))$ is the ultimate cost of the three, i.e., a convex function of $\mathcal{J}(h_t)$. For the agent, in state s_t the learning problem for desired x_t, becomes (6.20):

$$x_t(y_t) \triangleq \text{argmin}\, c\big(\mathcal{J}(h_t), \mathcal{J}(x_t), \mathcal{J}(h_t, x_t)\big) + \text{argmax}\, r^{min}(s_i). \tag{6.20}$$

We use mean field game theory [8] to know the best energy cost over long-term connectivity. With mean field GT, we don't need full analysis of all possible optimal power deterioration patten in Figure 6.2, we just need to know the distribution of similar energy cost experience. Thus assuming $\mathcal{J}(h_t, x_t)$ from (6.17) to (6.20) is the mean cost for each state, we exhaust it as the cost summary in each state for different user types and use it as input to mean field games trading off power, x_t, against channel gain,$\boldsymbol{x_t}$, to determine the most optimal target link SINR behaviour in the follow up sections.

6.4.2 Mean Field Game Concept and Consideration

Let the conditional PDF, $P(x_t, s_t|y_t, \Theta)$, with respect to EM be defined as [10]

$$P(x_t, s_t|y_t, \Theta) = \frac{P(y_t, s_t|x_t, \Theta).P(y_t, \Theta)}{m^l}, \tag{6.21}$$

such that the expected changed in x, given $\mathcal{J}(h_t, x_t)$, with respect to first order conditional PDF is [13]:

$$\frac{dx}{d\check{x}} = (\propto + \beta)P'(x_t, s_t|y_t, \Theta)\left(\frac{x - \check{x}}{\sigma^i}\right). \tag{6.22}$$

It must be emphasized that the quantity, which is needed having observed Y_T, is the conditional PDF $P(x_t, s_t|y_t, \Theta)$ given *X*, and not conditional distribution $P(y_t, s_t|x_t, \Theta)$ given Y, used earlier to estimate the needed received power pattern. However, considering the rapid variation, and sensitivity of mmWaves, obtaining accurate/full CSI, i.e., the conditional PDF is not practical. Besides, the EM estimations are subliminally [6] accurate. We use game-based mean-field approximation concepts with value functions in [3, 4] to provide one of the simplest ways to determine a more accurate pattern considering the rapid variation, and sensitivity of mmWaves.

A) The Game Value Function

Let $\hat{a}$ be the SINR value under consideration and $*\hat{a}$ denote the SINR estimates by EM for $x \in X$. We define the expected power value, x_t, as a variable of the immediate reward $r^{min}(*a, s_t, a)$ given the mean field value function,$V(x_t)$:

$$V(x_t) = \max_{\pi} E\left\{\sum_{k=0}^{K} \delta^k r^{min}(*a, s_k, a_k)|_{\pi(x_t|\theta\pi)}\right\}, \tag{6.23}$$

with policy, π, at each state.

$$\pi = \arg \max_{a_0,\dots a_k}\left[r^{min}(*a, s_k, a) + \sum_{s} P(s_{k+1}|s_k, a)\, V(s_{t+1})\right], \tag{6.24}$$

where $\delta \in (0, 1)$ is the discount factor and moves to closer to 1 as energy cost reduces. The long-term value function, $V(\check{x}_k)$, is approximated as:

$$V(\check{x}_k) = \sum_{a_0,\dots a_k} P_\pi(x_k|y_t, \theta(a_k))\left\{r^{min}(*a, s_k, a_k) + \delta^k \sum_{x} P(s_{k+1}|s_k, a, x)\, V(s_{t+1})\right\}, \tag{6.25}$$

where $P_\pi(x_k|a_k)$ is the probability, given policy, π, of obtaining a SINR value ,$*a$, (determined by EM) if ,x_k, power is received at state s_k. The mean field game then looks for a target link whose SINR value, a, satisfies condition $a \geq * a$ over observable signal, y_k with less than, x_k, power at state ,s_k. The discount factor, δ^k, is a tolerate factor to let the sum of rewards over S_K states, converge about the target link with the least energy cost values over S_K states.

B) Deterioration Path Projection Enhancement

To find the optimal power pattern over *k* states, we let $g(x)$: $\mathbb{R}^d \rightarrow \mathbb{R}$ represent the terminating cost factor and $r(x, a, *a)$: $\mathbb{R}^d x\, \mathbb{R}^m \rightarrow \mathbb{R}$ denote its running reward given the function value,$V(\check{x}_k)$, in each state over set, Y_K.

$$V(\check{x}_k) - \max_{a\in A} E\left\{\int_0^K \delta^k r(x, a, *a)ds(t) + \delta^k g(x)|Y_k\right\} = 0, \tag{6.26}$$

where the terminal energy cost factor, $g(x)$, corresponds to the outage costs, $c(*)$, [9] in (6.17) - (6.19) and the received power fail to satisfy receiver sensitivity threshold condition, $x_{k0} > x_{kmin}$,

or requires more transmission power, P, than maximum transmissible to meet x_{k0}, (see (6.4)-(6.8)).

To satisfy (6.26), hence identify the most probable received power pattern X, in Figure 2, the corresponding values of cumulative running reward up to state s_k must be greater than or equal to the values of the link terminating energy cost plus its long-term value function, $V(\check{x}_k)$ in (6.26). Thus, for a defined finite state space S , to calculate , the received power pattern $X = \{x_0, \dots x_k\}$ that satisfy SINR values $A = \{a_0, \dots a_k\}$, (generated by value function $Q(\Theta|\Theta^k)$ in EM) , the pattern $X = \{x_0, \dots x_k\}$ is calculated using the Hamilton–Jacobi–Bellman equation in (6.27):

$$x \triangleq V(s_t) + \sup_{hEU} \mathcal{H}(x, h^H p), \tag{6.27}$$

where the Hamiltonian function, $\mathcal{H}(x, h^H p)$, is defined by a Hamilton principle [9]. The principle states that the true evolution of points in $X(t)$, i.e., between x_1 and x_2, is described by *K* states $\{V_1(s_t), \dots, V_k(s_t)\}$ where $a \in A(t)$ is a stationary value (a SINR point of x where the variation of SINR is zero):

$$\mathcal{H}(x, h^H p, a) = 0. \tag{6.28}$$

This is achieved by ensuring any received power value change between, states i.e., $dx/d\check{x}$ is compensated by a feasible channel gain values, $h^H p$, to satisfy desired SINR. Even when the optimal received power given feasible channel gain is unattainable among the given target samples, the Hamilton Jacobi Bellman equation [12] for selected value *x* must satisfy eqn. (6.29) conditions in the kth state for the serving cell to sustain connectivity.

$$max \sum_{x \in X} V(s_t) \geq \mathbb{E}\left\{ \int_0^K \delta^k r^{min}(* a, s_k, a_k)|_{\pi(x_t|\theta\pi)} dk \right\}. \tag{6.29}$$

$$V(\check{x}_k) \geq V(x_t)$$

The sections below explain the game's innovation function to satisfy the conditions above.

C) Mean Game Innovation Function

Unlike Bayesian EM approximations, the interest of this variant is to show how coupled upward/downward the pattern of a target link might vary, which is the core of mean field game theory (in continuous time, with a continuous state space) [56]. Note that (6.29) is deterministic. The set of possible SINR values $A \subset \mathbb{R}^m$ is a subset of EM space described in Figure 6.2 describing X behaviour. The best pattern, X, with respect to HJB in (6.28), is assumed to be within the maximum and minimum of values in Figure 6.2, and not any other random value. Let (6.30) be the innovation value function of the path as

$$F_{,i\to j} = \max_{a\in A}\left[V_{ij}x_i x_j + H(\mathcal{u}, x_i, t)\right], \qquad (6.30)$$

where the Hamilton function $H(\mathcal{u}, x_i, t)$ is influenced by channel gain H value at the receiver for the corresponding energy-cost, $\mathcal{J}_{ij}$ (with efficiency V_{ij}), at the transmitter. Here the best values between s_{i+1} given s_i are decided by a Non-Cooperative Non-Atomic Game to optimize expected SINR at $x_j \in X_T$ as stated in the next section.

D) Non-Cooperative Non-Atomic Game Field

Given innovation values as components, Γ, of the game,

$$\Gamma < \{\mathcal{J}_{ij}\}, \{S_{ij}\}, H(\mathcal{u}, x_i, t) >, \qquad (6.31)$$

a non-cooperative game is played as follows. The value set of $\{\mathcal{V}_{ij}\}$, whose received power meet $x > x_{ko}$ condition but does not support any SINR above the threshold, are thought to be in NLOS. To thus determine optimal values in (6.31) for Γ , hence, the target link with optimal X pattern, the predicted target link values must satisfy $x > x_{ko}$ and high channel gain, *H*, conditions between s_i and s_j states. Correspondingly the energy cost following $s_i \rightarrow s_j$ transition must be efficiently low during transition but still meet give the desired SINR in every scenario. Technically, it implies, for the chosen, BS for HO, the BS-user target link will not need more transmit power, *P*, than maximum but will require high channel gain in NLOS scenarios to meet the desired SINR given the received power will be low. To quickly make a decision, the non-cooperative game, initially selected target links-based EM estimates in Figure 6.2. It then updates by the game values in (6.31) to at least satisfy $x_j > x_{kmin}$.

E) The Pure Nash Equilibrium

A Nash Equilibrium is a set of received power $\{\bar{x}_i, \dots, \bar{x}_N\}$ over N states $(\bar{s}_1, \dots, \bar{s}_N)$ such that it is "costly", i.e., in terms of energy, for a player/user i to deviate from receiving $\{\bar{x}_i, \dots, \bar{x}_N\}$ over $(\bar{s}_1, \dots, \bar{s}_N)$ to meet the threshold of the desired SINR over N states. It is not always easy to get pure Nash equilibria values for Target link selection. To that effect, we relax the innovation functions in (6.30) and Pure Nash Equilibria value conditions that a non-cooperative mean field game must satisfy to be selected for HO. We use mixed Nash equilibria values with a less strict innovation function as elaborated in the following sections.

F) The Mixed Nash Equilibrium

Nash equilibria in pure strategies do not necessarily exist. We must introduce the notion of mixed values. This means that each player anticipates a family of received power values with a certain probability, $\mathcal{P}(\mathcal{V}_i)$ of not meeting the desired SINR. Particularly, given $a^d \in \tilde{A}$ is the desired SINR value at state s. The sufficient and necessary condition for a^d given corresponding EM value $*$ needs be a symmetric Nash Equilibrium is:

$$r^{min}\left(* a^d, s_k, a^d(s)\right) + \delta^k \sum_x P(s_{k+1}|s_k, * a^d, a^d(s), x)\, V(s_{k+1}) \geq r^{min}\left(* a, s_k, a(s)\right) + \delta^k \sum_x P(s_{k+1}|s_k, * a^d, a^d(s), x)\, V(s_{k+1}). \quad (6.32)$$

By solving the above conditions, we find the constraints to the JMLS model parameters, Θ, desired for corresponding value x to the SINR a^d to be a symmetrical Nash Equilibrium. The mean-field approximation as the function, F^*, of the innovation over all time $t = 1, \dots, T$ is defined by (6.33) [13]:

$$F^* = \max_{\mu} \left\{ \sum_{i,j} V_{i,j} \mathbb{E}_{\mu}[X_i X_j] + \sum\nolimits_i H\left(\frac{x_{i+1}}{2}\right) \right\}, \quad (6.33)$$

where μ is the probability that the user value x_i at distance d and can be calculated based on the partial costs as:

$$\mu_{d,i}(t) = \frac{1}{1 + \exp\left(\sum_{i \neq j} V_{d,x_i} - V_{d,x_j}\right)}, \tag{6.34}$$

The HJB equation hence corresponding best pattern X_T is solved backwards in time, starting from a state with $t = T$ and ending at t = 0 [12] given **S**, for **X**.

G) Global and Local Deterioration Pattern

Global deterioration patterns are created by the aggregative collection of CSIs from different mmWave eNBs in the network. Local deterioration pattern is created with an aggregative collection of CSIs from local user experiences at individual target mmWave eNBs. A global deterioration pattern instance using CSI from various BSs. At HO, the divergence of individual target link pattern is compared to global optimal pattern. The higher the divergence, the less reliable the target link and connection is. We use the Kullback-Leibler (KL) divergence to measure the level of unreliability of a target cell and is explained in detail below.

H) Deterioration Pattern Analysis

We use the Kullback-Leibler (KL) divergence to measure the divergence of local pattern X, function values, $V_{Los}(\check{s}_t)$ and $V_{NLOS}(\check{s}_t)$, in LOS and NLOS, respectively, from the global path values $V_i(\check{s}_t)$. Thus, KL measures how either $x_{Los}(\check{s}_{i+t})$ or $x_{NLOS}(\check{s}_t)$ with reference to their respective value functions, $V_{Los}(s_i)$ and $V_{NLOS}(s_i)$, will deviate from desired $x(\check{s}_t) \in V_i(\check{s}_t)\ \forall\ \bar{X}$. The KL equations and three conditions used for prediction of target link behaviour are:

$$KL\big(V(s_t)||V(s_{t+1})\big) = \mathbb{E}\left[\log\left(\frac{V(s_t)}{V(s_{t+1})}\right)\right]. \tag{6.36}$$

1) Using forward KL [12], the difference between $V(\check{x}_k)$ and $V_{Los}(s_{i+1})/V_{NLos}(s_{i+1})$ is weighted by $V(\check{x}_k)$, If, for instance, the reward is zero, i.e., $V(s_t) = 0$ then the current global pattern $\bar{X}$ needs to be updated to know which target cell would deteriorate optimally. In other words, $\bar{x}_t$ as minimal receivable power is not desirable and will not satisfy $x_t > x_{ko}$ to avert outage.

2) Conversely, if $V(s_t) > 0$, then the $\log\left(\frac{V(s_t)}{V(s_{t+1})}\right)$ term values will contribute to deterioration pattern $(\bar{x}_1, \ldots, \bar{x}_N)$ updates using pattern $(x_1, \ldots, x_N)$. If the divergence is high, this is not good because our objective is to minimize KL divergence with discrepancy between global value,$\bar{x}_t$,

and target cell local value x_t. We measure $V(s_i)$ with either $V_{Los}(s_{i+1})$ or $V_{NLOS}(s_{i+1})$ and whichever target link gives minimal KL divergence value will cause less abrupt changes in received power requirement. It also implies the target link is able to get the desired SINR over different states by balancing the channel gain and receivable power deficits within each transmission range.

3) we use Reverse KL to assess the target cells meeting situation (1) and (2). The target cell whose KL difference between (1) and (2) lower than other cells is more reliable. A higher reverse KL implies a wide divergence on how a target link is likely to perform after HO. This also indicates to how much the likely target link predicted desired pattern $x_1, \ldots, x_N$ will likely not be meet given the optimal desirable deterioration pattern $(\bar{x}_1, \ldots, \bar{x}_N)$. This form of KL Divergence is known as *zero forcing*, as it forces the difference between global $V(s_t)$ and local function values $V(s_t)$ to approach 0 on some areas where the link of a target link is more reliable and stable.

I) Online Update of The Proposed Model

They are two ways in which the deterioration experience is updated. We consider two forms of link deterioration experience; 1) locally on individual mm-BS and 2) a global deterioration pattern, i.e., aggregative experience of distribute mm-BS. In this approach, the overall experience is independent of the number of mm-BSs and may be scaled as desired at the cost of additional communication overhead. To find X_T over Y_T, the online optimization process is summarized in Algorithm I. It follows the classical EM algorithms [7] and alternates it with game theory at each step. Given the current iteration, the new path, X_t, uses the previous pattern, X_{t-1}, as a warm restart. The coefficients computed during the previous iterations for X_{t-1} are used as aggregate information for X_t. The information from past coefficients is thus carried forward in matrices as initial input to the later parameters. For instance, the SINR values, $\hat{\gamma}_1, \hat{\gamma}_2, \ldots, \hat{\gamma}_t$, are carried forward in matrices as in step (13) of algorithm I as (6.37):

$$A_t \leftarrow A_{t-1} + \gamma^{min}\hat{\gamma} \text{ and } B_t \leftarrow B_{t-1} + \hat{\gamma}u_t. \qquad (6.37)$$

This enables updating paths based on past information without accessing the past training samples again. The projected path, X_T, can then be optimized by using these matrices and the previous **global values,** X_{T-1}, as input to X_T. This optimization strategy leads to faster convergence

performance and better paths than classical batch algorithms, scaling up gracefully to large data sets as more training samples build up. Algorithm 6.1 summarize the online training process with step 1-6 summarizing the EM process if estimating receivable power, and step 6-14 improving EM power pattern estimates on target links using game theory. Step 15-17 introduce KL divergence measure and selection of target links using the GT and JMLS estimates.

Algorithm 6.1: Online Deterioration Pattern Learning Algorithm

Input: A set of signals $Y = \{y_1, y_2, \dots, y_t\} \in Q$, T(number of iterations),user velocities v_1, v_2, v_3

Output: An overcomplete deterioration path $X \in Q(\Theta|\Theta^k)$using EM and then optimize X pattern with game theory

1. Initialize the path predictions X_o with K randomly selected training signals, Θ, where $A_0 \leftarrow 0$, $B \leftarrow 0$
2. for t=1 to T do
3. EM Inference: $Q(\Theta|\Theta^k)$
4. $\Theta^k = \arg\max_{\Theta} Q(\Theta|\Theta^k)$,
5. Estimate user x_t for each y_t to initially determine X_T
6. Infer: X_T using EM estimator

Pattern update: Non-Cooperative Non-Atomic Game:

7. Draw y_t to improve $Q(\Theta|\Theta^k)$ estimates for X_T
8. If $(\mathcal{H}(x, h^H p, a) = 0)$
9. else
10. $V(s_t) \leq V^*(s_t)$, in (14)
11. for each if do
12. $F^* = \max_{\mu} \left\{ \sum_{i,j} \mathcal{J}_{i,j} \mathbb{E}_{\mu}[X_i X_j] + \sum_i H\left(\frac{x_{i+1}}{2}\right) \right\}$
13. $A_t \leftarrow A_{t-1} + \gamma^{min}\hat{\gamma}$ and $B_t \leftarrow B_{t-1} + \hat{\gamma} u_t$
14. Update:X_T
15. use forward $KL(V(s_t)||V(s_{t+1}))$ to update pattern
16. Optimize pattern:X_T for local and global
17. Use Reverse $KL(V(s_t)||V(s_{t+1}))$ at HO

6.5 Performance Evaluation

In this Section we describe the system model and simulation platforms used to access our results.

6.5.1 Description of the Simulation Model

We evaluated the performance of our proposed HO scheme. We assume that the state information, e.g., distance, speed etc., can be known by any gNB in the network using a location-based service such as GLS [4] or HLS [5]. For our simulation we use the DC LTE-mmWave model introduced by the NYU and the University of Padova [3] where LTE BSs manage mm-BS. The model carefully considers an end-to-end mmWave cellular network performance. It uses ns-3 simulator featuring 3GPP channel model for frequencies above 6 GHz and a 3GPP-like cellular protocol stack [8]. A summary of simulation parameters is given in Table I. The initial stages of our model determine possible received power deterioration patterns for different users. We use the EM estimator on JMLS. The later stages involve optimizing the initial pattern using game theory. At HO selection stage, i.e., to choose the best target cell, individual cell deterioration patterns with respect to user type are compared to the global JMLS-mean field GT pattern optimized. We use reverse KL divergence [12] to understand the margin of divergence of the target cell pattern from that of the global optimized pattern. The smaller the divergence the more reliable the target cell is at and post HO. The learning algorithm is online and is developed on an augmented ns-3-OpenAI Gym [10] toolkit as summarized in Algorithm I. Open AI Gym is integrable with ns-3 and supports the teaching agents for the variety of applications.

Table 6.1: Simulation Parameter.

Parameter	**Value**
mmWave	28GHz
mmWave bandwidth	1GHz
3GPP Channel Scenario	Urban Micro, Urban Macro
MMWave max outage	-5dB
mmWave transmission Power	46dBm
mmWave max PHY Rate	3.2Gbps
X2 link latency	1ms
S1 link latency	10ms
RLC buffer Size	5MB
S1 MME link latency	10ms
User speed	[1,50] m/s
UDP Source rate	200Mbits/sec

6.5.2 Implementation Of the Simulation Model

Dynamic system level simulations are performed to evaluate performance. Figure 6.3 shows the flow chart of HO execution. The distribution of received power with respect to different states of the target link is predicted using EM estimations given a few initial values obtained from user

connection and expert data values. Initial optimal received power patterns are then generated. The patterns are classified into two sets. One set is assumed to be influenced by high gain (NLoS) and while the other set is influenced by high receivable power values (LoS) to achieve the same SINR. In the meantime, the total reward, which is the energy efficiency (EE) of each pattern state is also calculated and the following the steps used in Algorithm one and the flow chat in Figure 6.3.

Step 1: Using EM, the scheme estimates receivable power in LOS and NLOS.

Step 2: For the received power values in LOS set, the objective is to find ones that have least energy cost. A local optimization strategy set is obtained in this step by Hamilton–Jacobi–Bellman function conditions in (6.27) - (6.32).

Step 3: Received values from the partially NLOS user set and those from the LOS set are involved in a game. The user's received power values in the NLOS set(s) are assumed to low but compensable with high channel gain to obtain the same SINR value(s) obtained in LOS. Users with improved or degraded received power patterns given different states are moved to or from the NLOS set, while game values in the LOS sets are improved per state and user failing to meet desired received power values are moved to NLOS set. The scheme Skips to Step 4, if none of the parameters for users with regard to network improved into the LOS value set. Otherwise, if there is a new value for a user state, we get back to Step 2**.**

Step 4: Repeat Step 3 until the shifting/HO condition for each state is met, i.e., attainable or empirical minimum receivable power pattern. For the best target link, each user's received power about the mean at different states (distances) is not supposed to drop beyond user type pattern values to avoid link failure or outage.

Step 5: Calculate the total reward (Energy Efficiency-EE) according to the final energy cost and received power pattern. The power pattern corresponds to the highest EE values defining the pattern X at each state. The least considered viable EE values for NLOs scenarios in Figure 6.3, correspond to states where a user's received power is just above receiver sensitivity threshold. This is the worst-case scenario for a viable EE link in NLoS scenario. Beyond the receiver sensitivity threshold, the target link is considered to have states with higher outage probability.

The flow chart in 6.3 summaries the above steps where chart's first dashed box in Figure 6.3 is the LOS SINR value set and in the second box, it is the game's SINR parameters optimized using the

Hamilton–Jacobi–Bellman solution in (6.28). The worst case scenario for NLoS is when the received power is just above the receiver sensitivity threshold.

Figure 6.3: The flow chart of received power pattern estimations.

6.5.3 Simulation Results and Analysis

Wasteful HOs (or repeated HOs) refer to unnecessary handovers to the same serving BS. This is because the same serving BS triggers/refreshes a HO process instead of just maintaining the existing link or the system selecting a better target link that would support longer connectivity. An increase in the number of wasteful HOs turn to increase the system overhead, particularly the HO overhead. This because control plane data ends up using the bandwidth that would otherwise been used for user plane data. To that effect, we investigated the percentage of wasteful HOs over a period of time of training our proposed scheme. The results are shown in Figure 6.4. and our observation is that the percentage of wasteful/repeated HOs using our proposed HO scheme decreases over time. In fact, after the 70^{th} iteration, the training converges because the HO link selection is optimized. This implies that any link selection/HO process initiated after the 70^{th} iteration has less than 0.5% chance of being wasteful/being repeated. This performance is

attributed to the reliability of the proposed HO scheme to accurately predicted the behavioural pattern of mmWave links and ultimately select the most stable target links.

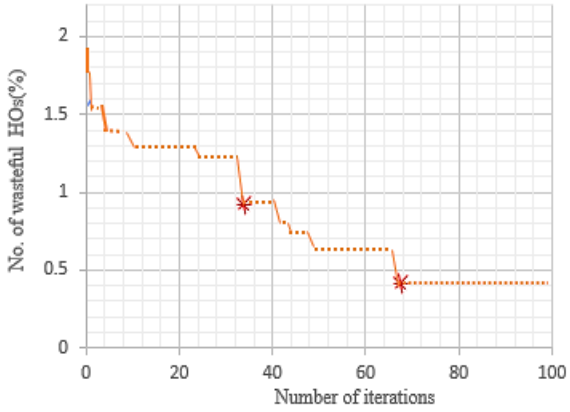

Figure 6.4: Wasteful/repeated HO vs. number of iterations.

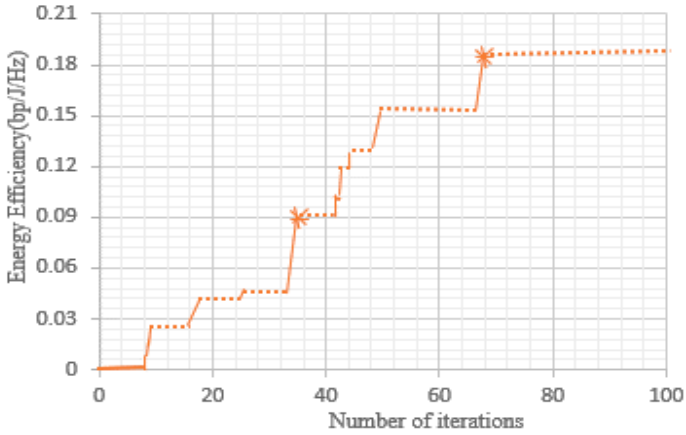

Figure 6.5: Energy Efficiency vs Number of iteration.

Figure 6.5 shows the energy efficiency (EE) in terms of bits received over the amount of energy transmitted. The EE increases over time, also reduces the energy cost. The EE increase also corresponds to the reduction in wasteful HOs over time as shown in Fig 6.4 For instance, steep increase in EE observed at around 40th episodes in Figure 6.5, correspond to reduction in wasteful HO observed around the 40th episode.

The mark, "*", in Figure 6.4 correspond to sharp changes in the learning path. This shows points of drastic learning improvement, and points to how the proposed scheme is able to make sufficient improvements with minimal training data, e.g., within less than 40 training episodes. Particularly, our proposed JMLS-MFGT technique has the ability to learn fast and select more LOS target links

(less wasteful HO) than NLOS links once a considerable small amount of training data samples/iteration have been accumulated.

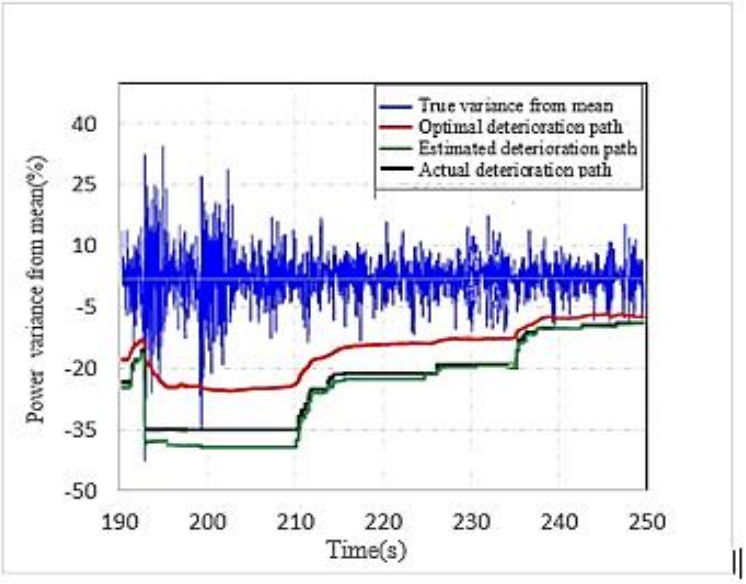

Figure 6.6: Expected received power drop about the mean (%) vs distance to meet desirable SINR.

The Figure 6.6 shows how the receivable power must vary over distance to ensure the user-BS link meets the desired SINR. The true variance(blue) lines shown in percentage (%) terms, indicate received power of a user from a BS and that they can have a variance of 5% about the mean and still meet the desired SINR. And the red line indicates the lowest allowable decrease, for a user-BS target link to still meet the desired SINR. These are pure Nash equilibria values. However, in the HO scheme, as earlier alluded, we acknowledge that the decrease in received power can be compensated by high channel gain to attain the same SINR. The green line is thus lower compensatory power limit (in green line) needed to meet the desirable SINR. Therefore, our scheme set a much lower compensatory power limit (in green line) needed to meet the desirable SINR using a mixed equilibrium approximation. This (green line) thus dictates the least receivable power decrease a user-BS link must have to meet the minimum desired SINR. Based on the black line results from our simulation, the simulation showed that the actual lowest user-target link power variations allowable (in black line) over distance for a reliable target link was indeed within the HO scheme's estimations. Therefore, any target link that satisfies at least the power variation pattern of the line in green had the capability of sustaining connectivity over given distance after HO.

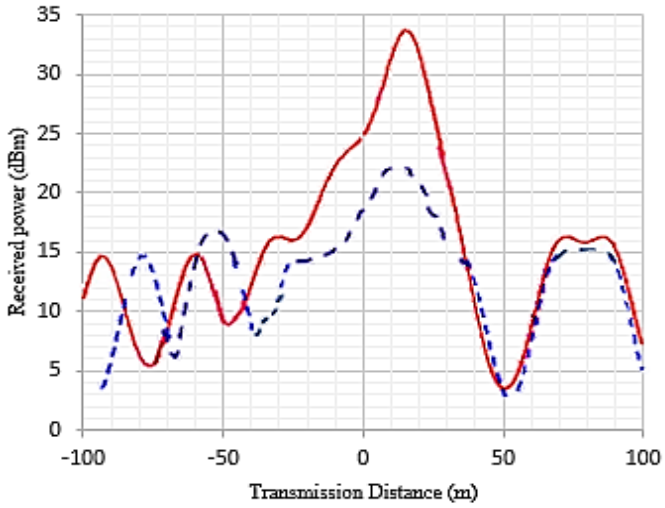

Figure 6.7: Targeted received power profile (in solid red) vs actual target cell power behaviour (in dotted blue line).

Figure 6.7 shows the user received power from the target cell vs. the user distance from the target cell in which a user is moving towards (-) and from (+) the mm-BS within a stretch of 100m. The dotted blue line shows the actual power a user received over a 100m coverage of a mm-BS. The solid red line shows the predicted power by the scheme a user will receive from a selected target mm-BS and meet the desired SINR over the transmission range. This closeness between the predicted and actual received power pattern proves that our proposed HO scheme can accurately predict the likely behaviour of the target link after a HO is executed. However, the imprecision predicted received power pattern especially when the user is beyond 50m from the mm-BS can attributed to the interplay power and channel gain have in determining the SINR (see (2)). Additionally, we analyzed each user-type data rate over the transmission distance and compared their results.

However, we acknowledge that at locations >50 m from BSs, as can be seen in Figure 6.7. the user experienced lower signal strength and rapid variations hence possibility of outages was high. This is expected with mmWave, particularly because outage is highly dynamic and environmentally dependent, one cannot, generalize too much from power variation measurements. Thus, equipped with limits of our initial selection criteria fact, we needed a HO decision that simultaneously predicts channel gain and received Power's efficiency (EE) variation over distance for HO selections. We used the KL divergence. We set the receiver sensitivity threshold to 5dBm and used KL divergence to predict (EE) distribution. The KL divergence values over the expected received power pattern X values determines the extent to which the target link's SINR values will deviate from desired. Figure 6.8 and Figure 6.9 shows KL divergence value distribution for two target cells plotted within a 100m range for different types of users going towards (-) and away (+) from the mm-BS. In Figure 8, the KL hence EE distribution across the 100m range is greater than the desired threshold for pedestrians and bicycles going towards (-) and away (+) from the mm-BS.

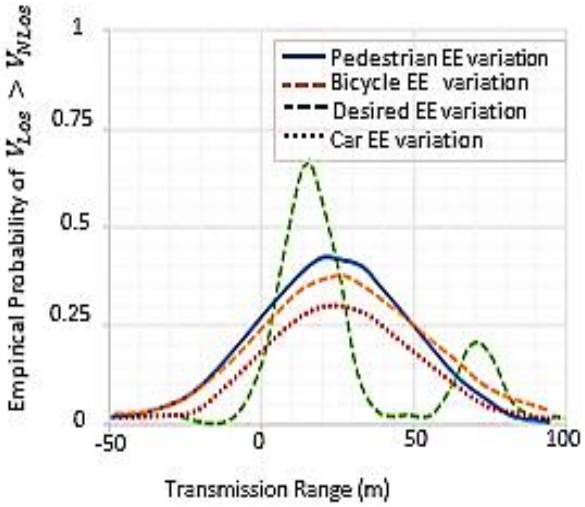

Figure 6.8: KL divergence values for user moving towards and away from a target cell.

In Figure 6.8, the target link's likely EE distribution after the HO is likely to be greater than the desired threshold, at least for pedestrian and cycler within the 50m transmission range. The consequence for this scenario is that the KL Divergence from the threshold would be big for a cycler and pedestrian user. The HO selection optimization algorithm using reverse KL would however need to select a different target link BS for a car user to sustain connectivity.

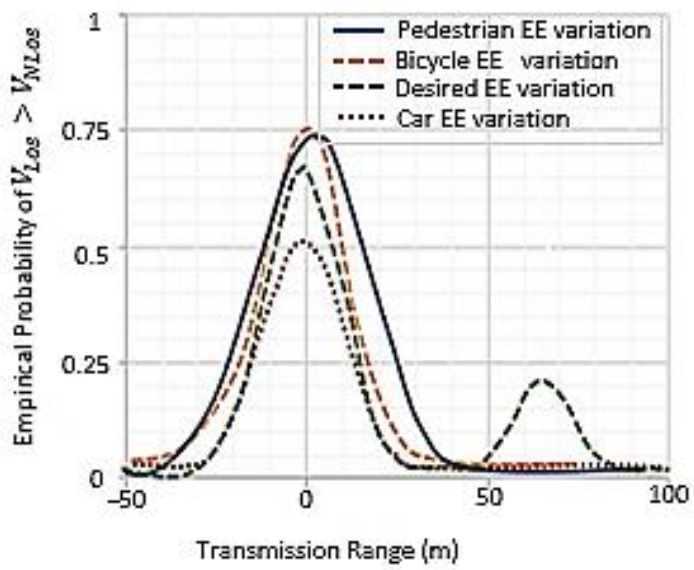

Figure 6.9: KL divergence values for users moving towards (-50) and away (50) a target cell.

For Figure 6.9., the target link EE distribution after HO is predicted to be more spread out, covering most of the 100m range. Although there are still some areas that are not fully covered (uncertain), this is the desired target cell for all user types. The likelihood of KL values Diverging i.e., the EE distribution, from the optimal or desired power values that would satisfy the SINR over the transmission range, is low. Although it is unlikely to meet the efficacy of the optimal the first cell, coverage will be longer and more stable than the previous target cell. Principally, the shape of the KL mountains is influenced by received power values in different states while their width determines reliability of the link over transmission distance and is influenced by the channel gain variation. The peaky EE(KL) points on the graphs are largely influenced by received power. Contrarily, the flatter sides of the graph, are largely influence by high channel gain. BSs with flatter graphs as in Figure 6.9 may lead to more reliable connectivity in both NLOS and LOS. Particularly because the links may still be capable of meeting the desired SINR even when receivable power is low due to brokage or distance.

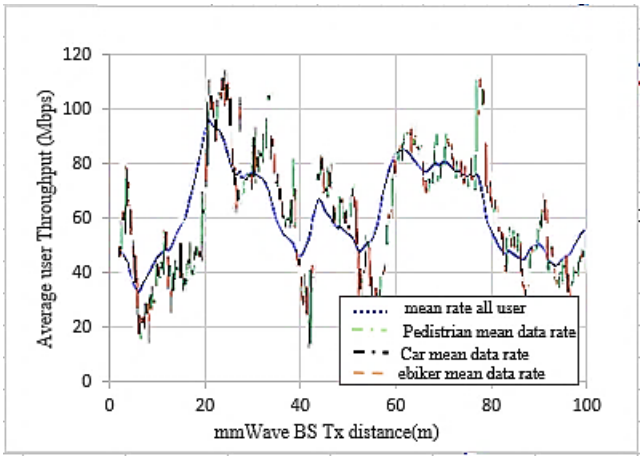

Figure 6.10: Mean throughput against distance in mm-BSs and different user types.

Figure 6.10 shows the mean throughput vs. distance from the connected mm-BS for different types of users. The pedestrian users have a higher data rate at each point. This is attributed to the JMLS-MFGT HO scheme to predict more accurately the behavior of the link for slower (low speed) users. Particularly, the channel state information hence training data from one state to the other gradually changes for slower users. This is unlike high-speed user training data where the change in CSI is more than often rapid, abrupt and inconsistence to learn rate. This ultimately negatively affects the incremental behavioural prediction of the pattern of a target link by the proposed scheme. To that effect, to quicken the behavioural prediction of target link, we fine-tuned the HO to only predict likely deterioration pattern of target links using pedestrian data. It only uses the mean data of all user type when the targeted link results with respect to pedestrian data leads into a link failure after the HO.

Further, additional analysis of the cause of spikes and lows/flat in data rate over distance in Figure 6.10 was done. Results show that the spikes are generally influenced with increase in received power e.g., due to LoS scenarios. The low data rate for instant at 10m and 20m (user-Bs distance), were mainly due to low received power and the high channel gain compensating some of the loss in received power to keep the SINR above threshold. By extension, this supports the notion where the KL divergence analysis in selecting reliable links chooses target links whose channel gain value over

distance will sustain SINR above threshold and any increase in received power above receiver sensitivity increase the data rate i.e. spikes.

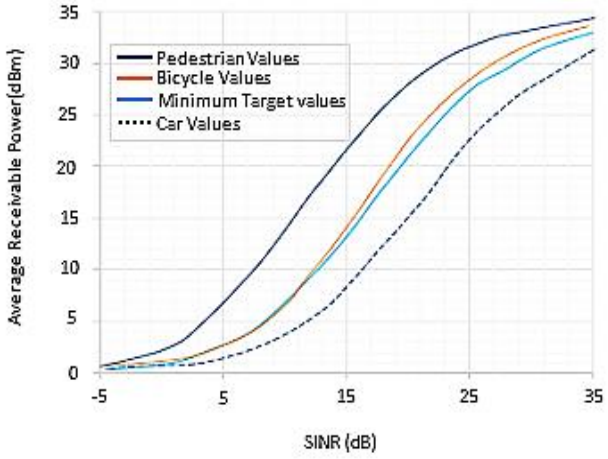

Figure 6.11: Received power vs SINR variation for different user types.

Figure 6.11. SINR vs Received power variation for different types of users in BS two in Figure 6.9 i.e., after HO. Considering cell capacities are often greater than 1 Gb/s for mmWave cells, a SINR for the lowest rate i.e., NLOS, would experience greater than 10 Mb/s as seen in Figure 6.10. Most SINR and received power from selected target links in Figure 6.11 would satisfy many of the rate requirements for 4G systems [5] [6]. Additionally, only about 10% of the assumed 100m coverage area of target BS is predicted to be under 0 dB. This means the outage probability would be less than 10% if our proposed HO mechanism is adopted.

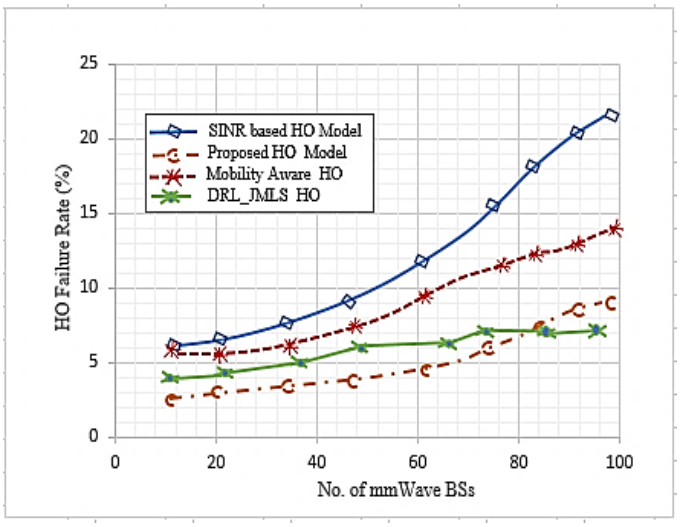

Figure 6.12: HO failure rate for different HO schemes

Figure 6.12 shows the HO failure rate against number of BSs for different HO schemes. We compare the performance of our HO scheme to those whose selection is based on mobility awareness [11], SINR-based [7] and DRL-JMLS [13]. The mobility awareness in [6] executes HO commands based on user speed. The SINR based in [4] executes HO commands based on the next mm-NB with the highest SINR/data rate. The DRL-JMLS is based on DRL and JMLS models to complete the HOs. From the results obtained in Figure 6.12, our proposed scheme experiences less HO failures with the number of mm-BS. For a larger number of mm-BSs, the HO failure rate for DRL-JMLS and our proposed scheme is similar because both have sufficient training data to make reliable HO policies. For a smaller number of BS hence limited training data, our HO scheme performs better.

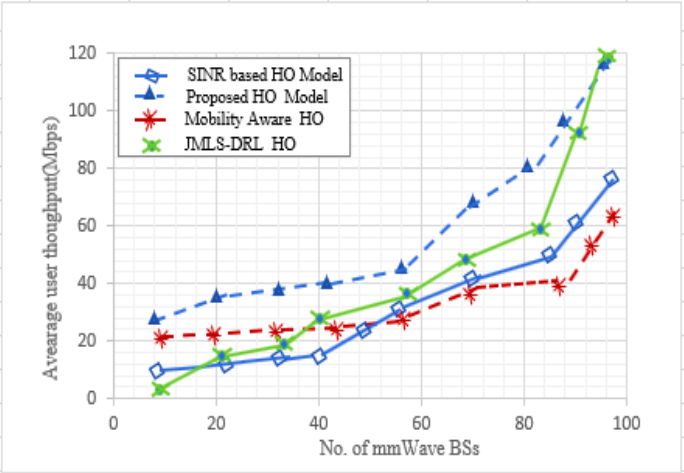

Figure 6.13: Average data rate vs no. of BSs for different HO schemes.

Figure 6.13 shows the average data rate of the network when different HO schemes are used. Similar to Figure 6.13, results , Figure 6.13, results show that the average data rate is higher in our proposed scheme when compared to the others. For a large number of mm-BSs, the average data rate for DRL-JMLS and our proposed scheme is also similar as both learn to select better target links, i.e., links with good LOS. It competes better with DRL-JMLS scheme when using fewer BS as it requires less data for training to make more reliable HO decisions. The scheme doesn't just select target BSs because of their high initial data rate/SINR/ received power. It also considers received power and channel gain variance post HO.

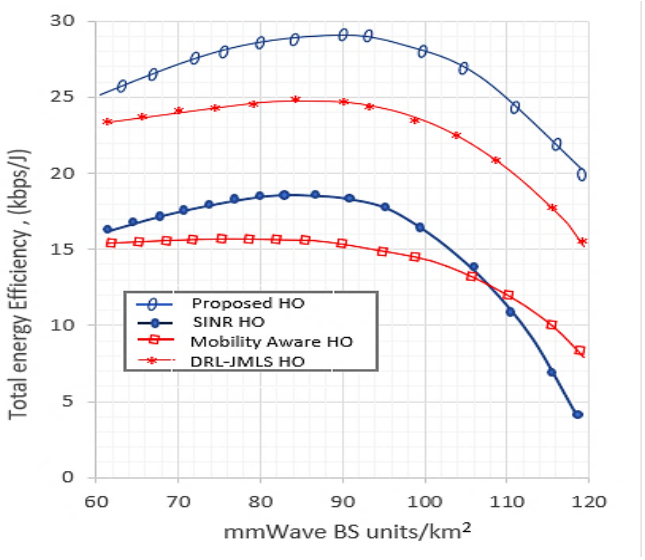

Figure 6.14. The Energy Efficiency vs mm-BS Density variation.

Figure 6.14 shows the variation of energy efficiency for different HO schemes against the reference density of mm-BS. As the mm-BSs increased, the EE rose to 90 BSs per square km. It then slowly decreased when the mm-BS number increased beyond 90 noting transmission power fixed. The decrease emanates from increasing distortion brought about by mmWave transmission sensitivity. We also observed that EE decreases with the increase in cellular number is due to the increase in cellular power thereby creating interference. Above all, for the same number of BS/km^2, the proposed HO scheme proved to be more EE than the DRL based in [13], SINR based in [7] and Mobility aware schemes in [11]. At peak, i.e., 90 BS/km^2, it is about 17% better than a DRL, and 30% - 40% better than SINR based and mobility aware HO schemes.

6.6 Conclusion

In this work, we investigated the performance of using the deterioration pattern to determine reliable links in 5G mobile networks. An efficient GT-JMLS modelling of likely target link variation/pattern was presented to solve this. Specifically, mean field-based game theory has been applied to optimize EM estimates in the selection matrix. Simulation results showed that our proposed joint trained GT-JMLS system significantly outperforms the existing HO algorithms with robust and longer stable connectivity. Thus, we conclude that the proposed design can be employed as an alternative of the classic HO optimization algorithm in practical systems. This algorithm is easy to implement, and the solving speed is fast.

Chapter 7

The 5G Mobile Network Connectivity-Aware Route Selection

The quest for autonomous cars has widened the scope for novel intelligent services needed to manage vehicles on a highway. Some of these services include e-Horizon and cognitive driving assistance services that are mainly focused on enhancing safety. Others include infotainment services rendered to enhance user quality of experience. The challenge for these services is how to access neighbour/far wide car data relevant to its safety. Enormous differences exist on how the safety content can be shared among autonomous cars. For those that have decided to share data through mobile networks, enormous responsibilities for ensuring supporting mobile networks and technologies guarantee connectivity safety and scalability exist. Mobile networks for instance will need to ensure that the content is shared efficiently and relevantly within the desired time. This implies networks must guarantee connectivity with the least minimal latency to enhance safety among autonomous cars. To that effect, smart routing schemes with Edge Computing (EC)

schemes such as: Vehicular Edge Computing (VEC) in [2] - [4], or 5G networks with mobile edge computing in [4] or Mobility-aware edge caching and computing in vehicle networks in [5] among other popular solutions have emerged. These ensure vehicular resources are accessed quicker at the edge. These technologies further leverage blockchain and AI technology [6] to enable real-time packet routing and caching of data. Further, VANET architectures with its On-Board Units (OBUs) and RSUs using AI are being integrated to further enhance VEC capability [7]. The principal concept in all these schemes is to bring network resources nearer to the cars and reduce delay in accessing the resources. The challenge however is despite smart routing schemes, there is usually no guarantee that network links will be available to support the ongoing routing of data in moving cars. The problem in 5G is further exacerbated by the vulnerability of mmWave links to user mobility and infrastructure, i.e., are easily blocked despite their abundant bandwidth rates. Furthermore, the alternative microwave bands, though known for longer range transmission and resilience, have insufficient bandwidth thus are not able to sustain the huge data flows/links between edge and autonomous cars in content exchange. Therefore, given the ability by autonomous cars to select their own route to destination, and also having taken into consideration band sensitivity problems in mmWaves, the limited bandwidth in microwave bands, short resident time of users owing to high user mobility among others. In this chapter, rather than autonomous cars simply choosing the routes to destination just on the basis of the shortest distance, we proposed a network availability aware route selection scheme that considers mobile network availability and connectivity guarantees to provide 5G coverage to autonomous cars on the route selected for/by a car to reach their destinations. This method offers guarantee that safety messages shared are delivered on time, and consequent driving decisions are accurate and catapults infotainment content sharing in autonomous cars.

7.1 Background

Because of its superior performance in continuous control problems, deep reinforcement learning (DRL) has found widespread application in path planning and other areas. In [23], for instance, the authors propose a DRL-based autonomous path planning model for intelligent path planning of unmanned vehicles. Then using DDPG as a foundation, the author of [11] proposes a framework for autonomous car-following planning on par with human performance. In this setup, driverless cars gain knowledge of their surroundings through trial and error. DRL is currently used in many

areas, including but not limited to network control (e.g., self-reconfiguration and optimization) [9], automatic driving [10, 11], financial prediction (e.g., predicting a financial meltdown or bankruptcy) [13, 14], and traffic control [15, 16]. The majority of these DRL have some experience with game theory. In fact, game-theoretic strategies are frequently used in the design of deep learning models. Example: the popular deep learning approach to classification problems can be viewed as a Stackelberg game. When it comes to solving difficult computer vision problems, the deep learning architecture known as the Generative Adversarial Network (GAN) has emerged as a frontrunner. GANs were originally developed from game theory. The problem with deep learning is that there are so many hyperparameters, making optimization problems much more difficult to solve. In order to achieve high levels of accuracy, they need to be set up properly, with features like accurate offline training data. Problems arise if the algorithm is not given enough time to run in order to solve the problem at hand and converge on one or more effective policies. Additionally, the existence of a variety of architecture designs complicates the DRL's ability to zero in on a single best solution.

7.2 Contributions

In this chapter we investigate aforementioned 5G mobile network heterogeneities using different driving paths that incur different transmission costs on data packet transmission as the user drives towards their destination. We want to minimize the cost of using the radio resources to transmit a burst of data, assuming that the Ultra-Reliable and Low-Latency Communications (URLLC) application needs them to be delivered with the required reliability and within a given latency:

- We propose a driving route selection approach that considers mmWave network availability and guarantee, i.e., not just the shortest route to destination. The approach considers a route and network selection as Mean field game. Here, the user’s interest is to reach the desired destination faster, however, the autonomous car’s main concern is safety and the shortest route is not always the one that has better network connectivity guarantees. In Mean Field Games, unlike other game theory concepts such as the GAN used in DRL [156], it is possible to select a representative player, average result and focus on its policy in selecting driving routes instead of considering player/user impact individually. This simplifies tremendously the computation of an equilibrium and ultimately Deep Learning decision policies if deployed in DNNs. Additionally, it

makes our solution faster than DRL and less computationally expensive. This is particularly so because it averts looping over several user actions to find the optimal policy [134].

- We develop a new heuristic algorithm based on user route selection to destination and network availability. According to the scores, we rank all paths to destination in the order of mmWave network reliability, resource-usage and cost-effectiveness. We define searching mean field algorithms that measures the network resources greedily to result. And this results in driving path searching order that guarantees mmWave network connectivity, minimal resource usage high reliability and reduced delay constraints.
- We adopt a HO system that adapt HO triggering time based on the effects of the moving vehicle's driving path relation to mmWave BSs as users approach their destination. Particularly, the HO scheme does not trigger a HO trial on a target cell anticipated to have a short residence time with the car or when the target cell is just an intermediate cell that the vehicle/user will quickly passes through.

7.3 Network Resources

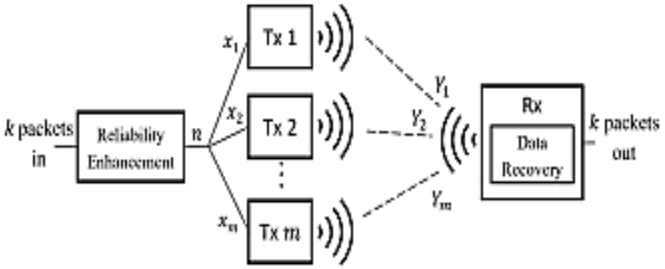

Figure 7.1: Reliability-enhanced transmission over multi-path network.

We consider the autonomous safety data user scenario where a burst of data is composed of δ packets and the multi-path network consists of M mm-BSs servicing K routes. paths, as shown in Figure 7.1. Each time the burst of packets are required to be delivered through the network to the receiver within a certain allowable delay and must satisfy a given reliability target ϵ. As mentioned earlier, we assume that the network is delay stringent and retransmission mechanisms like ARQ/HARQ do not react fast enough to meet the reliability target. Therefore, will require a network that is proactively reliable and a HO scheme that can delegate a certain number of cells

to a particular route upfront to protect data packet losses in the wireless 5G mmWave transmissions.

7.3.1 Network Resource Model

We formulate the Resource problem mathematically.

A) Outage Probability and Data Rates

Let π denote the access policy for a given set of network optimization parameter set Θ. The corresponding outage probability, P_π, given a set of observable signals Y_k, is defined as [2, 11]:

$$P_\pi(Y_k|\Theta) \triangleq P\left(\sum_{\ell}\sum_{s_k}^{S_K} b_l \log_2\left(1+\gamma_k(x)\right) \geq r^{m\zeta}(\hat{\gamma}_k)\right), \quad (7.1)$$

where γ_k and $\hat{\gamma}_k$ denote the actual and target SINR, respectively. $r^{m\zeta}$ represents the minimum target data rate over state $s_k \in S$. b_l is the bandwidth for the given channel link l. Assuming that all gNBs directionally transmit equal maximum power P. For a minimum receiver sensitivity of x_{kmin}, we denote a threshold x_{k0}, where $x_{k0} > x_{kmin}$ and any gNB link ℓ, that needs additional transmit power than P, to meet x_{k0}, is not established/loses connection if active, i.e., experiences a truncation outage. The distance $d = \left(\frac{P}{x_{k0}}\right)^{\frac{1}{\propto^{k_L}}}$ is the maximum user-BS transmission distance. α denotes the path loss exponent [25]. For a cut-off threshold, x_{k0}, maximum transmission distance in LOS and NLOS is denoted by $\left(\frac{P}{x_{k0}}\right)^{\frac{1}{\propto^{k_L}}}$ and $\left(\frac{P}{x_{k0}}\right)^{\frac{1}{\propto^{k_{NL}}}}$, respectively. Assuming a received power x_t is recorded for the ℓ link at state $s_k \in S$. The projected target receivable power x_{t+1} is:

$$x_{t+1} = \max \sum_{x_t,u_t} \left\{\frac{\hat{\gamma}}{\gamma^{min}} x_t - \frac{\alpha {x_t}^2}{\beta \hat{\gamma}^2}\right\}, \quad (7.2)$$

where $\{.\}^+ = \{max, 0\}$. x_t is current received power in LOS. $\hat{\gamma}$ and γ^{min} are targeted and measured SINR needed to satisfy $\mathbb{R}^m$. It must be noted that if there exists an infeasible SINR target in certain user state, the resulting power demand, x_{t+1}, by users may diverge to infinity. This is due to each user link attempt to meet its own required SINR no matter how high the power consumption can be. Thus α and β are power and SINR scaling factors respectively to substantially enhance reasonable deviations of x_{t+1} in NLOS. And the corresponding data rate is defined as:

$$r^m = b \log_2 \left(1 + \frac{P|h^H \boldsymbol{p}|^2}{(1+ d^{\propto})} F_x(|\theta_k^l|)\right), \tag{7.3}$$

$$\varphi_k^l(\cdot) = \frac{1.4\ x\ 10^4}{f_c(GHz).v(km/h)}, \tag{7.4}$$

where $\theta_k^l = \frac{2d \sin \varphi_k^l}{\lambda}$ denotes the angle of arrival for main lobe beam p. v, represents user velocity under 50 Km/h. And f_c, denotes the link's carrier frequency. $F_x(|\theta_k^l|)$ denotes the Fejér kernel value and $F_x(|\theta_k^l|) \rightarrow 1$,as user speed reduces to halt hence regulates SINR approaches to maximum value. F_x approaches 0 as v increases [4]. $|h^H \boldsymbol{p}|^2$ is channel gain vector for a channel model [9], [12]:

$$g_x = \rho_x^0 \mathscr{b}(\varphi_k^l)_x^0 + \sum_{l=1}^{L} \rho_x^l \mathscr{b}(\varphi_k^l), \tag{7.5}$$

where $\rho_x^0 \mathscr{b}(\varphi_x^0)_x^0$ and $\rho_x^l \mathscr{b}(\varphi_x^l)$ represent the LoS and NLoS vector for the *l*-th link between the BS and user *m*, given x, power is received. Moreover, φ_k^0 and φ_k^l denote LoS and NLoS spatial directions, respectively, described in (7.4),

B) Data Rate Optimization Problem

The minimum user data rate, $\mathbb{R}^m$, requirement problem for any route among μ_v possible routes to user destination is defined by state sequence, S, given outage and power constraints at different distance $d \in S$ from serving BS is defined as in (7.6):

$$\max_{\Theta} \sum_t \sum_{S_t, l} \left(1 - \mathbb{P}_{\mathfrak{y}b}\left(P_\pi^{m|x_t} + P_\pi^{m|u_t}\right) r_l^m(y)\right) \geq \mathbb{R}^m, \tag{7.6}$$

where $P_\pi^{m|x_t}$ and $P_\pi^{m|u_t}$ are LOS and NLOS conditional outage probability for a user in the m^{th} state, respectively. r_l^m is the maximum attainable data rate at user-BS distance d. $\mathbb{P}_{\mathfrak{y}b}$ is the blockage probability defined as the average number of blockages on chosen route to destination.

The projected energy cost E_c, for receiving x_{t+1}, given, $r^m > \mathbb{R}^m$ in (7.3) and (7.6) is satisfied is:

$$E_c = \beta \left\{ x_t \delta \frac{c(t-w)}{\mathbb{R}^m} + e_0 * \zeta c(t-w) \right\}, \tag{7.7}$$

where β represents the unit price per unit energy consumption, $c(t-w)$ represents the actual number of packets received by the user at *t* during window, *w*. $c\,(t-w)/\mathbb{R}^m$ is the latency. x_t is the current received power at time *t*. e_0 is the unit energy per packet, $e_0 * \zeta c(t-w)$ denotes energy lost due to lost packets (expected number less the actual number of received packets) at *t* during window *w*. Given receivable x_t, and transmittable P, power constraints, the is to ensure problem optimum packet delivery latency to ensure maximum link utility is formulated as [19]:

$$\max \sum_{P} \sum_{x_t,} \{x_{t+1} \delta c(t-w) - \zeta P E_c\}, \tag{7.8}$$

where δ expected latency scaling factor given, x_{t+1} is received *w*. ζE_c, is the latency discrepancy following x_t to x_{t+1} change as the user moves away from the serving BS.

C) Energy-efficiency Estimates

Let EE be the received packet data volume/unit time over power consumption. During user-BS association, there is no user plane data transmitted, we define the EE as, EE_0(t) and $EE_1(t)$ as the EE after association. EE_1(t) and EE_0(t) are denoted as:

$$EE_0(t) = {\mathbb{R}^m}/{P_c(t)}, \tag{7.9}$$

$$EE_1(t) = {r^m}/{(P_c(t) + P_o(t))}, \tag{7.10}$$

where $P_c > 0$ is the average circuit/control plane power consumption, $\mathbb{R}^m$ is the minimum data rate and r^m is the data rate at HO and P_o is the associated transmit power needed to meet x_{t+1} in (7.6). At HO, the energy efficiency analysis is done by comparing $EE_0(t)$ and $EE_1(t)$. With respect to (7.2), the rate increases with reduced association distance in LOS. Assuming $P_o(t)$ is the transmit power limit. At HO, a HO is only accepted $EE_1(t) > EE_0(t)$.

D) Distance-Weighted Degree-of-Connectivity

The degree of connectivity of a user is a measure of the number possible target gNBs from various points of the user-serving gNB transmission distances. For a mobile user, the closer the transmission boundary of the 1-hop gNB is, the number of HO. To avoid multiple HOs on the way

to destination, we calculate the distance-weighted connectivity of a gNB. It is defined as the function of the user-gNB Euclidean distance d_E, and transmission range, $d_x = \left(\frac{P}{x_{k0}}\right)^{\frac{1}{\propto k}}$ as:

$$\psi_d = \begin{cases} 1 & d_E \leq (0.6 * d_x) \\ 2.5\left(1 - \frac{d_E}{d_x}\right) & (0.6 * d_x) < d_E \leq d_x. \\ 0 & > d_x \end{cases} \quad (7.11)$$

Given the receiver sensitivity threshold x_{k0}, where $x_{k0} > x_{kmin}$, the location on the route to destination where the serving gNB-user link meets x_{k0} or SINR threshold and triggers a HO is awarded a ψ_d value. The closer the user-gNB distance point is where the link meets x_{k0} or SINR cut off limit the smaller the value of ψ_d awarded to the gNB on that route. Distances beyond maximum transmission distances are accorded $\psi_d = 0$ and the reference gNB is discarded for HO. The higher the value of ψ_d, the longer the gNB is likely to remain in LOS. The higher the total value of ψ_d on the route to destination, the smaller the number of HOs and the higher the chance of network connectivity up to destination. The distance weighted degree of connectivity of user $m \in M$ over total travelling distance to destination D is:

$$\lambda = \sum_{,m} \psi_d. \quad (7.12)$$

E) Route HandOver Pattern

The HO sequence for a user over a route is recorded using the HO record matrix MxM denoted $\boldsymbol{Q}$, with M representing the total number of cell users engage in their move. The record matrix, $\boldsymbol{Q}$, is defined as:

$$\mathbf{Q} = \begin{bmatrix} q_{11} & \cdots & q_{1m} \\ \vdots & \ddots & \vdots \\ q_{m1} & \cdots & q_{mm} \end{bmatrix}. \quad (7.13)$$

Initially, the elements in $\boldsymbol{Q}$ are equalled to 0. A mmWave HO from gNB i to gNB j increments q_{mm} by the value of ψ_d, depending on the user-gNB distance. The $\boldsymbol{Q}$, HOs, are transformed into HO probability matrix $\boldsymbol{P}$ for a given user path to destination.

$$\mathbf{P} = \begin{bmatrix} p_{11} & \cdots & p_{1m} \\ \vdots & \ddots & \vdots \\ p_{m1} & \cdots & p_{mm} \end{bmatrix}, \tag{7.14}$$

where $p_{ij} = q_{ij}/\sum_i q_{ij}$ and satisfies the conditions; $\sum_i p_{ij} = \lambda$ and $0 \leqslant p_{ij} \leq \psi_d$. Here, p_{ij} denotes the HO probability from gNB i to gNB j; $i,j \in S$, where the road space is given by state S**.** Given **X** is the expected receivable power pattern. The receivable power pattern set $\{\mathrm{X}_0, \mathrm{X}_1, \dots, \mathrm{X}_n\}$ at locations $\{d_0, d_1, \dots, d_n\} \in \boldsymbol{S_N}(\mu_v) = (s_0(n), s_1(n), s_2(n), \dots)$ reflecting consecutive user-gNB distances of HO over observation $N = \{0,1, \dots, n\}$.

F) The Game Concept

The game for each user, v, given a set of alternative routes, $\mu_v(\boldsymbol{S_N})$**,** to destination by e.g., Google Map applications is that:

1: Select the shortest route, $\boldsymbol{S}(\boldsymbol{N})$, that will still guarantee mmWave connectivity up to destination.

2: Select a set of gNBs in $\mathcal{Q}_i$ (7.13) on route $\boldsymbol{S}(\boldsymbol{N}) \in (\mu_v)$ that will optimise energy efficiency $EE_1(t) > EE_0(t)$ conditions in (7.9) and (7.10), cost in (7.7) and data rate condition in (7.6).

The optimal HO, *n,* selection for a user is one that optimise the immediate reward as: $r^n(s_n, a_n, \psi_d)$. Here the initial requirement for HO transitions in (7.13) is that no set of targets gNB cell and serving gNB cell, i.e., i and j, establish a HO at a symmetrical receivable power point, i.e., point of equal received power $x_i = x_j$; otherwise the HO transition from cell i to j in (7.13) is wasteful and accorded low ψ_d similar to NLOS condition.

However, the optimal route selection strategy is one that maximize condition 1 and 2 over the long term-reward and is defined as:

$$Reward = \max \mathbb{E}\left[\sum_{n=0}^{N} r^n(s_n, a_n, \lambda_n)\right]. \tag{7.15}$$

We formulate a non-cooperative game, were each player selfishly tries to find the shortest route and a route with high mmWave network availability i.e., meeting conditions 1 and 2. However,

the Nash equilibria in pure strategies do not necessarily exist. The shortest route might not always have the best connectivity and vice versa.

Note: The Nash equilibrium is where every participant in a noncooperative game can optimize their outcome based on the other players' HO decisions. Nash equilibrium is achieved in a game when no player has any incentive for deviating from the best long-term rewarding strategy in (7.15), even if they know the other players' strategies over the route has better immediate results. we must introduce the notion of mixed strategies – this means that each player type uses a family of strategies with a certain pattern of HO probability distribution. Given the *n*-step HO transition probability matrix, P^n, the HO probabilities at points *n*, $\boldsymbol{S(n)} = (s_0(n), s_1(n), s_2(n), \dots)$ is obtained by un-conditioning $P^{(n)}$. Let us denote $P(Q_i)$ the space of all HO probability measures on Q_i. A mixed strategy of player(users) switching from i to j will be a set of gNBs with $P(Q_i)$ needed to sustain connectivity on route, S . The set P(Q) is endowed with a sequence xN in $P(Q)$ converges to x, n $\in$ $P(Q)$.

7.3.2 Route Selection Problem Modeling

Knowing each player's best rewarding strategy hence route can be a daunting task to learn and requires huge training load. MFG involves identical and anonymous players making strategic decisions. MFGs can focus on a representative player's policy instead of all players' individual behaviour. This simplifies equilibrium calculation and ultimately the selection of routes to destination with high and efficient network guarantees. We're interested in computing Nash equilibria for user route selection and network availability guarantees as a game. As stated earlier on, it is not always guaranteed that the shortest route has the best connectivity possible. Thus, we we are cognizant of the fact that not all routes will be efficient. Particularly because of mobility, the location of each vehicle is time varying such that communication data rates, content transmission latency, the energy consumption and content delivery rates are time-varying on various routes leading to destination. The function of system is to match roads with a certain set of cells *m*.

Let $\boldsymbol{S_N(\mu_v)} = (s_0, s_1, s_n, \dots)$ denote the coordinates of each route at the n^{th} HO as given by Google map and N is the average number of HO taken by one vehicle v. $m \in M$ denotes the average number of cells involved in HO in each route $\boldsymbol{S_N(\mu_v)}$ up to s_n and $\boldsymbol{\mu_v(k)}$ is the distribution of k possible routes vehicle v can use at s_n to reach destination. $\boldsymbol{A}$ denote a set of SINR variation

over state space $\boldsymbol{S_N}$. Then $a_n \in A$ is the SINR value $\frac{\hat{\gamma}}{\gamma^{min}}$ in (see 7.6) at the $n-th$ HO. $\{d_0, d_1, \dots, d_n\}$ locations reflect consecutive points of handover over observation $0,1, \dots, n$ and $\{\mathrm{X}_0, \mathrm{X}_1, \dots, \mathrm{X}_n\}$ are receivable powers pattern between user and BS up to the point of HO d_n .

After taking the *k-th* route to destination given $\boldsymbol{\mu_v(k)}$, the vehicle will receive an immediate reward $\mathcal{r}^n(s_n, a_n)$. Since the objective of problem (7.5), (7.9) and (7.10) is to maximize network utility, we define the immediate reward as:

$$\mathcal{r}^n(s_n, a_n, k) = \begin{cases} \mathbb{E}\left[r^m \Big/ \left(P_c(t) + P_o(t)\right)\right] \\ \quad \text{if } EE_1(t) > EE_0(t) \text{ in (7.9) and (7.10)} . \\ plt(k), \qquad \text{Otherwise} \end{cases} \tag{7.15}$$

If the connection on the *k-th* route to destination satisfies constraints (7.9) and (7.10) condition in (7.15), the immediate reward is the current *k-th* road utility with data rate $r^m \geq \mathbb{R}^m$. Otherwise, the $q_{mm} \in \boldsymbol{Q}$ will receive a penalty and $\mathcal{r}^n(s_n, a_n, k) = plt(k)$, where $plt(k)$ is a negative constant -1. The system will need to look for another route after updating possible routes $\boldsymbol{\mu_v(k)}$.

The optimal route selection strategy is to maximize the long-term reward which can be defined as:

$$\mathcal{R} = \max \mathbb{E}\left[\sum_{n=0}^{N} \mathcal{r}^n(s_n, a_n, k_n)\right]. \tag{7.16}$$

Based on system state, action, and reward, we attempt to utilize Mean Field Game Theory to solve the proposed route selection problem.

7.3.3 Mean Field Modeling

As note earlier, we denote N_T the total Number of HO over finite time T and $n \in N$ a generic HO over a time space T . We focus on route space $\boldsymbol{S_N}$, the corresponding received power pattern and SINR variation over route $\boldsymbol{S_N}$, is denoted respectively by X and $\mathcal{A}$. Assuming ΔX denotes the set the power distributions of X over path $\boldsymbol{S_N}$. Each distribution may be regarded as a vector with length $|X|$. Th MFG's one-step strategy for a user is a HO, n with immediate reward r^n: $X * A * \Delta X \to R$. The corresponding HO probability term is denoted by: p^n: X*A*ΔX → ΔX, for $n = 0, 1, \dots, N_T$. The third argument corresponds to the current distribution of the BSs on the road. An initial distribution m_n of the BS is set and remain fixed over the paper.

For mean field, the two key quantities are player connectivity policy $\pi = (\pi^n)_n \in (\Delta_A^X)^{N_T - 1}$ and the corresponding distribution flow of roads of agents(users) $\mu = (\mu^n)_n \in \Delta_A^{N_T - 1}$. Since game HO

n and that π and μ are thus time-dependant objects. Given the possible destination routes μ, data, the goal for the game learning agent is to maximize the total network reward over policy π and the long-term reward in (7.16) is defined as:

$$\mathcal{R}(\pi, \mu) = \max \mathbb{E}_{\pi}\left[\sum_{n=0}^{N} r^{n}(s_n, a_n, \mu_n, x_n)\right], \tag{7.17}$$

$s.t.$: $a_n \sim \pi^n(.|x_n, s_n)$, $xn+1 \sim p^n(\cdot|s_n, a_n, \mu_n, x_n)$, $n \geq 0$. It must be noted that r^n and p^n are functions of the current route distribution data μ_n, at HO step, *n* which involves the mean field interactions. A policy π, is considered the best reward against a mean field road selection $S \in$ μ when it maximizes $\mathcal{R}(\pi, \mu)$. We define long term reward as the set of best Rewards $BR(\mu)$, to the mean field destination paths μ. Therefore while, μ, is not deduced from π, the ultimate decision of μ, is reliant on the network policy π. Technically, for: μ_0, given $n = 0, \ldots, N_T - 1$,HOs, the selection matrix is:

$$\mu_{n+1}(x) = \sum_{n, a', x'} \mu_n(x'_n)\, \pi^n(x'_n|x_n, s_n)\, p^n(x'_n|s_n, a_n, \mu_n, x_n), \tag{7.17}$$

which we can simply write $\mu_{n+1} = p_n^{\pi^n, \mu_n}\, \mu_n$, where $p_n^{\pi^n, \mu_n}\, \mu_n$ is the selection matrix given π.

Definition 1. Considering game conditions (1) and (2). If (7.16) is the Best Reward against one of the routes in $\hat{\mu}$, and (7.17) selection is triggered by (7.16), then the pair $(\hat{\pi}, \hat{\mu})$ denotes the Mean Field Nash Equilibrium (MFNE). Given the mean field routes for a destination, a typical MDP problem can be formulated for a representative player, and the MFGT problem can be analysed with standard DRL methods characterized through the $\mathcal{Q}$-function defined as: $\mathcal{Q}_{N_T-1}^{\pi,\mu}(\mathrm{x}, \mathrm{a}) = 0$ and for $n \geq N_T - 1$:

$$\mathcal{Q}_{N_T-1}^{\pi,\mu}(\mathrm{x}, \mathrm{a}) = \mathbb{E}\left[\sum_{n'>n} \left(r^{n'}(s_{n'}, a_{n'}, \mu_{n'}, x_{n'})\middle| s_n, a_n, x_n\right) = (s, a, x)\right]\pi^i. \tag{7.18}$$

It satisfies the Bellman equation: for $n \geq N_T - 1$,

$$\mathcal{Q}_{N_T-1}^{\pi,\mu}(\mathrm{x}, \mathrm{a}) = r^n(s_n, a_n, \mu_n, x_n) + \mathbb{E}_{x', a'}\left[\sum_{n'>n} \mathcal{Q}_{N_T-1}^{\pi,\mu}(\mathrm{x}', \mathrm{a}')\right], \tag{7.19}$$

where $a_n \sim \pi^n(.|x_n, s_n)$, $x_{n+1} \sim p^n(\cdot|s_n, a_n, \mu_n, x_n)$, $n \geq 0$, with the convention $Q^{\pi,\mu}_{N_T-1}(\mathrm{x}, \mathrm{a}) = 0$ for $n \geq N_T - 1$. The optimal $Q^{*,\mu}$ denotes the best Q-function value function of best policy π^* against μ. It is described by:

$$Q^{*,\mu}_{N_T-1}(\mathrm{x}, \mathrm{a}) = \max \mathbb{E}_\pi \left[\sum_{n=0}^{N} r^n(s_n, a_n, \mu_n, x_n) \right], \tag{7.20}$$

and satisfying the optimal Bellman equation: for HO $n \geq N_T - 1$,

$$Q^{*,\mu}_{N_T-1}(\mathrm{x}, \mathrm{a}) = r^n(s_n, a_n, \mu_n, x_n) + \mathbb{E}_{x',a'} \left[\max_{a'} Q^{*,\mu}_{n+1}(\mathrm{x}', \mathrm{a}') \right], \tag{7.21}$$

where $Q^{*,\mu}_{N+1}(.,.) = 0$.

7.3.4 Fine Tunning

Rather than calculating BR whenever path is chosen, an alternative strategy is to review a policy at each iteration. This does not require iterating over each user data to determine the ideal one. This procedure is more efficient than computing a BR f or finite road destinations. Therefore, at each iteration *k*, the policy iteration method for MFG, f irst evaluates the current policy π^k, by computing $Q^{k+1} = Q^{\pi_k,\mu_k}$, and then at each iteration $k + 1$, it selects using the greedy policy $\pi_{k+1}(.|x)$ that maximizes $Q^{k,\mu}(\mathrm{x}, .)$ Considering the road data is known before hand, then a backward induction, can be used to perform the evaluation at HO selection. This allows data from previous cycles to be used and stabilize the learning process. The Online method for MFG can be summed up as follows: after initializing q_0^n and π_0^n0 n for $n = 0, \ldots, N_T$, repeat at each iteration k:

1. Road update: $\mu^k = \mu^{\pi^k}$
2. Q-function update: $Q^k = Q^{\pi^k, \mu^k}$
3. Regularized HO Q update: $\bar{q}^{K+1} = \bar{q}^{K+1} + \frac{1}{t} Q^k$
4. MFGT Policy update: $\pi_n^{k+1}(\cdot|x)$

Algorithm 7.1:

Input: Google Map data of study area $\boldsymbol{S}$, total BS number M numbers of MFGT routes K and DQN route estimation iterations,
Output: cumulated $\mathcal{Q}$ value function, policy π
Initialize the parameters θ^0 in (7.1) – (7.8), SINR HO $\mathcal{Q}$ and $\mathcal{P}$ in (13)-(14)
Set $\pi^0(\mathrm{a}|(\mathrm{n},\mathrm{x})) \sim \left(\frac{1}{\tau}\widehat{\mathcal{Q}}_{\theta^o}((x_n), a_n)\right)$ for $s \in \mathrm{S}$
Calculate initial $\mathcal{J}(\pi, \mu)$ in (7.16) for n
for *route* $k = 1, \ldots, K$ **do**
  Generate Route Distribution: u^K with π^{k-1}
  DRL Value function: Initialize θ^k
    for $\ell = 1, \ldots, \mathcal{L}$ **do**
      Sample: minibatch of N_B HO transitions over $K \in S$:
        $\{(n_i, x_i), a_i, r_i(x_i, a_i, \mu_{ni}^k), (n_i + 1, x_i')\}_{i=1}^{N_B}$
        with $n_i \leq N_T$. received power projection $x_i' \sim p^n(.\,|x_i, a_i, \mu_{ni}^k)$ and a_i is chosen by $\in -greedy$ policy based on $\widehat{\mathcal{Q}}_{\theta^k}$ **in Algorithm 7.2**
      Update θ^k using gradient step:
      $\theta^k = \frac{1}{N_B}\sum_i^{N_B}\left|\widehat{\mathcal{Q}}_\theta((n_i, x_i), a_i) - T_i\right|^2$
      where T_i is given in (7.26)
    end for
    Policy: for all n, x, a
        $\pi^K(\mathrm{a}|(\mathrm{n},\mathrm{x})) \sim \left(\frac{1}{\tau}\widehat{\mathcal{Q}}_{\theta^k}((n, x), a)s\right)$
    end for
Return:
$\pi^K, \widehat{\mathcal{Q}}_{\theta^k}$
$update\ \mathcal{Q}$

We suppose that a representative agent in a finite horizon MFG setting is encoded by a policy, either explicitly or implicitly (through a Q-function), and is able to realise an episode in the sense that: for $n = 0, \ldots, N_T$, the agent observes x_n and chooses a **BS** with SINR $a_n \sim \pi_n(\cdot\,|\mathrm{x}_n)$ and mean field GT returns a realization of $\mathrm{x}_{n+1} \sim \mathrm{p}_n(\cdot\,|\mathrm{x}_n, \mathrm{a}_n, \mu_n)$and $r^n(s_n, a_n, \mu_n, x_n)$ for some n in the range [0,...,N T]. The mean field route data is not directly observed by the agent; rather, as a parameter of the HO transition and reward functions $\mathrm{p}_n(\cdot\,|\mathrm{x}_n, \mathrm{a}_n, \mu_n)$ and $r^n(s_n, a_n, \mu_n, x_n)$. The agent can approximate the Q-functions $\mathcal{Q}^{\pi,\mu}$and $\mathcal{Q}^{*,\mu}$ following (7.20) and (7.21) using these as samples, as shown in (7.20) and (7.21), instead of using trajectories that begin at any other set of

coordinates (time, state). To that effect, we began the simulation with a trajectory that began at 0 and state x0 ~ m0 instead of starting from any pair (x, a). To make it easier to implement, we just on the state as (n, x_n), and ignore time. This is justified as the road network is fixed. This allows us to employ regular Q-learning. It is important to remember, nevertheless, that for some stationary Bellman parameters, the Bellman equations are not fixed-point equations. In big road networks, it is not practical to do a detailed evaluation of each and every pair of states (x, a). For reasons of both memory efficiency and generalisation, MFGT allows us to approximate the Q-functions with nonlinear functions like neural networks, such $\mathcal{Q}_{\theta}^{\pi,\mu}$ and $\mathcal{Q}_{\theta}^{*,\mu}$, parameterized by θ. The parameters of the neural network are then trained by replacing the variables in (7.18), (7.21), and (7.22) with the minimization of a loss. Specifically, the quantities are minimised over time, which is considered an input. Namely, treating time as an input, one minimizes over θ the quantities:

$$\widehat{\mathbb{E}}\left[\left|\mathcal{Q}_{\theta,n}^{\pi,\mu}(\mathrm{x},\mathrm{a}) + r^{n}(s_n, a_n, \mu_n, x_n) + \mathcal{Q}_{\theta_t,n+1}^{\pi,\mu}(\mathrm{x}',\mathrm{a}')\right|^2\right], \tag{7.21}$$

$$\widehat{\mathbb{E}}\left[\left|\mathcal{Q}_{\theta,n}^{\pi,\mu}(\mathrm{x},\mathrm{a}) + r^{n}(s_n, a_n, \mu_n, x_n) + \max_{a'}\mathcal{Q}_{\theta_t,n+1}^{*,\mu}(\mathrm{x}',\mathrm{a}')^2\right|\right], \tag{7.22}$$

where $\widehat{\mathbb{E}}$ is an empirical expectation based on samples and θ_t is the parameter of a target network.

7.3.5 Deep Reinforcement Learning for Mean Field Optimization

Our fixed road network serves as inspiration as we swap out the Q-function at iteration k with a neural network whose parameters K are tuned to minimise a loss function (7.25). Since our goal is to gain knowledge through time, we will choose (*n*, *x*) as the initial condition. To be specific, given samples of transitions:

$$\left\{\left((n_i, x_i), a_i, r^{n}(s_n, a_n, \mu_n, x_n), (n_i' + 1, x_i')\right)\right\}_i^{N_B} \tag{7.23}$$

where $a_n \sim \pi^n(.|x_n, s_n)$, $x_i' \sim p^{n_i}(\cdot|s_{n_i}, a_{n_i}, \mu_n^k, x_{n_i})$, $n \geq 0$, the parameter θ^K is trained using stochastic gradient descent:

$$\nabla_{\theta\pi}\mathcal{J} \approx \max_{\pi}\mathbb{E}\left[\nabla_x \mathcal{Q}(a_t, x_t|\Theta_Q)|_{s_t,\pi(x_t|\theta\pi)}\nabla_{\theta\pi}\mathcal{Q}(a_t, -x_t|\theta\pi)\right], \tag{7.24}$$

to minimize the empirical loss:

$$\frac{1}{N_B}\sum_i \left|\hat{\mathcal{Q}}_{\theta}\left((n_i, x_i), a_i\right) - T_i\right|^2, \tag{7.25}$$

where the target T_i is:

$$T_i = -\mathcal{r}^n\left(s_{n,i}, a_{n,i}, \mu_{n,i}^k, x_{n,i}\right) - \alpha t \log\left(\pi^{k-1}\left(a_i|(n_i, x_i)\right)\right) - \sum_{a'} \pi^{k-1}\left(a_i'|(n_i, x_i)\right)\left[\hat{Q}_{\theta^k}\left((n_{i+1}, x_n'), a_n'\right)\right] - \alpha t \log\left(\pi^{k-1}\left(a'|(n_i, x_i)\right)\right), \quad (7.26)$$

,

As a result of passing n and x as inputs to the DRL Q-network, our nomenclature has changed. This method of learning is similar to DQN, with the exception of two modifications to the target: (1) it integrates the penalty for deviating from the prior policy, and (2) we do not take the argmax over the next action, but rather a mean based on the previous policy. A simpler version of the procedure is described in Algorithm 7.1 with Algorithm 7.2 as its greedy search scheme.

Algorithm 7.2: Path cost-effectiveness score based greedy search

1: Calculate the cost-effectiveness scores (e.g., n_i or a_i)
2: Rank paths by their scores to obtain a path order in K to destination
3: Construct mapping $x_i' \sim p^n\left(. \,|x_i, a_i, \mu_{ni}^k\right)$ through line (12) and define function $n_i \leq N_T$
4: Find the smallest number of HO n within $[0, \sum_{i=1}^m N_i]$ such that $EE_1(t) > EE_0(t)$ is satisfied at all HO using linear search
5: **if** a $EE_1(t) > EE_0(t)$ in (7.15) **then**
6: **return** n^*, x^*, a^*
7: **else**
8: Declare route infeasible
9: **end if**

7.3.6 Proof-Of-Utility Consensus in Vehicular Networks

In this section, we present the details of route and network connectivity consensus selection mechanism for approved route. We propose how to evaluate BS utility for road verifier selection.

A) Consensus

Inspired by the ability of autonomous cars to independently select routes to user destination, the MFGT route-network consensus mechanism is made up of two parts: 1) delegate selection, 2) route and verification, as shown in Figure 2. Given Google Map data, we use the utility of base station to select HO delegates and routes.

1) ***Delegate Selection:*** Considering vehicles can't trust each other with data but still need to communicate data with one another. For the purposes of route selection, we characterize BSs as

dominant delegate. Delegates play a crucial role in consensus algorithms (i.e., road selection and network validation), and so should be non-participating nodes [36]. Since BSs are not directly involved in route selection, they do not directly benefit from efficient route selection process It makes them suitable as neutral nodes that can be considered to be dominant delegate candidates. Thus, at each step in the route selection process, the vehicles/user communicates the SINR value (Sound Reference Signal-SRS) [56]. The SINR value is used as a yardstick of gNB-user link quality. With its increased utility bandwidth from mmWave bands, BS is a reliable tool for ensuring a fair reward is being given to a reliable route among many that can be taken by the user to reach destination. Figure 7.2 elaborates on the methodology behind utility analysis. Each car sends one SRS along the route per HO, and the BS's voting power is inversely proportional to the number of HO it conducts.

The top candidates in the number of successful HO with reference to a specific route receive the most votes and become the dominate delegates on that route. Any other BS that a delegate successfully HOs a user to given a specific route is regarded as part of the route sequenced BSs (see equation (7.13) that can be used without reverifying their ability to support connectivity. Non delegate BSs are called verifiers. Thus, the selection scheme does not need to check their efficiency or verify their viability once the delegate has been chosen unless a HO failure is recorded which also reduces the votes of the delegate. This simplifies the MFGT learning as it is repeated largely only on delegates . It is important to note that at first, routes are chosen using the shortest route and BS HOs are performed using the SINR. However, as more data is acquired, the benefit of employing the dominant delegate BSs is analysed on the basis of MFGT based reward. The more delegate a route has the more reliable is the network availability in that specific route.

2)Route Selection and Verification*:* Delegates and Verifiers have two distinct roles in the Route Selection and verification process: leader and verifier. The leader is in charge of recruiting verifier on a route hence creating HO sequences. The verifier is in charge of checking the validity of the network. One of the delegates with the highest *n* HO takes the reins as the leader during each route selection procedure, while the others serve as the route's verifiers. Each delegate has an equal chance of becoming the leader and choosing the optimal route to reach their portion of the destination or the entire destination.

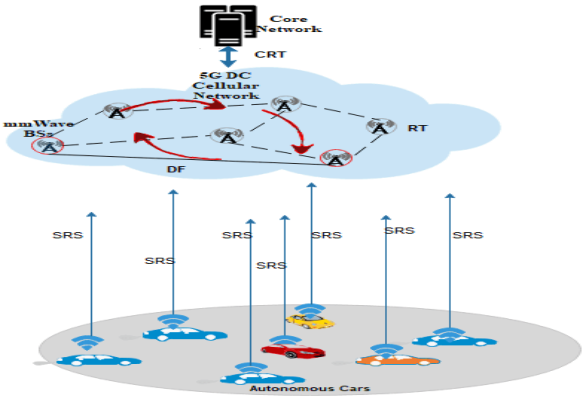

Figure 7.2: mmWave user-BS SRS transmission.

Based on Figure 7.2, for instance, the delegating BS with the largest accumulated Reward is the route selection leader. In our method of route selection and verification, the leading pointer/delegating BS is used to amass a subset of route data, which is analysed to estimate rewards and produce a path with unverified network availability. Broadcasting route data, verifying the routes, and confirming them are the three stages that make up the route verification protocol. At the start of the route broadcast phase, the top delegate sends out the data about the *n* HO messages to Game. The user's location, the time, and the route's long-term reward, up to n, are all included in the shared messages. Each verifier begins the route verification step by checking the signature of the SRS message. Verifiers then send their audit results, signed by each verifier. MFGT learners ensure the accuracy of route cached matches Google Maps routes.

After receiving audit findings from other verifiers, each verifier compares its own results to those received, and then sends a confirm message to the leader and the core network. The comparison result is the responsibility of the central network, which keeps records of all audit results obtained from other verifiers. The correct reward for each route is determined by the core after receiving all confirm messages from delegates and analysing them. To store and participate for a HO given (7.13), the core network will select all delegates in a specific route to participate in sequential HOs. Periodically, BSs outside the delegate commission will be updated on the most recent rewards from adjacent delegates. After the newly chosen route has been verified and included to the route as supporting networks on the route, the BSs used as delegates are continually monitored of their performance. At HO, other target BS that compete for user connectivity are skipped for HO till the delegate fails to HO. The supporting BSs can be directly engaged for HO by the dominant delegate.

The delegates do not just have the highest number of HOs but are also considered to be more energy efficient and offer the highest distance weighted degree of connectivity.

B) Utility Evaluation

Delegates BSs are those with the highest number of successful and high rewarding HO. The verifier BS handoff might can be skipped if necessary. This allows the vehicle to bypass a handoff to a delegate cell. This method works well in a city setting when the cells are stacked on top of one another. When the vehicle travels over two or more intermediate non delegate cells, it is common for the handover to skip those cells without disrupting network connectivity. Therefore, the HO procedure is sped up by the handover probability matrix with three or more steps.

Transfers are made using the one-step handover probability matrix P1 when a suitable destination delegate BS m cannot be chosen from Q m. This means the handoff is made to the delegate/verifier cell where the vehicle is most likely to move. If the target BS cannot be determined using P1, then the usual SINR based handover procedure is implemented. What is different to normal SINR HO, however, is that the HO can prompt a change of route to destination if the chosen BS is a delegate in another route. To reduce searching time, when the link fail on the dominant delegate, the quick exploitability of policy π is used: The exploitability of a policy π between using the k^{th} path with cell m_o and *k+1* path with the same cell m_o is defined as:

$$E(\pi) = \max_{\pi} \mathcal{R}\left(\pi_k^{m_0}; \mu_k^{\pi}\right) - \mathcal{R}\left(\pi_{k+1}^{m_1}; \mu_{k+1}^{\pi}\right). \quad (7.28)$$

An exploitability of 0 means that $\pi_k^{m_0}$is an MFNE policy and the user must maintain the initial route of choice to destination to receive maximum reward and the opposite is true when the exploitability is large. The user must switch to another path where the new serving BS is a delegate. . Similar notions are widely used in computational game theory [101],[134]

7.4 Performance Evaluation

7.4.1 Simulation Platform

The simulation of our scheme is carried out with NS-3, which is augmented with Open AI Gym to serve as a network simulator and MFGT-DRL platform, respectively. SUMO is used to carry out simulations of the urban traffic mobility. The LTE-mmWave DC module Ns-3 is modified and utilised to simulate a dual linked mode using radio technology operating at sub-6 GHz, as well as

a mmWave [10]. The DC module supports both Fast Switching (FS) and data rate-biased HO. Fast switch mechanism allows direct gNB to gNB HO without the Core network getting involved. The data rate/SINR based HO involves the core network using SINR of SRSs to select target networks. FS switching is usually used as a secondary HO. Thus we use core based HO on dominant delegates and FS switching on verifiers respectively(see details in [78]) . RRC and X2 procedures and primitives are implemented based on these two modes of operation, respectively. Users utilising FS are considered to be in the RRC CONNECTED state in relation to two eNBs; however, they will only send data to one of the two eNBs based on the DC control procedures that are in place. The MFGT control algorithms controls the selection given the SUMO data. The user considers the expected total reward for connecting to BSs along each of the potential routes to the destination before deciding which path to take to get there. By activating Route Mobility [67] in ns-3, one can acquire information regarding alternate routes to the destination as well as the coordinates of that destination. At HO every gNB has a SINR that is more than the minimal level (as calculated by DL MFGT for each state) necessary to fulfil the desired SINR over the provided ser states. The gNB who decides the highest number of successful HO is game leader/delegate dominant. The verifiers are allocation policy in accordance with what they observe. This allows for the automatic selection of a route network BS coordinator from among mmWave gNBs, as well as the control of route selection linked to mobility. The results of the simulation are shown in Table 7.1.

Table 7.1: Simulation Parameter.

Parameter	Value
mmWave	28GHz
mmWave bandwidth	1GHz
3GPP Channel Scenario	Urban Micro, Urban Macro
MMWave max outage	-5dB
mmWave transmission Power	46dBm
mmWave max PHY Rate	3.2Gbps
X2 link latency	1ms
S1 link latency	10ms
RLC buffer Size	5MB
S1 MME link latency	10ms
User speed	[1,50] m/s
UDP Source rate	200Mbits/sec
Propagation Model	TwoRayGroumd
Traffic Type	Constant Bit Rate(CBR)
Data Traffic Rate	6Mbps
Foreground Traffic (Packet size)	512 bytes
Background Traffic (Packet size)	1500,2500,4000,5000,6500 bytes
HELLO, Interval,	1 second
No of users	180
No of BSs	

7.4.2 Simulation

Figure 7.3: Manhattan Grid Road Network.

To test the efficacy of the proposed MFGT-DRL enabled section technique, we use an actual dataset given by Uber [37]. To model how cars might move to various locations, we use Uber's dataset's trajectory. This dataset includes details on 14.3 million Uber pickups in NYC from

January to June 2015 and 4.5 million pickups in NYC from April to September 2014. Figure 7.F3 depicts the location of our 100 sample cars within the larger observation region, which spans from 40°66'8671 to 40°67'8719 and 73°93'0269 to 73°95'0915. The size of the detected area is roughly a kilometre. Table III lists the settings used in the simulation. At the beginning of each simulation, vehicles are placed at random in the road system. The vehicles in the simulation immediately begin to move in accordance with a predefined mobility model as soon as the simulation begins. It is assumed in this model, known as the Car-following model, that a vehicle will travel at the same rate as the vehicle in front of it. On average, the vehicles travel 14 metres in one second, or 50 kilometres per hour. Studies have shown that mmWave BS can communicate up to 200 metres away, although the current configuration only allows for 100 metres. The simulation is conducted on NS-3, which makes advantage of the platform's mmWave Physical and MAC layer. At the media access control (MAC) layer, only Best Effort (BE) access is supported; no Quality of Service (QoS) features are enabled. The model of propagation employed here is the Two-Ray Ground model. We presumed that any network node with a location service, like GLS [4] or HLS [5,] could figure out where the final destination was. In order to study the impact of rearranged neighbours in the mmWave network, simulations of HELLO messaging intervals were conducted. As long as the neighbour item still has the potential to be a suitable forwarder, it will not be immediately deleted from the neighbours table if a HELLO or SRS message is lost or delayed due to collision or congestion in the network. To avoid adding to network congestion and to permit adequate sensitivity to changes in the link state without jeopardising the precision of implementing mmWave BS to BS HO, the waiting time is set to the mean dwell time. This was done so that the network wouldn't get even more clogged, and so that changes in the connection state could be responded to in a fair amount of time. The simulation features 180 different automobiles, all of which are constructed in the same fashion. We choose a dense network to test our method on because we want to see how it does in busy environments. The foreground connections, which number 10, are used to test the efficacy of the proposed approach. UDP is used to generate a Constant Bit Rate (CBR) of 512 bytes once every 0.5 seconds for each connection. More importantly, we inject background traffic ranging in size from 0 to 0.4 gigabytes to disrupt it because it can have a negative impact on routing metrics [28]. Additionally, the network is constructed in a way that causes congestion and interference between data flows, as the source and

destination pairs of foreground and background flows are predefined in a way that causes them to cross paths with each other. Our performance review findings are presented in the next part through comparison and analysis.

7.4.3 Simulation Results and Analysis

Experimental data from the simulation was analysed and the comparison with other HO mechanism, doing the route selection and HO selection separately were done given the total iteration time, the optimal decision time, and convergence steps to illustrate the accuracy and effectiveness of the improved model. Figure 7.4 shows the variation of the average user throughput against the number of mmWave BSs. As shown Figure 7.4, the proposed enhanced DRL HO scheme also known as the DRL-MFGT scheme, has the highest average data rate in comparison to SINR [6] using shortest route and DRL[13] scheme. Additionally, the scheme results show that 70 gNB (mmWave BSs) are enough to cover the study area and user number of 180. Technically, the change in the average data rate between 70 and 80 BS is very minute, i.e., less than 5%. The implications is that it is costly to add more BS than to gain the 5% average rate. This further affirms our proposed method is cost efficient in reducing the number of mmWave BS installations per unit area.

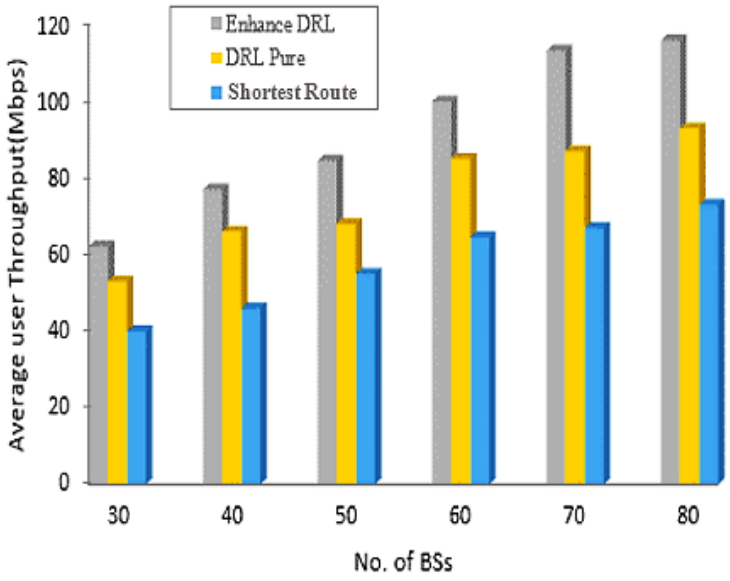

Figure 7.4: Average traffic throughput vs number of BSs for different HO schemes.

Figure 7.5 compares the HO success and failure rate of the proposed scheme at 10 different destinations. Ten random destinations for different users are randomly chosen. The results show that the average HO failure rate is less that 15%.

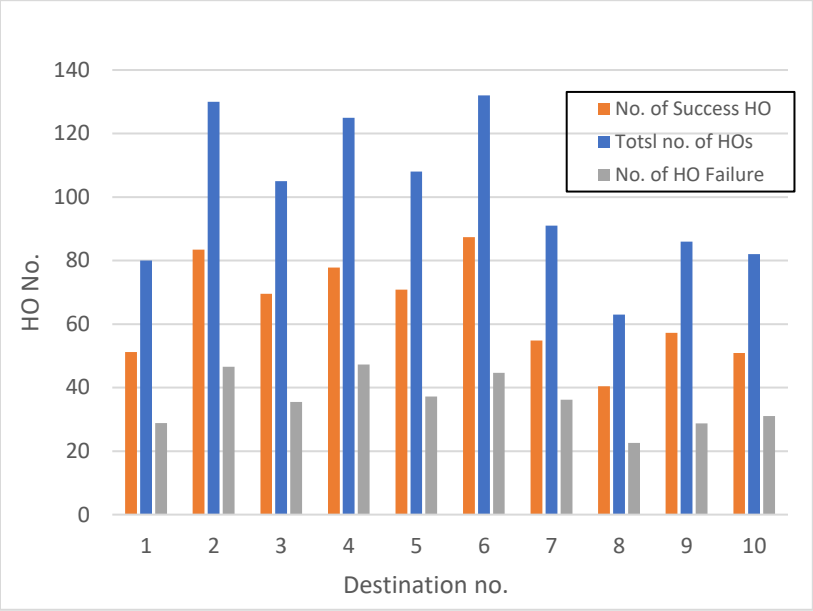

Figure 7.5: Handoff Success compared to Handoff failure.

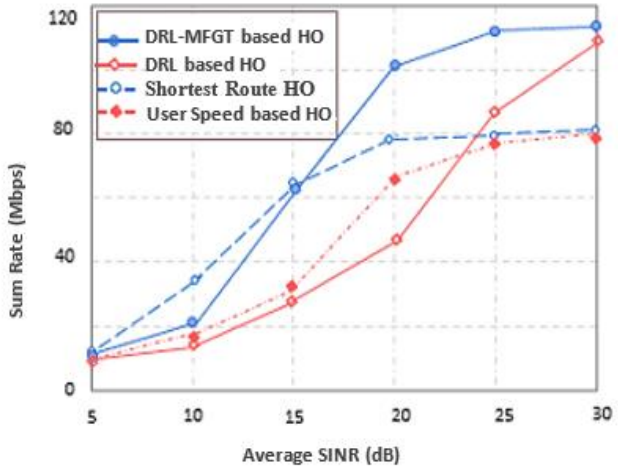

Figure 7.6: Sum rate vs outage variation given channel gain diversity impact.

Figure 7.6 shows the sum rate variation with averaged received power when different HO schemes are used. The analysis show that DRL-MFGT can greatly boost data rate in mmWave

communications. This is particularly because most of the transmitted power with reference to average rate variation is received. Keeping in mind that SINR variation greatly affects data rate, the received power variation also has significant bearing on the diversity of the channel gain to sustain a reasonable SINR for a link. The great improvement in average data rate of our scheme can be attributed to our scheme selecting target links with high channel gain diversity.

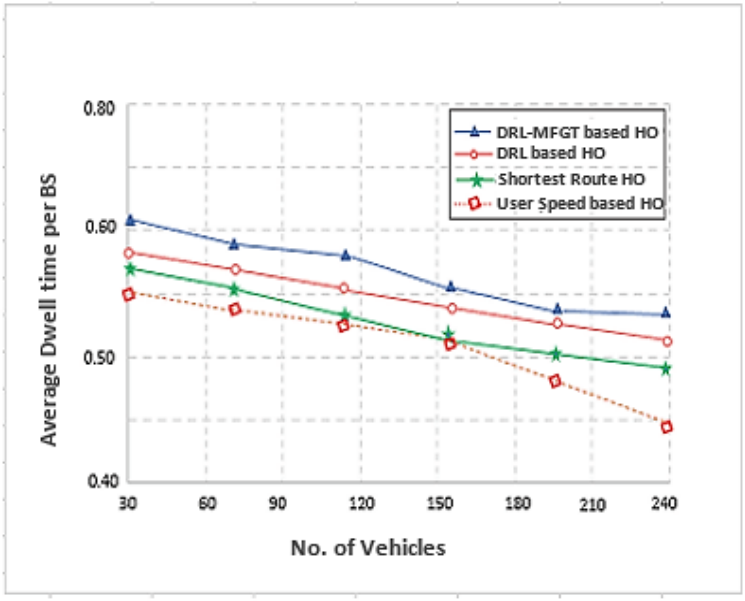

Figure 7.7: Compares DRL, Distance based and MFGT-DRL HO.

Figure 7.7 shows how the average dwell time varies with increase in the number of users. The results indicate that an increase in number of users reduces the average dwell time of a user per cell. Owing to the ability to select more viable target gNBs, our proposed scheme performed better than other competing HO schemes.

Figure 7.8 shows our evaluation of data packet success at different training episodes. Here we took the actual success ratio of each epoch when using DRL and DRL-MFGT HO scheme. As can be observed in Figure 7.8, the quick improvement in packet ratio success aligns with our proposed mechanism. The imagined forecast of behaviour change of several links on various driving pathways by scheme gives it advantage of choosing links that would hardly drop packets. There is a larger chance of packet success from the beginning due to the additional examination of linkages on routes, which increases the number of delegates supporting network availability in the route. In fact, the superior success rate may be attributable to the proposed scheme's training plan structure,

in which the training data is kept until a predetermined threshold of prediction uncertainty is reached, after which new policies are derived and average. Nonetheless, DRL builds initial low packet success owing less route knowledge and the need to loop through every user data to make credible policies. Longer routes to the source may reduce success rates because chance of pre-selected delegate gNB unsuccessfully implementing HO in very volatile transmission. In Figure 7.8, for subsequent epochs, the real success ratio of all considered algorithms in fact levels up. This shows just how fast/early the proposed mechanism formulates stable policies.

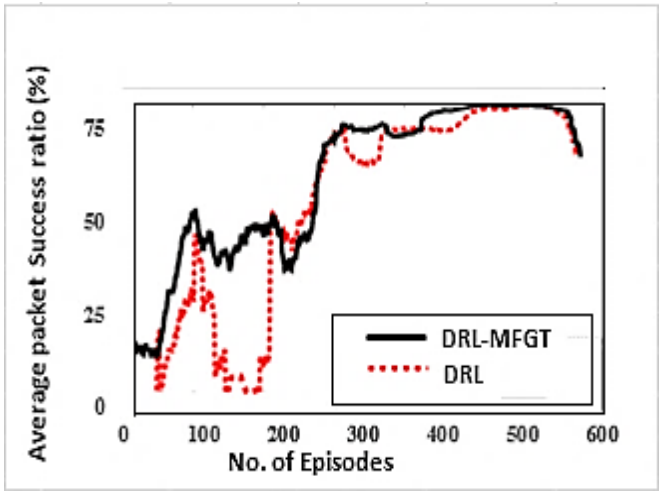

Figure 7.8: Compares DRL and Game inspired packet transfer success ratio.

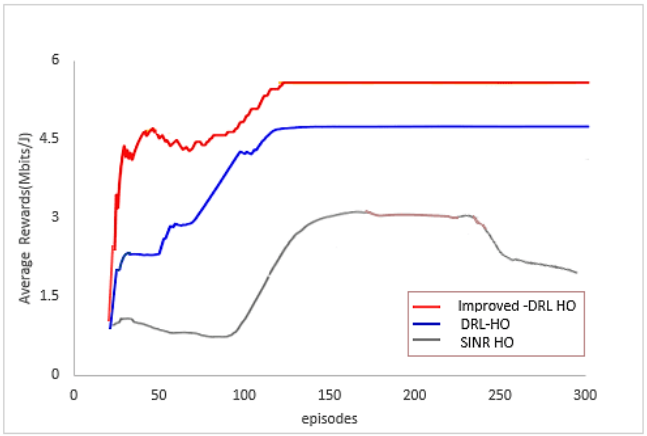

Figure 7.10: Average EE of network with Improved DRL and other algorithms.

Figure 7.9 shows the energy efficiency convergence rate over the number of episodes for different schemes. The results of EE for DRL and improved DRL (DRL-MFGT) methods show steady

convergence as shown by the plot curves. However, the DRL method's highest average EE is about 4.5 Mbits/J lower than that achieved by the MFGT-DRL approach. The enhanced method outperforms the others in terms of convergence rate and provides an overall network with an average EE that is greater than 4.5 Mbits/J.

7.5 Conclusion

The proposed procedure can thwart redundant handover to an intermediate cell through which the vehicle just passes. In the simulations, the performance gains of the proposed scheme have been measured in terms of the number of HO failures, HOs and average throughput over users, BS and training episodes. The results show that the proposed scheme has reduced the number of handover occurrences while maintaining adequate throughput. This chapter has further shown the advantages of combining route selection in autonomous cars with network connectivity awareness. It technically makes the reliance on 5G mmWave link more credible. This will ultimately increase safety as users are able to stay connected longer on high-speed links. On another note, routing selection is another important issue in autonomous cars that requires careful planning. Network disruptions are always likely to cause anxiety about the safety of autonomous cars. In fact, route recalculation is a costly process in terms of time and resource consumption. Therefore, routing recovery process should be integrated in handoff schemes viability to guarantee safety in future cars. On another note, it will be interesting to use blockchain in such systems. Taking advantage of its distributed ledger could be something that would improve the scale of this solution in future.

Chapter 8

Conclusion and Future Works

In this thesis, we conducted a simulation study of Handovers to overcome challenges and support the theoretical finding in 5G mobile networks.

In the first instance, this research explored the suitability of using Jump Markov linear system modeling to predict the target link's behaviour in real-time communication. With limited CSI, the scheme was able to predict the mmWave network behavior by using meta data techniques. The peculiarity of the proposed method is that it enables 5G mobile networks effectively execute reliable handover in a complex HetNet with very limited CSI. The intellectual capabilities of the proposed method are provided by EM and a Viterbi based estimation using Deep Neural Network. This allows the HO scheme to periodically predict target link pattern changes and their reliability with very limited data. The HO scheme adaptively selects mmWave target cells with a more gradual deterioration pattern. During the HO execution, the scheme accounts how the global mmWave link behaviour pattern (generated from all the gathered CSI) compares with local predictions (individual target links). In order to compare the two patterns at specific user-BS distances, we adopted the KMO approximation technique. The HO analysis and selection is then

viewed as a mixture of the estimates corresponding to local and global data. Results showed the proposed scheme selects cells with better HO failure rate, data rate and dwell time than traditional HO models.

In the second AI HO solution, we proposed an intelligent HO using DRL and JMLS. By integrating deep reinforcement learning and JMLS in 5G mmWave networks, we first proposed a framework where different users exhibiting different network behaviour when using mmWaves BS. To learn the dynamics of the network topology and time-variant channel condition, we exploited JMLS to maintain a reasonable training overhead before more stable policies are formulated by DRL. The interference metrics in JMLS modelling accelerates target link behaviour verification while the DRL policies result into more stable HO policies. This scheme simply compensates the slow decision-making process in DRL with JMLS inference decisions. At the same time, it self corrects the errors in the inference decisions using DRL policies. The performance parameters studied are throughput, latency, packet delivery ratio, etc. for different user topologies. The study found the HO scheme offers target links with better data rate, HO failure rate and latency than existing DRLs. Thus, the analysis shows that our proposed HO can achieve high data rate and less HO failure. Simulation results were based on a real dataset from Uber. And results compared to normal DRL indicate that the scheme significantly outperforms most ordinary HO benchmark policies.

Another method proposed involved the Mean Field Game Theory. The Mean Field Game Theory techniques are promising solutions to meet the increasing complexity and overheads associated with 5G wireless systems. The chapter 7 highlighted the use of mean field and JMLS. The resource optimization technique uses JMLS modelled estimates as meta data to Mean Field approximation to optimize mmWave HetNets and proposed solutions using various parameter techniques including energy efficiency and data rate optimization. Various parameters e.g. received power and channel gain variation were incorporated to build the HO system models and the simulation results validated the effectiveness of HO model based in proposed frameworks under mmWave technology, dual mode BSs, and using dual slope path loss models. The advantage of MFGT is that it offers distributed based HO decision. Each target link is analyzed based on individual cell behaviour and thereafter the results are strategically extrapolated to predict how a group of individual will behave under certain scenarios. This makes the training less complex and faster

than DRL methods. The results showed a HO scheme that is more energy efficient , requires less overheard than DRL schemes.

In another proposed HO mechanism, considering the traditional path planning algorithms cannot select routes to destination-based network availability. This study suggested a MFGT-DRL-based autonomous path planning method for unmanned cars. The autonomous car path planning and network aware method based on improved DRL is suggested by combining Mean Field Game Theory and DDPG. The enhanced DRL is contrasted with other traditional DRL techniques. The first step is it gathers user destination data using an google path planning platform. Secondly, the model is developed and trained in conjunction with Mean Field Game Theory and user experience to select the route with the most reliable links. The results of the simulation demonstrate that cars respond appropriately to new network available routes, successfully carry out autonomous path planning based on network availability and implement an end-to-end learning strategy. The outcomes demonstrated that the improved network availability on routes and can achieve continuous operation output, has less navigation error, and the convergence speed is good enough to support the model and accuracy, further validating the efficacy of the suggested approach. The biggest hurdle though how to deal with road traffic congestion.

8.1 Future Works

With regard to use of JMLS, almost all mmWave behaviour exhibit a non-linear dynamic in state evolution. This is despite the fact that we linearized every inference. Therefore, future research should concentrate on creating a HO model for non-linear cases. Secondly, the proposed model's validity for newer 5G/6G model and a real-data cases needs to be investigated. On another note, AI, machine learning, deep learning and natural language processing will play an important role in future communication and gaming. The task to combine them and make 5G more effective is inevitable. The ubiquitous deployment of mmWave cells in 5G mobile network will have many challenges such as dynamic environmental conditions, poor signal, interference, energy efficiency, availability of free spectrum, etc. Smart HO schemes are required to accurately predict real-time information about the important data terms link failure, the sequence of link deterioration, etc. for HO analysis purpose. Further, self-discovery, self-configuration and automatic HO deployments will be important pursuing with block chain technology. The distributed blockchain ledger, its

traceability, authenticity, confidentiality, and intrusion prevention offers great opportunities for 5G HO models.

Further, AI-based techniques if fined tuned proved to be able to acclimatize and be competent for the smart and faster HO decisions in 5G HetNets. In our future work, we plan to optimize the features used in the classification and regression stages of the ML-based workload orchestrator by using an intelligent optimization algorithm. This will pave way on how Federate learning can be customized in tacking and analyzing multiple competing deterioration factors in mmWave communication. In addition, we want to implement a deep learning-based blockchain approach HO solution to solving the 5G mobile network workload orchestration and balancing problem e.g. in V2V offloading in which the vehicles have an on-board computation unit. One instance of using such solutions could be blockchain involved in vehicular networks from different cities across the world safely sharing traffic data and decision-making processes. This will reduce the training time when a vehicle is brought in a totally new environment. For instance, a vehicle in Japan can passively learn the driving 'culture and traditions' in New Zealand prior to importation. Once imported, it doesn't need to start anew learning and adapting to New Zealand Road rules. This could be effective in avoiding accidents and traffic congestion than expecting a car to adapt first when it first comes to New Zealand. The decentralized operation is a core feature of Blockchain. The Blockchain operates with the decentralized transaction validation and approval mechanism. Furthermore, the transaction ledger is decentralized, and each node contains a copy of the ledger. In contrast with the centralized storage and validation systems, the intruders require more effort to take over the control of decentralized validation systems and the decentralized ledger. The blocks of transactions are cryptographically linked, and it is hard to alter them due to the decentralization.

References

[1] V.Nguyen, et. al, “Understanding Millimeter Wave Spectrum for 5G Networks,” *5G Americas*, Bellevue, Washington, Dec 2020 [online]. Available: https://www.5gamericas.org/wp-content/uploads/2020/12/InDesign-Understanding-mmWave-for-5G-Networks.pdf.

[2] J. Du and R.A. Valenzuela,” How Much Spectrum is Too Much in Millimeter Wave Wireless Access,” arXiv:1704.04274v1 [cs.IT], Apr. 2017.

[3] I.F. Akyildiz, C. Han and S. Nie, "Combating the Distance Problem in the Millimeter Wave and Terahertz Frequency Bands,” *IEEE Communications Magazine*, vol. 56, no. 6, pp. 102-108, June 2018, doi: 10.1109/MCOM.2018.1700928.

[4] S. Rangan, T. S. Rappaport and E. Erkip, "Millimeter-Wave Cellular Wireless Networks: Potentials and Challenges,” *Proceedings of the IEEE*, vol. 102, no. 3, pp. 366-385, March 2014, doi: 10.1109/JPROC.2014.2299397.

[5] F. Devoti, I. Filippini and A. Capone, "Facing the Millimeter-Wave Cell Discovery Challenge in 5G Networks with Context-Awareness,” *IEEE Access*, vol. 4, pp. 8019-8034, 2016, doi: 10.1109/ACCESS.2016.2628917.

[6] B. Shubyn and T. Maksymyuk, "Intelligent Handover Management in 5G Mobile Networks based on Recurrent Neural Networks," *2019 3rd International Conference on Advanced Information and Communications Technologies (AICT)*, 2019, pp. 348-351, doi: 10.1109/AIACT.2019.8847734.

[7] M. Polese, M. Giordani, M. Mezzavilla, S. Rangan and M. Zorzi, "Improved Handover Through Dual Connectivity in 5G mmWave Mobile Networks," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 9, pp. 2069-2084, Sept. 2017, doi: 10.1109/JSAC.2017.2720338.

[8] S. Niknam, A. A. Nasir, H. Mehrpouyan and B. Natarajan, "A Multiband OFDMA Heterogeneous Network for Millimeter Wave 5G Wireless Applications," *IEEE Access*, vol. 4, pp. 5640-5648, 2016, doi: 10.1109/ACCESS.2016.2604364.

[9] M. Mezzavilla et al., "End-to-End Simulation of 5G mmWave Networks," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 3, pp. 2237-2263, third quarter 2018, doi: 10.1109/COMST.2018.2828880.

[10] J. Park, S. Y. Jung, S. -L. Kim, M. Bennis and M. Debbah, "User-Centric Mobility Management in Ultra-Dense Cellular Networks under Spatio-Temporal Dynamics," *in 2016 IEEE Global Communications Conference (GLOBECOM)*, 2016, pp. 1-6, doi: 10.1109/GLOCOM.2016.7842367

[11] M. Joud, M. García-Lozano and S. Ruiz, "User specific cell clustering to improve mobility robustness in 5G ultra-dense cellular networks," in *2018 14th Annual Conference on Wireless On-demand Network Systems and Services (WONS)*, 2018, pp. 45-50, doi: 10.23919/WONS.2018.8311661.

[12] M. Chiputa, P.H.J Chong et al., "Cognitive Radio Technologies in 5G mm Wave Cellular Networks: Challenges and Solutions", *Advances in Engineering Research*, Victoria M. Petrova, New York, NY, USA, Nova Science Publishers, 2022, pp. 1-23.

[13] M. Chiputa, M. Zhang, G. Ali, P.H.J. Chong, H. Sabit, A. Kumar, H. Li, "Enhancing Handover for 5G Mobile Networks using Jump Markov Linear System and Deep Reinforcement Learning," *Advanced Antenna Techniques for IoT and 5G Applications Sensors, MDPI Sensors Journal*, vol. 22, no.3, pp. 145-167, Jan 2022.

[14] M. Chiputa, P.H.J. Chong and A. Kumar, "An Autonomous Deterioration Prediction Based Handover Model for 5G Networks," *2021 IEEE International Conference on Computing (ICOCO)*, 2021, pp. 168-173. doi: 10.1109/ICOCO53166.2021.9673555.

[15] M. Chipta, M. Uttarwar and P. H. Joo Chong, "Intelligent Handover using User-Mobility Pattern Analysis for 5G Mobile Networks," *proceedings 2021 Conference on Information Communications Technology and Society (ICTAS)*, 2021, pp. 5-10, doi: 10.1109/ICTAS50802.2021.9395027.

[16] M. Chiputa, P.H. Joo Chong, S.U. Rehman, A. Kumar," Investigating Mobility Robustness in 5G Networks Using User-Adaptive Handoff Strategies," in *Springer Ad Hoc Networks. ADHOCNETS 2019. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, vol 306.,2019. Cham. https://doi.org/10.1007/978-3-030-37262-0_8.

[17] T. Bai and R.W. Heath, Jr.,"Coverage and rate analysis for millimeter wave cellular networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 2, pp. 1100-1114, Feb. 2015.

[18] W. Abbas and M. Zorzi," Context information based initial cell search for millimeter wave 5G cellular networks." May 2016 [Online]. Available: https://arxiv.org/abs/1605.01930

[19] A. Redondi, I. Filippini, and A. Capone, "Context management in energy efficient radio access networks," in *Proc. IEEE 24th Tyrrhenian Networks Digit. Commun. -Green ICT (TIWDC)*, Sep. 2013, pp. 1-5.

[20] A. Capone, I. Filippini, and V. Sciancalepore,"Context information for fast cell discovery in mm-Wave 5G networks," in *Proc. Eur. Wireless 21th Eur. Wireless Conf.*, 2015, pp. 1-6.

[21] R. Chávez-Santiago, M. Szydełko, A. Kliks et al.,"5G: The Convergence of Wireless Communications," *Springer Wireless Personal Communication*, vol.83, no.1 - 17, 2015.

[22] N. Yong, Y. Li, D. Jin, et al.," A Survey of Millimeter Wave (mmWave) Communications for 5G: Opportunities and Challenges," arxiv: 1502.07228, Feb. 2015. [online]. Available: http://arxiv.org/abs/1502.07228v1

[23] A. Çalhan and C. Ceken, "An adaptive neuro-fuzzy based vertical handoff decision algorithm for wireless heterogeneous networks," in *Proc. IEEE 21st Int. Symp. PIMRC*, Sep. 2010, pp. 2271-2276.

[24] B. Ma and X. Liao, "Speed-adaptive vertical handoff algorithm based on fuzzy logic in vehicular heterogeneous networks," in *Proc. 9th Int. Conf. FSKD*, May 2012, pp. 371-375.

[25] S. B. Kotsiantis, “Supervised machine learning: A review of classification techniques," *Informatica,* vol. 31, no. 3, pp. 249-268, 2007.

[26] L. P. Kaelbling, M. L. Littman, and A. W. Moore, “Reinforcement learning: A survey," *J. Artif. Intell. Res.*, vol. 4, pp. 237-285, 1996.

[27] M. A. Khan, H. Tembine, and A. V. Vasilakos, “Game dynamics and cost of learning in heterogeneous 4G networks," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 1, pp. 198_213, Jan. 2012.

[28] C. Watkins and P. Dayan, “Q-learning,” *Mach. Learn.*, vol. 8, nos. 3-4, pp. 279-292, 1992.

[29] P. T. Semov, A. Mihovska, R. Prasad, and V. Poulkov, “Use of positioning information for performance enhancement of uncoordinated heterogeneous network deployment," in *Proc. 3rd Int. Conf. VITAE,* Jun. 2013, pp. 1-6.

[30] J. Park, S. Y. Jung, S. -L. Kim, M. Bennis and M. Debbah, "User-Centric Mobility Management in Ultra-Dense Cellular Networks under Spatio-Temporal Dynamics," *2016 IEEE Global Communications Conference (GLOBECOM)*, 2016, pp. 1-6, doi: 10.1109/GLOCOM.2016.7842367.

[31] Francesco Devoti et al., “Facing the Millimeter-Wave Cell Discovery Challenge in 5G Networks with Context-Awareness,” IEEE Access Special Section on Deployment and Management Of Small Heterogeneous Cells For 5g Volume 4, pp.819-834, Nov. 2016.

[32] A. Capone, I. Filippini, and V. Sciancalepore, “Context information for fast cell discovery in mm-Wave 5G networks," in *Proc. Eur. Wireless 21th Eur. Wireless Conf.*, 2015, pp. 1- 6.

[33] Z. Zhang, K. Long, J. Wang, and F. Dressler, “On swarm intelligence inspired self-organized networking: Its bionic mechanisms, designing principles and optimization approaches," *IEEE Commun. Surv. Tuts.*, vol. 16, no. 1, pp. 513-537, First Quarter 2014.

[34] O. G. Aliu, A. Imran, M. A. Imran, and B. Evans, “A survey of self organisation in future cellular networks," *IEEE Commun. Surv. Tuts.*, vol. 15, no. 1, pp. 336-361, First Quarter 2013.

[35] Z. Zhang, W. Huangfu, K. Long, X. Zhang, X. Liu, and B. Zhong”On the designing principles and optimization approaches of bio-inspired self-organized network: A survey," *Sci. China Inf. Sci.*, vol. 56, no. 7, pp. 1-28, Jul. 2013.

[36] A. Imran, A. Zoha, and A. Abu-Dayya, ``Challenges in 5G: How to empower SON with big data for enabling 5G," *IEEE Netw.*, vol. 28, no. 6, pp. 27-33, Nov./Dec. 2014.

[37] *HetNet/Small Cells*. [Online]. Available: http://www.3gpp.org/hetnet, accessed Aug. 2014.

[38] X. Su, J. Zeng, and C. Xiao, "Key technologies for SON in next generation radio access networks," in *Proc. 23rd ICCCN*, Aug. 2014, pp. 1-8.

[39] M. Peng, D. Liang, Y. Wei, J. Li, and H.-H. Chen, "Self-configuration and self-optimization in LTE-advanced heterogeneous networks," *IEEE Commun. Mag.*, vol. 51, no. 5, pp. 36_45, May 2013. L. T. W. Ho, I. Ashraf, and H. Claussen, ``Evolving femtocell coverage optimization algorithms using genetic programming," in *Proc. IEEE 20th Int. Symp. PIMRC*, Sep. 2009, pp. 2132-2136.

[40] S. Sun, M. Kadoch, and T. Ran, "Adaptive SON and cognitive smart LPN for 5G heterogeneous networks," *Mobile Netw. Appl.*, Jan. 2015, doi: 10.1007/s11036-014-0563-2

[41] W. Guo and T. O'Farrell, "Relay deployment in cellular networks: Planning and optimization," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 8, pp. 1597-1606, Aug. 2013.

[42] C.-S. Tseng, C.-C. Chen, and C. Lin, "A bio-inspired robust adaptive random search algorithm for distributed beamforming," in *Proc. IEEE ICC*, Jun. 2011, pp. 1-6.

[43] K. M. Passino, "Biomimicry of bacterial foraging for distributed optimization and control," *IEEE Control Syst. Mag.*, vol. 22, no. 3, pp. 52-67, Jun. 2002.

[44] W. Li, P. Yu, Z. Jiang, and Z. Li, "Centralized management mechanism for cell outage compensation in LTE networks," *Int. J. Distrib. Sensor Netw.*, vol. 2012, pp. 1-8, Jul. 2012.

[45] Z. Zhao, J. Chen, and N. Crespi, "Apolicy-based framework for autonomic reconfiguration management in heterogeneous networks," in *Proc. 7th Int. Conf. Mobile Ubiquitous Multimedia*, 2008, pp. 71-78.

[46] S. Fan, H. Tian, and C. Sengul, "Self-optimization of coverage and capacity based on a fuzzy neural network with cooperative reinforcement learning," *EURASIP J. Wireless Commun. Netw.*, vol. 2014, no. 1, pp. 57:1-57:14, Apr. 2014.

[47] M. Calef_ and L. Paura, "Bio-inspired link quality estimation for wireless mesh networks," in *Proc. IEEE Int. Symp. WoWMoM*, Jun. 2009, pp. 1-6.

[48] A.H. Tsai, L.C.Wang, J.H. Huang, and T.-M. Lin, “Intelligent resource management for device-to-device (D2D) communications in heterogeneous networks," in *Proc. 15th Int. Symp. WPMC*, Sep. 2012, pp. 75-79.

[49] A. Engels, M. Reyer, X. Xu, R. Mathar, J. Zhang, and H. Zhuang, “Autonomous self-optimization of coverage and capacity in LTE cellular networks," *IEEE Trans. Veh. Technol.*, vol. 62, no. 5, pp. 1989-2004, Jun. 2013.

[50] M. Peng, Y. Liu, D. Wei, W. Wang, and H.-H. Chen, “Hierarchical cooperative relay based heterogeneous networks," *IEEE Wireless Commun.*, vol. 18, no. 3, pp. 48-56, Jun. 2011.

[51] M. Liu, Z. Li, X. Guo, and E. Dutkiewicz, “Performance analysis and optimization of handoff algorithms in heterogeneous wireless networks," *IEEE Trans. Mobile Comput.*, vol. 7, no. 7, pp. 846-857, Jul. 2008.

[52] R. Chai, J. Cheng, X. Pu, and Q. Chen, “Neural network based vertical handoff performance enhancement in heterogeneous wireless networks," in *Proc. WiCOM*, Sep. 2011, pp. 1 - 4.

[53] M. Polese, M. Mezzavilla, and M. Zorzi, “Performance Comparison of Dual Connectivity and Hard Handover for LTE-5G Tight Integration,” in Proceedings of the 9th EAI International Conference on Simulation Tools and Techniques (SIMUTOOLS), 2016, pp. 118–123.

[54] 3GPP, “Study on New Radio (NR) Access Technology - Physical Layer Aspects - Release 14,” TR 38.802, 2017.

[55] M. Polese, M. Mezzavilla, and M. Zorzi, “Performance Comparison of Dual Connectivity and Hard Handover for LTE-5G Tight Integration,” in Proceedings of the 9th EAI International Conference on Simulation Tools and Techniques (SIMUTOOLS), 2016, pp. 118–123.

[56] M. Polese, M. Giordani, M. Mezzavilla, S. Rangan, and M. Zorzi, “Improved Handover Through Dual Connectivity in 5G mmWave Mobile Networks,” IEEE Journal on Selected Areas in Communications, vol. 35, no. 9, pp. 2069–2084, Sept 2017.

[57] Y. Huo, X. Dong, and W. Xu, “5G Cellular User Equipment: From Theory to Practical Hardware Design,” IEEE Access, vol. 5, pp. 13 992– 14 010, July 2017.

[58] Y. Huo, X. Dong, W. Xu, and M. Yuen, “Cellular and WiFi Co-design for 5G User Equipment,” ArXiv e-prints, March 2018. [Online]. Available: https://arxiv.org/abs/1803.06943

[59] D. J. Love and R. W. Heath, “Equal Gain Transmission in Multiple-input Multiple-output Wireless Systems,” IEEE Transactions on Communications, vol. 51, no. 7, pp. 1102–1110, July 2003.

[60] M. Rebato, M. Mezzavilla, S. Rangan, F. Boccardi, and M. Zorzi, “Understanding Noise and Interference Regimes in 5G Millimeter-Wave Cellular Networks,” in *22th European Wireless Conference*, 2016.

[61] F. Boccardi et al., "Spectrum Pooling in MmWave Networks: Opportunities, Challenges, and Enablers," in IEEE Communications Magazine, vol. 54, no. 11, pp. 33-39, November 2016.

[62] Wei, J. Yuan, D. W. K. Ng, M. Elkashlan, and Z. Ding, “A survey of downlink non–orthogonal multiple access for 5G wireless communication networks,” *ZTE Commun.*, vol. 14, no. 4, pp. 17–26, Oct. 2016.

[63] D. E. Papanikolaou, N. E. Papanikolaou, G. T. Pitsiladis, A. D. Panagopoulos and P. Constantinou, "Spectrum sensing in mm-wave cognitive radio networks under rain fading," *Proceedings of the 5th European Conference on Antennas and Propagation (EUCAP)*, Rome, 2011, pp. 1684-1687.

[64] D. Maamari, N. Devroye, and D. Tuninetti, “Coverage in mmWave cellular networks with base station co-operation,” *IEEE Trans. Wireless Commun.*, vol. 15, no. 4, pp. 2981–2994, Apr. 2016.

[65] J. G. Andrews, T. Bai, M. N. Kulkarni, A. Alkhateeb, A. K. Gupta, and R. W. Heath, Jr., “Modeling and analyzing millimeter wave cellular systems,” *IEEE Trans. Commun.*, vol. 65, no. 1, pp. 403–430, Jan. 2017.

[66] M. Xiao et al., "Millimeter Wave Communications for Future Mobile Networks," in IEEE Journal on Selected Areas in Communications, vol. 35, no. 9, pp. 1909-1935, Sept. 2017.

[67] X. Yang, M. Matthaiou, J. Yang, C. Wen, F. Gao and S. Jin, "Hardware-Constrained Millimeter-Wave Systems for 5G: Challenges, Opportunities, and Solutions," in *IEEE Communications Magazine*, vol. 57, no. 1, pp. 44-50, January 2019.

[68] M. Mezzavilla et al., "Public Safety Communications above 6 GHz: Challenges and Opportunities," in *IEEE Access*, vol. 6, pp. 316-329, 2018.

[69] O. Semiari, W. Saad, M. Bennis and M. Debbah, "Integrated Millimeter Wave and Sub-6 GHz Wireless Networks: A Roadmap for Joint Mobile Broadband and Ultra-Reliable Low-Latency Communications," in *IEEE Wireless Communications*.

[70] R. Chakraborty, N. Kumari, M. Mousam and A. Mukherjee, "The Future of 5G and Millimeter Waves," in *2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, Coimbatore, India , 2018, pp. 1679-1683.

[71] Y. Wu, A. Khisti, C. Xiao, G. Caire, K. Wong and X. Gao, "A Survey of Physical Layer Security Techniques for 5G Wireless Networks and Challenges Ahead," in *IEEE Journal on Selected Areas in Communications*, vol. 36, no. 4, pp. 679-695, April 2018.

[72] P. Yang, Y. Xiao, Y. Guan, Z. Liu, S. Li, and W. Xiang, “Adaptive SM-MIMO for mmWave communications with reduced RF chains,” *IEEE J. Sel. Areas Commun*., vol. 35, no. 7, pp. 1472–1485, Jul. 2017.

[73] X. Ma, F. Yang, S. Liu, J. Song, and Z. Han, “Design and optimization on training sequence for mmWave communications: A new approach for sparse channel estimation in massive MIMO,” IEEE J. Sel. Areas Commun., vol. 35, no. 7, pp. 1486–1497, Jul. 2017.

[74] C. Wang, C. Qin, Y. Yao, Y. Li, and W. Wang, “Low complexity interference alignment for mmWave MIMO channels in three cell mobile network,” *IEEE J. Sel. Areas Commun.*, vol. 5, no. 7, pp. 1513–1523, Jul. 2017.

[75] Z. Zhou, J. Fang, L. Yang, H. Li, Z. Chen, and R. S. Blum, “Low-rank tensor decomposition-aided channel estimation for millimeter wave MIMO-OFDM systems,” *IEEE J. Sel. Areas Commun.*, vol. 35, no. 7, pp. 1524–1538, Jul. 2017.

[76] Y. Yao, X. Cheng, C. Wang, J. Yu, and X. Chen, “Wideband circularly polarized antipodal curvedly tapered slot antenna array for 5G applications,” *IEEE J. Sel. Areas Commun.*, vol. 35, no. 7, pp. 1539–1549, Jul. 2017.

[77] Q. Xue, X. Fang, and C.-X. Wang, “Beamspace SU-MIMO for future millimeter wave wireless communications,” *IEEE J. Sel. Areas Commun.*, vol. 35, no. 7, pp. 1564–1575, Jul. 2017.

[78] [25] L. Zhao, D. W. K. Ng, and J. Yuan, “Multi-user precoding and channel estimation for hybrid millimeter wave systems,” *IEEE J. Sel. Areas Commun*., vol. 35, no. 7, pp. 1576–1590, Jul. 2017

[79] Sumsang Whitepaper, ' Samsung Unveils the World's First 5G FWA Commercial Solutions at MWC 2018', Feb.2018. [Online]. Available: https://www.samsung.com/global/business/networks/insights/news/samsung-unveils-the-worlds-first-5g-fwa-commercial-solutions-at-mwc-2018/C.[Accessed: 16- Aug- 2018]

[80] D.Bosnjak,” Google Can Now Test Millimeter-Wave Tech Throughout US ,” March 19, 2016. [Online] Available: https://www.androidheadlines.com/2016/03/google-can-now-test-millimeter-wave-tech-throughout-us.html. Accessed 16-Aug-18

[81] A.Tiwari,” Facebook demonstrates record-breaking data rate using mmW technology” Nov. 2016 [Online]. Available https://code.fb.com/connectivity/facebook-demonstrates-record-breaking-data-rate-using-millimeter-wave-technology/ Accesed:16-Aug-18

[82] M.Alleven, “Verizon applies for STA to test at 28 GHz in Texas, New Jersey ,” Aug 14, 2017,[Online] .Available : https://www.fiercewireless.com/wireless/verizon-applies-for-sta-to-test-at-28-ghz-texas-new-jersey Accessed:16-Aug-2018

[83] M.Alleven, ,”T-Mobile continues tests in 3.5 GHz CBRS band, seeks extension on STA in Vegas,” Mar 19, 2018 [Onlline] Available: https://www.fiercewireless.com/wireless/t-mobile-continues-tests-3-5-ghz-cbrs-band-seeks-extension-sta-vegas Accessed: 16-Aug-2018

[84] M.Polese et. al “Improved Handover Through Dual Connectivity in 5G mmWave Mobile Networks,” *IEEE Journal On Selected Areas In Communications*, Vol. 35, No. 9, pp 0733-8716 September 2017

[85] M.Alleven, “Huawei Launches 5G Simplified Solution ,” Feb 26, 2019 ,[Online] Available : https://www.huawei.com/en/press-events/news/2019/2/huawei-5g-simplified-solution Accessed:26-March-2019

[86] R. J. Weiler, W. Keusgen, I. Filippini, and A. Capone, ``Split control plane functionality in millimeter-wave overlay access," in *Proc. IEEE 1st Int. Conf. 5G Ubiquitous Connectivity (5GU)*, Nov. 2014, pp. 134-139

[87] C. Jonho and B. Lin, “On the communication range of millimetre wave small cell for energy efficiency of a Heterogeneous Network,” *Journal of communication and information networks,* vol.1, no. 1 Jun 2016,p.111-124.

[88] M. Pan et al., “Spectrum Harvesting and Sharing in Multi-Hop CRNs under Uncertain Spectrum Supply,” *IEEE JSAC,* vol. 30, no. 2, 2012, pp. 369–78.

[89] X. Huang et al., “Coolest Path: Spectrum Mobility Aware Routing Metrics In Cognitive Ad Hoc Networks,” *Proc. Int’l. Conf. Distrib. Computing Sys*., 2011, pp. 182–91.

[90] I. Christian et al., “Spectrum Mobility in Cognitive Radio Networks,” IEEE Commun. Mag., vol. 50, no. 6, June 2012, pp. 114–21.

[91] Y. Huang, H. Jiang, H. Hu, and Y. Yao, “Design of learning engine based on support vector machine in cognitive radio,” in *Computational Intelligence and Software Engineering, 2009. CiSE 2009. International Conference on. IEEE,* 2009, pp. 1–4.

[92] J. Qadir, “Artificial intelligence based cognitive routing for cognitive radio networks,” *Artificial Intelligence Review,* vol. 45, no. 1, pp. 25–96, 2016.

[93] N. Abbas, Y. Nasser, and K. El Ahmad, “Recent advances on artificial intelligence and learning techniques in cognitive radio networks,” *EURASIP Journal on Wireless Communications and Networking,* vol. 2015, no. 1, p. 174, 2015.

[94] C. Yang et al.,” Advanced Spectrum Sharing in 5G Cognitive Heterogeneous Networks,” *IEEE Wireless Comm.,* pp.94 – 101, April 2016

[95] M. Caleffi and L. Paura, “Bio-inspired link quality estimation for wireless mesh networks,” in *IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks & Workshops, 2009. WoWMoM 2009. IEEE,* 2009, pp. 1–6.

[96] X. Tan, H. Huang, and L. Ma, “Frequency allocation with artificial neural networks in cognitive radio system,” in *TENCON Spring Conference, 2013 IEEE. IEEE, 2013*, pp. 366–370.

[97] T. Zhang, M. Wu, and C. Liu, “Cooperative spectrum sensing based on artificial neural network for cognitive radio systems,” in *Wireless Communications, Networking and Mobile Computing (WiCOM), 2012 8th International Conference on. IEEE*, 2012, pp. 1–5.

[98] S. Fan, H. Tian, and C. Sengul, “Self-optimization of coverage and capacity based on a fuzzy neural network with cooperative reinforcement learning,” *EURASIP Journal on Wireless Communications and Networking*, vol. 2014, no. 1, pp. 1–14, 2014.

[99] N. Ahad, J. Qadir, and N. Ahsan, “Neural networks in wireless networks: Techniques, applications and guidelines,” in *Journal of Network and Computer Applications*, vol. 68, pp. 1–27, 2016.

[100] Z. Ding, P. Fan and H. V. Poor, "Random Beamforming in Millimeter-Wave NOMA Networks," in *IEEE Access*, vol. 5, pp. 7667-7681, 2016.

[101] Francesco Devoti et al., “Facing the Millimeter-Wave Cell Discovery Challenge in 5G Networks with Context-Awareness,” IEEE Access Special Section on Deployment and Management Of Small Heterogeneous Cells For 5g Volume 4, pp.819-834, Nov. 2016.

[102] Z. Ding, P. Fan, and H. V. Poor, “Impact of user pairing on 5G nonorthogonal multiple-access downlink transmissions,” IEEE Trans. Veh. Technol., vol. 65, no. 8, pp. 6010–6023, Aug. 2016.

[103] M. N. Kulkarni, A. Ghosh, and J. G. Andrews, ``A comparison of MIMO techniques in downlink millimeter wave cellular networks with hybrid beamforming," *IEEE Trans.* Commun., vol. 64, no. 5, pp. 1952-1967, May 2016.

[104] G. Lee, Y. Sung, and J. Seo, ``Randomly-directional beamforming in millimeter-wave multiuser MISO downlink," *IEEE Trans. Wireless Commun.*, vol. 15, no. 2, pp. 1086-1100, Feb. 2016.

[105] A. Capone, I. Filippini, and V. Sciancalepore, ``Context information for fast cell discovery in mm-Wave 5G networks," in *Proc. Eur. Wireless 21th Eur. Wireless Conf.*, 2015, pp. 1-6.

[106] M. R. Akdeniz et al., ``Millimeter wave channel modeling and cellular capacity evaluation," IEEE J. Sel. Areas Commun., vol. 32, no. 6, pp. 1164-1179, Jun. 2014

[107] G. Feng, E. Patelli & M. Beer," Survival signature-based sensitivity analysis of systems with epistemic uncertainties," *Institute for Risk and Uncertainty*, University of Liverpool, Liverpool, UK,2015, pp.1547 – 1552.

[108] A. Talukdar, M. Cudak, and A. Ghosh, ``Handoff rates for millimeterwave 5G systems," in *Proc. IEEE 79th Veh. Technol. Conf. (VTC Spring)*, May 2014, pp. 1_5.

[109] A. Ghosh, "The 5G mmWave radio revolution," *Microwave Journal.*, vol. 59, no. 9, pp. 22–36, Sep. 2016.

[110] GmbH, N. R. , GPP ,"LTE—A standardization in Release 12 and beyond," *White paper. Horizon 2020-Work Program 2014–2015. Leadership in enabling and industrial technologies information and communication tech,*2013 [Online]. Available: http://www.nomor.de/uploads/fd/24/fd24709a64bc490a083a8eba6d3cc2cb/NoMoR_LTE-A_Rel12_and_Beyond_2013-01.pdf. European Commission.

[111] S. Sun, M. Kadoch, and T. Ran, ``Adaptive SON and cognitive smart LPN for 5G heterogeneous networks," *Mobile Netw. Appl*., Jan. 2015, doi: 10.1007/s11036-014-0563-2

[112] M. Youssef, M. Ibrahim, M. Abdelatif, L. Chen, and A. V. Vasilakos, ``Routing metrics of cognitive radio networks: A survey," *IEEE Commun. Surveys Tuts*., vol. 16, no. 1, pp. 92-109, First Quarter 2014.

[113] S.-S. Byun, I. Balashingham, A. V. Vasilakos, and H.-N. Lee, ``Computation of an equilibrium in spectrum markets for cognitive radio networks," *IEEE Trans. Comput*., vol. 63, no. 2, pp. 304-316, Feb. 2014.

[114] T. Jiang, H. Wang, and A. V. Vasilakos, ``QoE-driven channel allocation schemes for multimedia transmission of priority-based secondary users over cognitive radio networks," *IEEE J. Sel. Areas Commun*., vol. 30, no. 7, pp. 1215-1224, Aug. 2012.

[115] Z. Zhang, K. Long, J. Wang, and F. Dressler, ``On swarm intelligence inspired self-organized networking: Its bionic mechanisms, designing principles and optimization approaches," *IEEE Commun. Surv. Tuts.*, vol. 16, no. 1, pp. 513-537, First Quarter 2014.

[116] O. G. Aliu, A. Imran, M. A. Imran, and B. Evans, ``A survey of self organisation in future cellular networks," *IEEE Commun. Surv. Tuts.*, vol. 15, no. 1, pp. 336-361, First Quarter 2013.

[117] Z. Zhang, W. Huangfu, K. Long, X. Zhang, X. Liu, and B. Zhong``On the designing principles and optimization approaches of bio-inspired self-organized network: A survey," *Sci. China Inf. Sci.*, vol. 56, no. 7, pp. 1-28, Jul. 2013.

[118] A. Imran, A. Zoha, and A. Abu-Dayya, ``Challenges in 5G: How to empower SON with big data for enabling 5G," *IEEE Netw.*, vol. 28, no. 6, pp. 27-33, Nov./Dec. 2014.

[119] *HetNet/Small Cells*. [Online]. Available: http://www.3gpp.org/hetnet, accessed Aug. 2014.

[120] X. Su, J. Zeng, and C. Xiao, ``Key technologies for SON in next generation radio access networks," in *Proc. 23rd ICCCN*, Aug. 2014, pp. 1-8.

[121] M. Peng, D. Liang, Y. Wei, J. Li, and H.-H. Chen, ``Self-configuration and self-optimization in LTE-advanced heterogeneous networks," *IEEE Commun. Mag.*, vol. 51, no. 5, pp. 36_45, May 2013. L. T. W. Ho, I. Ashraf, and H. Claussen, ``Evolving femtocell coverage optimization algorithms using genetic programming," in *Proc. IEEE 20th Int. Symp. PIMRC*, Sep. 2009, pp. 2132-2136.

[122] S. Sun, M. Kadoch, and T. Ran, ``Adaptive SON and cognitive smart LPN for 5G heterogeneous networks," *Mobile Netw. Appl.*, Jan. 2015, doi: 10.1007/s11036-014-0563-2

[123] W. Guo and T. O'Farrell, ``Relay deployment in cellular networks: Planning and optimization," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 8, pp. 1597-1606, Aug. 2013.

[124] C.-S. Tseng, C.-C. Chen, and C. Lin, ``A bio-inspired robust adaptive random search algorithm for distributed beamforming," in *Proc. IEEE ICC*, Jun. 2011, pp. 1-6.

[125] K. M. Passino, ``Biomimicry of bacterial foraging for distributed optimization and control," *IEEE Control Syst. Mag.*, vol. 22, no. 3, pp. 52-67, Jun. 2002.

[126] W. Li, P. Yu, Z. Jiang, and Z. Li, ``Centralized management mechanism for cell outage compensation in LTE networks," *Int. J. Distrib. Sensor Netw.*, vol. 2012, pp. 1_8, Jul. 2012.

[127] Z. Zhao, J. Chen, and N. Crespi, ``Apolicy-based framework for autonomic reconfiguration management in heterogeneous networks," in *Proc. 7th Int. Conf. Mobile Ubiquitous Multimedia*, 2008, pp. 71-78.

[128] S. Fan, H. Tian, and C. Sengul, ``Self-optimization of coverage and capacity based on a fuzzy neural network with cooperative reinforcement learning," *EURASIP J. Wireless Commun. Netw.*, vol. 2014, no. 1, pp. 57:1_57:14, Apr. 2014.

[129] M. Calef_ and L. Paura, ``Bio-inspired link quality estimation for wireless mesh networks," in *Proc. IEEE Int. Symp. WoWMoM*, Jun. 2009, pp. 1-6.

[130] A.-H. Tsai, L.-C.Wang, J.-H. Huang, and T.-M. Lin, ``Intelligent resource management for device-to-device (D2D) communications in heterogeneous networks," in *Proc. 15th Int. Symp. WPMC*, Sep. 2012, pp. 75-79.

[131] A. Engels, M. Reyer, X. Xu, R. Mathar, J. Zhang, and H. Zhuang, ``Autonomous self-optimization of coverage and capacity in LTE cellular networks," *IEEE Trans. Veh. Technol.*, vol. 62, no. 5, pp. 1989-2004, Jun. 2013.

[132] M. Peng, Y. Liu, D. Wei, W. Wang, and H.-H. Chen, ``Hierarchical cooperative relay based heterogeneous networks," *IEEE Wireless Commun.*, vol. 18, no. 3, pp. 48_56, Jun. 2011.

[133] M. Liu, Z. Li, X. Guo, and E. Dutkiewicz, ``Performance analysis and optimization of handoff algorithms in heterogeneous wireless networks," *IEEE Trans. Mobile Comput.*, vol. 7, no. 7, pp. 846_857, Jul. 2008.

[134] R. Chai, J. Cheng, X. Pu, and Q. Chen, ``Neural network based vertical handoff performance enhancement in heterogeneous wireless networks," in *Proc. WiCOM*, Sep. 2011, pp. 1-4.

[135] A. Çalhan and C. Ceken, ``An adaptive neuro-fuzzy based vertical handoff decision algorithm for wireless heterogeneous networks," in *Proc. IEEE 21st Int. Symp. PIMRC*, Sep. 2010, pp. 2271-2276.

[136] B. Ma and X. Liao, ``Speed-adaptive vertical handoff algorithm based on fuzzy logic in vehicular heterogeneous networks," in *Proc. 9th Int. Conf. FSKD*, May 2012, pp. 371_375.

[137] S. B. Kotsiantis, ``Supervised machine learning: A review of classification techniques," *Informatica*, vol. 31, no. 3, pp. 249_268, 2007.

[138] L. P. Kaelbling, M. L. Littman, and A. W. Moore, ``Reinforcement learning: A survey," *J. Artif. Intell. Res.*, vol. 4, pp. 237-285, 1996.

[139] M. A. Khan, H. Tembine, and A. V. Vasilakos, ``Game dynamics and cost of learning in heterogeneous 4G networks," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 1, pp. 198_213, Jan. 2012.

[140] C. Watkins and P. Dayan, "Q-learning," *Mach. Learn.*, vol. 8, nos. 3-4, pp. 279_292, 1992.

[141] P. T. Semov, A. Mihovska, R. Prasad, and V. Poulkov, ``Use of positioning information for performance enhancement of uncoordinated heterogeneous network deployment," in *Proc. 3rd Int. Conf. VITAE,* Jun. 2013, pp. 1-6.

[142] J. Park, S. Y. Jung, S. -L. Kim, M. Bennis and M. Debbah, "User-Centric Mobility Management in Ultra-Dense Cellular Networks under Spatio-Temporal Dynamics," *2016 IEEE Global Communications Conference (GLOBECOM),* 2016, pp. 1-6, doi: 10.1109/GLOCOM.2016.7842367.

[143] H. Munir *et al*., "Computationally Intelligent Techniques for Resource Management in MmWave Small Cell Networks," in *IEEE Wireless Communications*, vol. 25, no. 4, pp. 32-39, AUGUST 2018, doi: 10.1109/MWC.2018.1700400.

[144] N. Shlezinger, N. Farsad, Y. C. Eldar and A. J. Goldsmith, "ViterbiNet: A Deep Learning Based Viterbi Algorithm for Symbol Detection," in IEEE Transactions on Wireless Communications, vol. 19, no. 5, pp. 3319-3331, May 2020, doi: 10.1109/TWC.2020.2972352.

[145] T.T. Le, F.Chatelain and C.Bérenguer," Jump Markov Linear Systems for deterioration modeling and Remaining Useful Life estimation," *Safety and Reliability of Complex Engineered Systems,* Taylor & Francis Group, London,2015.

[146] J. S. Lu, D. Steinbach, P. Cabrol, and P. Pietraski, ''Modeling human blockers in millimeter wave radio links,*'' ZTE Commun.,* vol. 10, no. 4, pp. 23–28, Dec. 2012.

[147] S. Chia, M. Gasparroni, and P. Brick, ''The next challenge for cellular networks: Backhaul,'' *IEEE Microw. Mag.,* vol. 10, no. 5, pp. 54–66, May 2009.

[148] T. Cho, D. Cline, C. Conroy, and P. Gray, ''Design considerations for low-power, high speed CMOS analog/digital converters,'' *in Proc. IEEE Symp. Low Power Electron*., Oct. 1994, pp. 70–73.

[149] J. Murdock and T. S. Rappaport, ‘‘Consumption factor and power-efficiency factor: A theory for evaluating the energy efficiency of cascaded communication systems,’’ *IEEE J. Sel. Areas Commun.*, 2013, DOI: 10.1109/JSAC.2014.141204.

[150] C.-Y. Chen, J. Wu, J.-J. Hung, T. Li, W. Liu, and W.-T. Shih, ‘‘A 12-bit 3 GS/s pipeline ADC with 0.42mm and 500 mW in 40 nm digital CMOS,’’ *IEEE J. Solid-State Circuits*, vol. 47, no. 4, pp. 1013–1021, Apr. 2011.

[151] H.-L. Park, Y.-G. Kwon, M.-H. Choi, Y. Kim, S.-H. Lee, Y.-D. Jeon, and J.-K. Kwon, ‘‘A 6b 1.2 GS/s 47.8 mW 0.17 mm2 65 nm CMOS ADC for high-rate WPAN systems,’’ *J. Semicond. Sci. Technol.*, vol. 11, no. 2, pp. 95–103, Jun. 2011.

[152] M. Abouelseoud and G. Charlton, ‘‘System level performance of millimeter-wave access link for outdoor coverage,’’ *in Proc. IEEE Wireless Commun. Netw. Conf.*, 2013, pp. 4146–4151.

[153] A. Ghosh, ‘‘Can mmwave wireless technology meet the future capacity crunch?’’ *IEEE Int. Conf. Commun.*, Budapest, Hungary, Jun. 2013. [Online]. Available: http://www. ieee-icc.org/2013/Mmwave_Spring_ ICC2013_Ghosh.pdf

[154] Qualcomm, LTE Advanced: Heterogeneous networks, White Paper, Jan. 2011. [Online]. Available: http://www.qualcomm.com/ media/documents

[155] A. Damnjanovic, J. Montojo, Y. Wei, T. Ji, T. Luo, M. Vajapeyam, T. Yoo, O. Song, and D. Malladi, ‘‘A survey on 3GPP heterogeneous networks,’’ *IEEE Wireless Commun.*, vol. 18, no. 3, pp. 10–21, Jun. 2011.

[156] Qualcomm, Neighborhood small cells for hyperdense deployments: Taking HetNets to the next level, Feb. 2013. [Online]. Available: http://www.qualcomm.com/ media/documents/files/qualcomm-researchneighborhood- small-cell-deploymentmodel. pdf

[157] R. Ford, C. Kim, and S. Rangan, ‘‘Optimal user association in cellular networks with opportunistic third-party backhaul,’’ Apr. 2013, arXiv preprint.

[158] T. Zwick, T. Beukema, and H. Nam, ‘‘Wideband channel sounder with measurements and model for the 60 GHz indoor radio channel,’’ IEEE Trans. Veh. Technol., vol. 54, no. 4, pp. 1266–1277, Jul. 2005.

[159] F. Giannetti, M. Luise, and R. Reggiannini, ‘‘Mobile and personal communications in 60 GHz band: A survey,’’ *Wireless Pers. Commun.*, vol. 10, pp. 207–243, 1999.

[160] P. Smulders and A. Wagemans, ‘‘Wideband indoor radio propagation measurements at 58 GHz,’’ *Electron. Lett.*, vol. 28, no. 13, pp. 1270–1272, Jun. 1992.

[161] T. Manabe, Y. Miura, and T. Ihara, ‘‘Effects of antenna directivity and polarization on indoor multipath propagation characteristics at 60 GHz,’’ *IEEE J. Sel. Areas Commun.*, vol. 14, no. 3, pp. 441–448, Apr. 1996.

[162] H. Xu, V. Kukshya, and T. S. Rappaport, ‘‘Spatial and temporal characteristics of 60 GHz indoor channel,’’ *IEEE J. Sel. Areas Commun.*, vol. 20, no. 3, pp. 620–630, Apr. 2002.

[163] A. Elrefaie and M. Shakouri, ‘‘Propagation measurements at 28 GHz for coverage evaluation of local multipoint distribution service,’’ *in Proc. Wireless Commun. Conf.,* Aug. 1997, pp. 12–17.

[164] S. Seidel and H. Arnold, ‘‘Propagation measurements at 28 GHz to investigate the performance of local multipoint distribution service (LMDS),*’’ in Proc. Wireless Commun. Conf.*, Nov. 1995, pp. 754–757.

[165] T. S. Rappaport, E. Ben-Dor, J. Murdock, and Y. Qiao, ‘‘38 GHz and 60 GHz angle-dependent propagation for cellular & peer-to-peer wireless communications,’’ *in Proc. IEEE Int. Conf. Commun.*, Jun. 2012, pp. 4568–4573.

[166] T. S. Rappaport, E. Ben-Dor, J. Murdock, Y. Qiao, and J. Tamir, ‘‘Cellular broadband millimeter wave propagation and angle of arrival for adaptive beam steering systems,’’ *in Proc. IEEE Radio Wireless Symp.*, Jan. 2012, pp. 151–154.

[167] J. Murdock, E. Ben-Dor, Y. Qiao, J. Tamir, and T. S. Rappaport, ‘‘A 38 GHz cellular outage study for an urban outdoor campus environment,’’ *in Proc. IEEE Wireless Commun. Netw. Conf., Apr*. 2012, pp. 3085–3090.

[168] G. R. MacCartney, Jr., J. Zhang, S. Nie, and T. S. Rappaport, ‘‘Path loss models for 5G millimeter wave propagation channels in urban microcells,’’ *in Proc. IEEE Global Telecommun. Conf.,* Dec. 2013.

[169] J. A. Wells, ‘‘Faster than fiber: The future of multi-Gb/s wireless,’’ IEEE Microw. Mag., vol. 10, no. 3, pp. 104–112, May 2009.

[170] 3GPP, ‘‘Further advancements for E-UTRA physical layer aspects,’’ TR 36.814 (release 9), 2010.

[171] S. Pinel, S. Sarkar, P. Sen, B. Perumana, D. Yeh, D. Dawn, and J. Laskar, ‘‘A 90 nm CMOS 60 GHz radio,’’ *in Proc. IEEE Int. Solid-State Circuits Conf.,* 2008, DOI: 10.1109/ISSCC.2008.4523091.

[172] C. Marcu, D. Chowdhury, C. Thakkar, J.-D. Park, L.-K. Kong, M. Tabesh, Y. Wang, B. Afshar, A. Gupta, A. Arbabian,S. Gambini, R. Zamani, E. Alon, and A. Niknejad, ‘‘A 90 nm CMOS low-power 60 GHz transceiver with integrated baseband circuitry,’’ *IEEE J. Solid-State Circuits*, vol. 44, no. 12, pp. 3434–3447, Dec. 2009.

[173] A. Alkhateeb, O. E. Ayach, G. Leus, J. Robert, and W. Heath, ‘‘Hybrid precoding for millimeter wave cellular systems with partial channel knowledge,’’ in *Proc. Inf. Theory Appl. Workshop,* Feb. 2013, DOI: 10.1109/ITA.2013.6522603.

[174] A. Lozano, ‘‘Long-term transmit beamforming for wireless multicasting,’’ in *Proc. Int. Conf. Acoust. Speech Signal Process.*, 2007, vol. 3, pp. III-417–III-420.

[175] Q. Li, H. Niu, G. Wu, and R. Q. Hu, ‘‘Anchor-booster based heterogeneous networks with mmWave capable booster cells,’’ *in Proc. IEEE Globecom Workshop,* Dec. 2013.

[176] C. Bontu and E. Illidge, ‘‘DRX mechanism for power saving in LTE,’’ IEEE Commun. Mag., vol. 47, no. 6, pp. 48–55, Jun. 2009.

[177] D. Parker and D. Z. Zimmermann, ‘‘Phased arraysVPart I: Theory and architecture,’’ *IEEE Trans. Microw. Theory Tech.,* vol. 50, no. 3, pp. 678–687, Mar. 2002.

[178] K.-J. Koh and G. M. Rebeiz, ‘‘0.13-m CMOS phase shifters for X-, Ku- and K-band phased arrays,’’ *IEEE J. Solid-State Circuits*, vol. 42, no. 11, pp. 2535–2546, Nov. 2007. Vol. 102, No. 3, March 2014 | Proceedings of the IEEE 383

[179] K.-J. Koh and G. M. Rebeiz, ‘‘A millimeter wave (4045 GHz) 16-element phased-array transmitter in 0.18-m SiGe BiCMOS technology,*’’ IEEE J. Solid-State Circuits*, vol. 44, no. 5, pp. 1498–1509, May 2009.

[180] P. E. Crane, ‘‘Phased array scanning system,’’ U.S. Patent 4 731 614, Mar. 15, 1988.

[181] S. Raman, N. S. Barker, and G. M. Rebeiz, ‘‘A W-band dielectric-lens-based integrated monopulse radar receive,’’ *IEEE Trans. Microw. Theory Tech.*, vol. 46, no. 12, pp. 2308–2316, Dec. 1998.

[182] X. Guan, H. Hashemi, and A. Hajimiri, ‘‘A fully integrated 24-GHz eight-element phased-array receiver in silicon,’’ *IEEE J. Solid-State Circuits*, vol. 39, no. 12, pp. 2311–2320, Dec. 2004.

[183] G. Fettweis, F. Guderian, and S. Krone, ‘‘Entering the path towards terabit/s wireless links,’’ in Proc. IEEE Design Autom. Test Eur., 2011, DOI: 10.1109/DATE.2011.5763075.

[184] H. Zhang, S. Venkateswaran, and U. Madhow, ‘‘Analog multitone with interference suppression: Relieving the ADC bottleneck for wideband 60 GHz systems,’’ *in Proc. IEEE Global Telecommun. Conf.,* Nov. 2012, pp. 2305–2310.

[185] S. Tavildar, S. Shakkottai, T. Richardson, J. Li, R. Laroia, and A. Jovicic, ‘‘FlashLinQ: A synchronous distributed scheduler for peer-to-peer ad hoc networks,’’ *in Proc. Allerton Conf. Commun. Control Comput.,* Allerton, IL, USA, Oct. 2010, pp. 514–521.

[186] A. So and B. Liang, ‘‘Effect of relaying on capacity improvement in wireless local area networks,’’ *in Proc. IEEE Wireless Commun. Netw. Conf.,* Mar. 2005, vol. 3, pp. 1539–1544.

[187] R. Schoenen, W. Zirwas, and B. H. Walke, ‘‘Capacity and coverage analysis of a 3GPP-LTE multihop deployment scenario,’’ *in Proc. IEEE Int. Conf. Commun*., May 2008, pp. 31–36.

[188] A. B. Saleh, S. Redana, B. Raaf, and J. Hamalainen, ‘‘Comparison of relay and Pico eNB deployments in LTE-advanced,’’ *in Proc. IEEE Veh. Technol. Conf.,* Sep. 2009, DOI: 10.1109/VETECF.2009.5378828.

[189] S. W. Peters, A. Y. Panah, K. T. Truong, and R. W. Heath, ‘‘Relay architectures for 3GPP LTE-advanced,’’ *EURASIP J. Wireless Commun. Netw.*, 2009, DOI: 10.1155/2009/618787.

[190] G. D. Durgin and T. S. Rappaport, ‘‘Theory of multipath shape factors for small-scale fading wireless channels,’’ *IEEE Trans. Antennas Propag.*, vol. 48, no. 5, pp. 682–693, May 2000.

[191] M. Vetterli, P. Marziliano, and T. Blu, ‘‘Sampling signals with finite rate of innovation,’’ *IEEE Trans. Signal Process.*, vol. 50, no. 6, pp. 1417–1428, Jun. 2002.

[192] G. Taubock and F. Hlawatsch, ‘‘A compressed sensing technique for OFDM channel estimation in mobile environments: Exploiting channel sparsity for reducing pilots,’’ *in Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, May 2008, pp. 2885–2888.

[193] J. Haupt, W. U. Bajwa, G. Raz, and R. Nowak, ‘‘Toeplitz compressed sensing matrices with applications to sparse channel estimation,’’ *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5862–5875, Nov. 2010.

[194] Y. Barbotin, A. Hormati, S. Rangan, and M. Vetterli, ‘‘Estimating sparse MIMO channels having common support,’’ *in Proc. IEEE Int. Conf. Acoust. Speech Signal Process*., May 2011, pp. 2920–2923.

[195] J. I. Tamir, T. S. Rappaport, Y. C. Eldar, and A. Aziz, ‘‘Analog compressed sensing for RF propagation channel sounding,’’ *in Proc. IEEE Int. Conf. Acoust. Speech Signal Process*., 2012, pp. 5317–5320.

[196] 3GPP, ‘‘Evolved universal terrestrial radio access (E-UTRA) and evolved universal terrestrial radio access network (E-UTRAN); overall description; stage 2,’’ TS 36.300 (release 10), 2010.

[197] G. Yuan, X. Zhang, W. Wang, and Y. Yang, ‘‘Carrier aggregation for LTE-advanced mobile communication systems,’’ IEEE Commun. Mag., vol. 48, no. 2, pp. 88–93, Feb. 2010.