Radio Resource Allocation and Beam Management under Location Uncertainty in 5G mmWave Networks

by

Yujie Yao

A thesis submitted to the University of Ottawa
in partial fulfillment of the requirement for the degree of

Master of Applied Science in Electrical and Computer Engineering

Abstract

Millimeter Wave (mmWave) plays a critical role in the Fifth-generation (5G) new radio due to the rich bandwidth it provides. However, one shortcoming of mmWave is the substantial path loss caused by poor diffraction at high frequencies, and consequently highly directional beams are applied to mitigate this problem. A typical way of beam management is to cluster users based on their locations. However, localization uncertainty is unavoidable due to measurement accuracy, system performance fluctuation, and so on. Meanwhile, the traffic demand may change dynamically in wireless environments, which increases the complexity of network management. Therefore, a scheme that can handle both the uncertainty of localization and dynamic radio resource allocation is required. Moreover, since the localization uncertainty will influence the network performance, more state-of-the-art localization methods, such as vision-aided localization, are expected to reduce the localization error.

In this thesis, we proposed two algorithms for joint radio resource allocation and beam management in 5G mmWave networks, namely UK-means-based Clustering and Deep Reinforcement Learning-based resource allocation (UK-DRL) and UK-medoids-based Clustering and Deep Reinforcement Learning-based resource allocation (UKM-DRL). Specifically, we deploy UK-means and UK-medoids clustering method in UK-DRL and UKM-DRL, respectively, which is designed to handle the clustering under location uncertainties. Meanwhile, we apply Deep Reinforcement Learning (DRL) for intra-beam radio resource allocations in UK-DRL and UKM-DRL. Moreover, to improve the localization accuracy, we develop a vision-aided localization scheme, where pixel characteristics-based features are extracted from satellite images as additional input features for location prediction.

The simulations show that UK-DRL and UKM-DRL successfully improve the network performance in data rate and delay than baseline algorithms. When the traffic load is 4 Mbps, UK-DRL has a 172.4% improvement in sum rate and 64.1% improvement in latency than K-means-based Clustering and Deep Reinforcement Learning-based resource allocation (K-DRL). UKM-DRL has 17.2% higher throughput and 7.7% lower latency than UK-DRL, and 231% higher throughput and 55.8% lower latency than K-DRL. On the other hand, the vision-aided localization scheme can significantly reduce the localization error from 17.11 meters to 3.6 meters.

To

my beloved parents,
Shuhua Xiong and Aiping Yao

Acknowledgements

Firstly, I would like to express my particular gratitude to my respectable supervisor, Professor Erol-Kantarci, for providing me with direction and encouragement during my research period. Her patient instruction and constructive suggestions have been really valuable to me. She has walked me through all the stages of the research. Without her consistent and illuminating instruction, this thesis could not have reached its present form.

Secondly, I would like to acknowledge the other members of the Networked systems and Communications Research Lab (NETCORE). Especially, I would like to thank Hao Zhou, Medhat Elsayed, and Hind Mukhtar for their wise counsel during my thesis.

Moreover, I would like to thank the University of Ottawa for offering me a platform to invest my interest and achieve self-development. Particular thanks go to all the professors who have taught me for their instruction and generous support during these years.

In the end, I would like to express the most heartfelt gratitude to my parents who have provided me with great help and support.

Table of Contents

List of Tables viii

List of Figures ix

List of Abbreviations xi

List of Symbols xiv

1 Introduction 1

1.1 Motivation 1

1.2 Thesis Contributions 3

1.3 Organization of the Thesis 3

2 Background 5

2.1 5G New Radio 5

2.2 Localization 7

2.3 Beam Management 8

2.4 Radio Resource Management 8

2.4.1 Spectrum and Power Allocation 8

2.4.2 Beamforming and User Clustering 10

2.5 Machine Learning Techniques 11

2.5.1 Clustering 11

2.5.2 Neural Network . . . 14
2.5.3 Convolutional Neural Network . . . 15
2.5.4 Long-short Term Memory Neural Network . . . 16
2.5.5 Reinforcement Learning . . . 18
2.5.6 Deep Reinforcement Learning . . . 19

3 Literature Review **22**
3.1 Clustering-based Network Management . . . 22
3.2 Reinforcement Learning-based Network Management . . . 25
3.3 Localization in Wireless Networks . . . 28
3.3.1 Machine Learning-based Localization . . . 28
3.3.2 Vision Aided-Localization . . . 29
3.4 Overview . . . 31

4 Methodology **32**
4.1 Datasets for Localization . . . 33
4.1.1 Raymobtime Dataset . . . 33
4.1.2 Mobile Communication System Measurements and Satellite Images Dataset . . . 35
4.2 System Model and Proposed Techniques . . . 37
4.2.1 System Model . . . 37
4.2.2 Vision-aided Localization . . . 39
4.2.3 UK-means-based User Clustering . . . 47
4.2.4 UK-medoids-based User Clustering . . . 49
4.2.5 Deep Reinforcement Learning-based Radio Resource Allocation . . . 51

5 Results **55**

5.1 UK-means-based Clustering and Deep Reinforcement Learning-based Radio Resource Allocation . . . 55

5.1.1 Simulation Settings . . . 55

5.1.2 Results . . . 60

5.2 UK-medoids based Clustering and Deep reinforcement learning based radio resource allocation . . . 64

5.2.1 Simulation Settings . . . 64

6 Conclusion and Future Work **75**

6.1 Conclusion . . . 75

6.2 Future Work . . . 76

List of Tables

4.1 Information Provided by Raymobtime 35
4.2 Mean value of Color Characteristics 41
4.3 Range of Three Categories 42

5.1 Simulation Parameters 57
5.2 Simulation settings of FFNN 65
5.3 Simulation Parameters 66
5.4 Measurements of Localization Results 68

List of Figures

2.1 The architecture of 5G cellular network 6
2.2 Resource grid of 5G networks 9
2.3 Structure of a neuron. 13
2.4 Structure of a two-layered feed-forward network 15
2.5 Structure of a simple CNN 16
2.6 The structure of LSTM 17
2.7 Standard reinforcement learning model 18
2.8 Architecture of deep reinforcement learning method 20

3.1 System model of [41] 23
3.2 System model of [58] 24
3.3 System model of [61] 26
3.4 Framework of neural network in [62] 27
3.5 Diagram of visual localization algorithm [25] 30

4.1 Overall System Architectures 32
4.2 Rosslyn map [80] 34
4.3 Structure of Raymobtime. 34
4.4 Location distribution of mobile communication system measurements and satellite images dataset 36
4.5 Network model 38

4.6 Two examples of provided satellite images 40
4.7 Three selected objects for characterising an image 41
4.8 Histograms of $|R-B|$ for three objects 43
4.9 Histograms of $|R-G|$ for three objects 44
4.10 Histograms of $|G-B|$ for three objects 45
4.11 Histograms of $|2G-R-B|$ for three objects 46
4.12 Diagram of LSTM-based Deep Q-learning. 52

5.1 Distribution of test data 58
5.2 Distribution of predicted data 58
5.3 Distribution of selected data 59
5.4 Coverage rate comparison under varying number of beams 61
5.5 Coverage rate comparison under varying beam angle 62
5.6 Sum rate comparison under varying traffic loads 63
5.7 Delay comparison under varying traffic loads 64
5.8 Coverage rate comparison under varying beam angle. 69
5.9 Coverage rate comparison under varying beam angle 69
5.10 Sum rate comparison under various traffic loads 72
5.11 Delay comparison under various traffic loads 73

List of Abbreviations

Acronym	Meaning
4G	Fourth-generation
5G	Fifth-generation
AoA	Angle of Arrival
AoD	Angle of Departure
BDMA	Beam Division Multiple Access
BS	Base Stations
BSs	Base Stations
CNN	Convolutional Neural Network
CQI	Channel Quality Indicator
D2D	Device-to-Device
DBSCAN	Density-Based Spatial Clustering of Applications with Noise
DRL	Deep Reinforcement Learning
EM	Expectation Maximization
eMBB	enhanced Mobile Broad-Brand
FFNN	Feed-forward Neural Network
gNB	5G-NodeB
GPS	Global Positioning System
INS	Inertial Navigation System
IoT	Internet of Things
IoV	Internet of Vehicles
K-DRL	K-means-based Clustering and Deep Reinforcement Learning-based resource allocation
LOS	Line-of-Sight
LSTM	Long-short Term Memory
MAC	Media Access Control
MAE	Mean Absolute Error

MDP	Markov Decision Process
MIMO	Multi-input and Multi-output
ML	Machine Learning
mMTC	massive Machine Type Communication
mmWave	Millimeter Wave
MSE	Mean Square Error
NFV	Network Function Virtualization
NLOS	Non-line-of-sight
NOMA	Non-orthogonal Multiple Access
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
PDF	Probability Density Function
QCI	Quality Class Indicator
QoS	Quality of Service
R-CNN	Recurrent Convolutional Neural Network
RB	Resource Block
RF	Radio Frequency
RL	Reinforcement Learning
RMSE	Root Mean Squared Error
RNNs	Recurrent Neural Networks
RRM	Radio Resource Management
RSRP	Reference Signal Received Power
RSRQ	Reference Signal Received Quality
RSS	Received Signal Strength
RSSI	Received signal strength indication
SINR	Signal to Interference and Noise Ratio
SSIM	Structure Similarity Index Matrices
TDD	Time Division Duplex
ToA	Time of Arrival
TTI	Transmission Time Interval
UAVs	Unmanned Aerial Vehicles
UE	User Equipment
UEs	User Equipments
UGVs	Unmanned Ground Vehicles
UK-DRL	UK-means-based Clustering and Deep Reinforcement Learning-based resource allocation
UKM-DRL	UK-medoids-based Clustering and Deep Reinforcement Learning-based resource allocation

ULA	Uniform Linear Array
URLLC	ultra Reliable and Low Latency Communication

List of Symbols

α	The complex gain of a path
$\hat{\boldsymbol{y}}$	Output vector of a neural network
$\boldsymbol{a}$	Antenna array response
$\boldsymbol{B}$	Set of beams
$\boldsymbol{b}$	Bias for a neuron
$\boldsymbol{C}$	Set of clusters
$\boldsymbol{H}$	The mmWave channel between gNB and UE
$\boldsymbol{h}$	The channel model vector between gNB and UE
$\boldsymbol{p}$	A record of UE location
$\boldsymbol{U}$	Set of UEs
$\boldsymbol{x}$	Input vector of a neural network
$\boldsymbol{y}$	Exact data vector from a database
δ	Uncertain distance function
Δ_1	Difference between two latitudes
Δ_2	Difference between two latitudes
Δd	Distance between two points
ϵ	Exploration probability

Γ	Training interval of a nueral network
γ	Discount factor
κ	State-pair function
Λ	Longitude
λ	Wavelength
Λ_0	Central longitude
ω	Weights of a neural network
ϕ	Angle of departure
ϕ	Latitude
Π	Policy
Θ	Angle of arrival
θ	Any value of angle within uncertainty region
ξ	Pass loss component
a	Action
b	Beam
C	Cluster
c	Center of a cluster
d	Distance between gNB's antennas
Er	Error function
i	Indicator of UEs
j	Indicator of clusters/beams
K	The number of path between gNB and UE
k	Indicator of resource block group

L	The total number of paths between gNB and UE
l	Indicator of path between gNB and UE
L_ω	Gradient descent algorithm
M	The number of user locations
m, n	Indicators of two different locations
N	The number of formed clusters/beams
n_d	The number of data in a dataset
N_t	The number of transmission antennas
Q	Q-function
q	CQI feedback
R	Uncertainty region
r	Reward
R_d	Radius of uncertainty region
r_d	Any value of radius within uncertainty region
R_E	Radius of the Earth
S	SINR
s	State
S^{QoS}	SINR QoS requirement
t	Time step
T^{QoS}	Delay QoS requirement
T_u^q	Delay
U	UE
u	Indicator of UE

V	Expected discounted reward
X	Number of rows of an image during sampling process
x	x-axis value
Y	Number of columns of an image during sampling process
y	y-axis value

Chapter 1

Introduction

1.1 Motivation

5G network is distinguished by multi-scenario applications, a huge number of connected devices, and tremendous data volume. Especially, with a significant amount of connected devices, 5G network is expected to support higher user density, lower latency, and higher spectrum efficiency [1, 2]. MmWave communication is developed to overcome various deterrents to obstacles of mobile networks, especially the high data rate challenge and low latency prerequisite [3]. Compared with widely used frequency bands, MmWave bands use the frequencies between 30 and 300 GHz and offer rich bandwidth and lower interference between adjoining Base Stations (BS). Nonetheless, one crucial deficiency of MmWave is the path loss resulted by Non-line-of-sight (NLOS) transmission [4].

To this end, highly directional beams which can furnish large antenna array gains with less inter-beam interference are adopted to tackle this problem. Beam Division Multiple Access (BDMA) is generally adopted to divide the antenna beams according to the location of User Equipment (UE), prompting an expansion in system capacity [5]. A typical way of beam management is to have all users divided into numerous clusters, and each cluster is served by a separate beam [6]. Power is apportioned to each beam and the radio resource is allocated within the beam. Therefore, user location turns into a key pre-requisite in beam management and Radio Resource Management (RRM).

A popular technique for outdoor localization is Global Positioning System (GPS). GPS relies on the received signals from satellites for localization. In some wireless networks, GPS is applied for offering position information for more advantage. For example, position

information is required to perform handover for mobile users to guarantee uninterrupted service [7]. However, the application of GPS is restricted to Line-of-Sight (LOS) circumstances, and it becomes ineffective in indoor scenarios, urban canyons, underground, and other environments. Over these years, efforts were made to apply machine learning on Radio Frequency (RF) signals as an alternative to GPS technology in wireless communications [8]. Recently, computer vision has been introduced to supplement localization in wireless communications, and visual information can overcome issues of classic localization procedures [9].

On the other hand, although some localization techniques have been proposed to replace GPS, localization error is still unavoidable in practice, which may be caused by measurement accuracy, system performance fluctuation, and so on. Localization uncertainty will diminish the performance of the network. For instance, assuming the BS detects the distorted location of UE and performs clustering based on distorted location. Since the direction of beams are formed according to the distorted location, a few User Equipments (UEs) may not be covered by the highly directional beams from BS when the exact location is considered, which may bring higher packet loss and longer delay. As a result, a clustering method that has the ability to handle localization uncertainty is required.

Moreover, wireless networks nowadays are faced with heterogeneous service requirements and dynamic traffic demands. Three typical types of services are included in 5G network: Enhanced Mobile Broad-Brand (eMBB), Ultra Reliable and Low Latency Communication (URLLC), and Massive Machine Type Communication (mMTC) [10]. The Quality of Service (QoS) requirements of different types of users are diverse. EMBB users need high throughput transmission, URLLC users require extremely low latency and high reliability, and mMTC supports connections for a large number of devices. To this end, a radio resource allocation scheme that can intelligently satisfy the QoS requirements of different types of users is expected, and Machine Learning (ML) has become a promising solution [11]. The recent successes of ML in a variety of domains demonstrate its ability to successfully tackle complicated problems that have long eluded deterministic techniques. Moreover, the process of radio resource allocation can be formulated as ML problems [12]. Specifically, Reinforcement Learning (RL) is applied to exploit the wireless environment without any prior knowledge, which avoids the complexity of building a dedicated optimization model. Moreover, combined with a deep neural network, DRL overcomes the limitations of tabular-based RL, which has been generally applied for wireless network optimizations [13].

1.2 Thesis Contributions

In this thesis, two RRM schemes for handling both the localization uncertainty and different service types of users are proposed. Two real datasets are applied in simulation to examine the effect of proposed schemes. For the second dataset, a vision-aided localization technique is proposed to increase localization accuracy. The localization technique is a substitute under outdoor localization scenarios when the performance of GPS is limited. Especially in urban canyons where a large amount of tall buildings cause the degrade of GPS signal. The contributions are summarized as follows.

- A UK-means-based clustering and DRL-based radio resource allocation technique named UK-DRL is proposed as the first scheme for joint beam management and resource allocation [14] under localization uncertainty. UK-DRL can mitigate the localization uncertainty problem by utilizing Probability Density Function (PDF). It can improve the sum rate of a wireless network and reduce the latency.
- A UK-medoids-based clustering and DRL-based radio resource allocation technique named UKM-DRL is proposed as the second scheme for joint beam management and resource allocation [15] under localization uncertainty. UKM-DRL can mitigate the localization uncertainty problem by utilizing uncertainty distance function between two uncertain objects. Compared with UK-DRL, UKM-DRL proves a significant benefit in improving the network performance in the case of localization uncertainty. The improvement comes from two parts. The first reason is the better estimation of distances when utilizing uncertainty distance function. The second reason is the selection of actual data points as centers instead of deterministic objects computed by PDF.
- A vision-aided localization scheme is proposed to replace GPS for outdoor localization. It improves the localization accuracy to the meter level. Pixel characteristic-based features are extracted from satellite images to improve the accuracy of localization. Therefore, three new features are added to the original features provided by the dataset. Input features are put into a Feed-forward Neural Network (FFNN) to predict the location of vehicles. The increased number of input leads to lower localization errors.

1.3 Organization of the Thesis

The rest of this thesis is organized as follows.

Chapter 2 introduces the background of this thesis, including 5G network, RRM and related machine learning techniques. Chapter 3 presents related works in the areas of machine learning-enabled RRM and localization techniques. Chapter 4 defines the proposed schemes in this thesis, including vision-aided localization method, UK-means-based clustering, UK-medoids-based clustering, and DRL based radio resource allocation method, respectively. Chapter 5 illustrates the simulation settings and simulation results. Chapter 6 concludes the thesis and discusses possible future work.

Chapter 2

Background

This chapter contains a concise summary of the subjects that are related to the proposed schemes. Section 2.1 introduces the background of the wireless communication systems and 5G new radio. Section 2.2 gives a summary of widely applied localization techniques. Section 2.4 describes the RRM in 5G wireless networks. Section 2.5 summarizes several machine learning techniques that are utilized in this thesis.

2.1 5G New Radio

Wireless communication systems utilize radio wave as the medium of information to implement transmissions. Compared to wired infrastructure, wireless communication systems have a lower cost and higher flexibility. However, it also encounters challenges such as path loss and inter-symbol interference. Wireless networks have been advancing over the years with the aim of higher speed and more types of services [16]. The first generation has a speed of 2.4 Kbps whereas a speed higher than 1 Gbps is being pursued for the 5G.

To meet the expectations in the near future, 5G technology is developed to address challenges including higher capacity and data rate, ultra reliability, lower end-to-end latency, enormous device connectivity [17], and so on. Then MmWave is utilized for the signal transmission as one of the promising candidates. The abundant bandwidth in the high-frequency range provides users with a higher transfer rate. Ultra reliability and lower end-to-end latency enable the application of emerging technologies such as augmented and virtual reality, unmanned vehicles, and telemedicine. 5G aims to realize the interconnection of all things at any time, anywhere. For example, Internet of Vehicles (IoV) allows

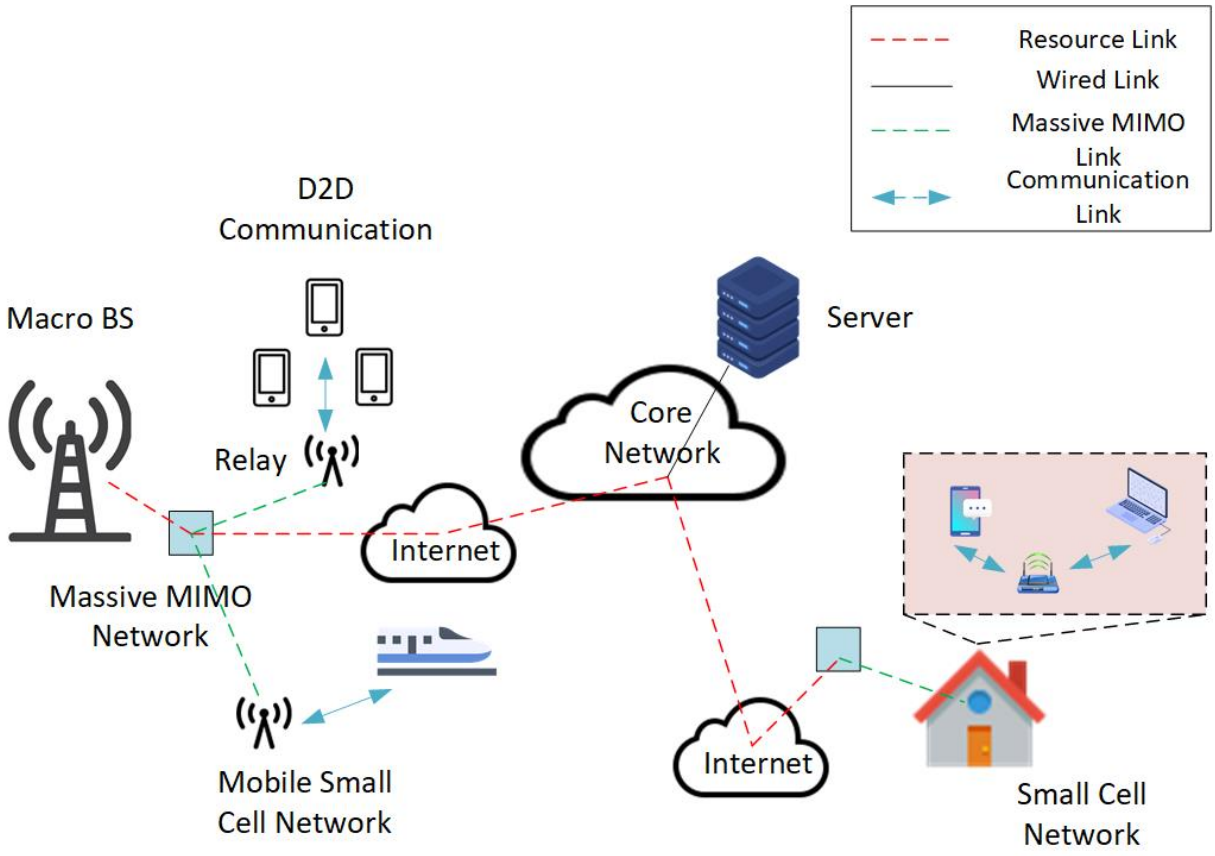


Figure 2.1: The architecture of 5G cellular network

vehicles to exchange information with each other via the Internet, and Internet of Things (IoT) realizes the ubiquitous connection between things and things, things and people.

Considering the shortened transmission distance and smaller coverage area brought by the utilization of a high-frequency band, more base stations are required to cover the same area compared to the Fourth-generation (4G). The 5G cellular architecture become heterogeneous and includes macro cells and micro cells [17]. Macro base stations are larger in size. They apply massive Multi-input and Multi-output (MIMO) [18] technique and are able to serve a large number of users while covering a wide area. Micro base stations are usually installed in dense areas to support UEs in a small range.

Fig 2.1 shows the architecture of a general 5G network. As it is presented, a small cell network is usually a confined space, such as the network in a house or a moving train.

However, the massive MIMO network shown in the picture is composed of a mobile small cell network and a Device-to-Device (D2D) communication network. They are connected via massive MIMO links. All the networks shown in Fig 2.1 connect with each other via the Internet, which is a main characteristic of 5G network. UEs in a 5G network require various QoS standards. For instance, UEs in the house need enhanced broadband while UEs in the moving train need ultra reliable and low latency communication. The core network of 5G network utilizes slicing to enable segmentation of the core network and Network Function Virtualization (NFV). NFV enables network services such as routers, firewalls, and load balancers that traditionally run on proprietary hardware to be packaged as virtual machines on commodity hardware. In this way, service providers can run their networks on standard servers.

2.2 Localization

GPS is a commonly used outdoor localization and navigation system. GPS provides worldwide time and location data for both military and civilian users by utilizing different carrier frequencies [19]. It is utilized in an assortment of aims including personal navigation, aircraft navigation, automotive navigation, and maritime navigation. The GPS system consists of at least 24 operational satellites orbiting roughly 19,000 kilometres above the earth to achieve 24-hour-a-day coverage [20]. GPS positioning is based on the basic triangulation principle of satellites. The GPS receiving device measures the distance by measuring the transmission time of the radio signal and uses the distance to determine the position of the satellite in space. The localization accuracy of GPS is affected by various elements including satellite and receiver clock error, multipath error, atmospheric effects, and radio frequency interference [21]. GPS has a significant limitation in that it requires a receiver sensitivity of -130 dBm, which makes it unusable in congested metropolitan areas, tunnels, and indoor environments [22].

Some works have been investigated to replace GPS using RF signals and ML. For instance, in an indoor environment, [23] used deep belief network to extract features from Received Signal Strength (RSS) measurements for location prediction. In [24], the localization process was formulated as a Markov Decision Process (MDP), and deep Q-learning was utilized to determine the strategy that gradually localizes the target by mapping observable states to actions. Vision-aided localization solutions are used for outdoor scenarios. For example, [25] proposed an improved localization algorithm based on GPS, Inertial Navigation System (INS), and visual localization. [26] proposed two localization schemes based on deep learning and landmark detection to replace GPS.

2.3 Beam Management

The main purpose of beam management is to obtain and maintain an optimal set of transmission/reception points and UE beams available for downlink and uplink transmission/reception. The procedure of beam management include four important phases: beam sweeping, beam measurement, beam reporting, and beam determination [27]. Beam sweeping is to cover a spatial area with a set of transmit and receive beams at pre-specified time intervals and directions. Beam measurement refers to the evaluation of the received signal quality at the 5G-NodeB (gNB) or UE. Beam reporting refers to the process that the UE uses to send beam quality and beam decision information. Beam determination is to select appropriate beams at gNB or UE according to the measurement results.

Challenges of beam management are listed in [28], including high overhead during beam sweeping, efficient panel switching, high mobility and robustness. In this work, we focus on the high mobility issue. We fix the number of beams instead of forming multiple beams and selecting from formed beams. The angle of beam is also fixed. We only consider the situation where UE is connected to gNB via one beam. Position is the only information utilized to determine beam direction.

2.4 Radio Resource Management

The aim of RRM is to provide service quality assurance for wireless users by flexibly allocating and adjusting available resources [29]. Various constraints, such as system throughput and QoS, are applied while devising RRM scheme under different situations. In this thesis, we consider the spectrum and power allocation, beamforming, and user clustering.

2.4.1 Spectrum and Power Allocation

5G adopts Orthogonal Frequency Division Multiple Access (OFDMA) scheme [30] to serve multiple users in a cell. BS allocates bandwidth resources by transmitting to UEs at different time slots and frequencies. A resource grid is composed of time and frequency resources. As illustrated in Fig 2.2, the grid's time direction is divided into a number of Orthogonal Frequency Division Multiplexing (OFDM) symbols, while the frequency direction is divided into a number of orthogonal subcarriers. Each Resource Block (RB) is made up of 12 contiguous subcarriers. And the interference is minimal among adjacent blocks since they are orthogonally divided [31]. As the minimum duration of a transmission,

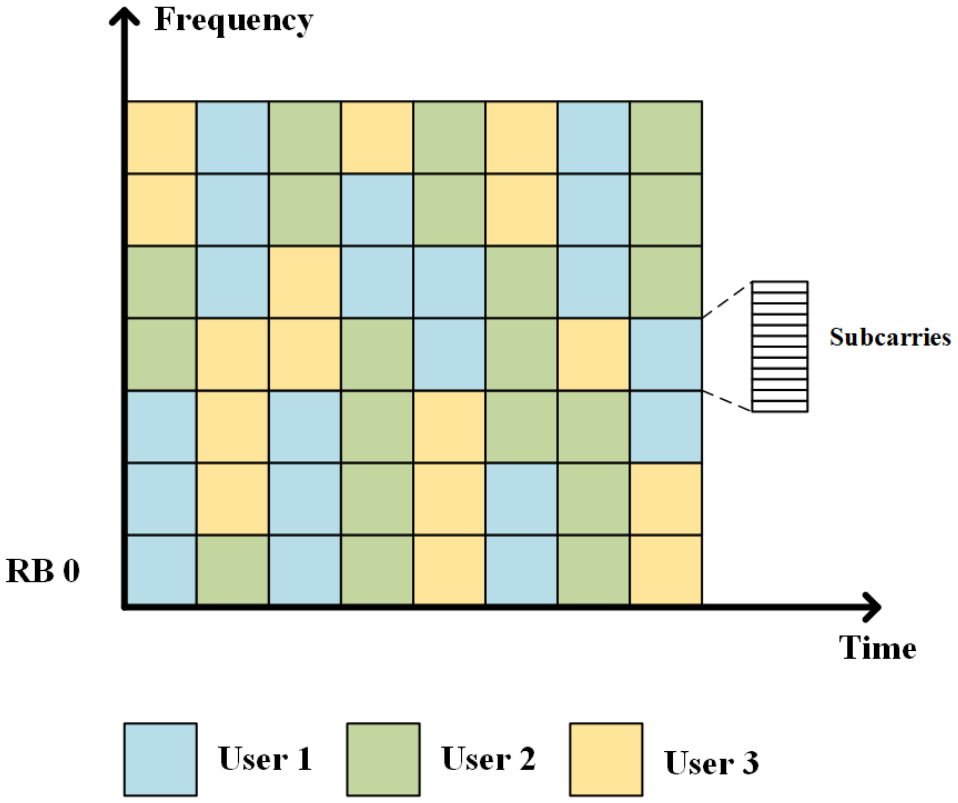


Figure 2.2: Resource grid of 5G networks

a time slot is also named by Transmission Time Interval (TTI). A RB constitutes the minimum unit of allocation to a user [32]. Therefore, the assignment of resource allocation is to allocate RB to served users in each TTI.

Meanwhile, high-density wireless networks are posing considerable issues due to increased interference caused by co-channel cell proximity. For example, inter-cell interference in uplink transmission is caused by users in various cells sharing the same RB. Inter-cell interference in downlink transmission occurs in adjacent cells when one cell utilizes high transmission power. Power control is a key criterion in 5G systems for less consumption of power and interference mitigation [29]. For instance, [33] has reduced power consumption using Q-learning. [34] has improved energy efficiency but ignore the interference management. The save on resources is at the expense of interference. In addition, authors of [35] declare that joint multi-tier and power control will be potential schemes to mitigate the interference caused by multi-layer deployment in 5G.

2.4.2 Beamforming and User Clustering

As aforementioned, MIMO is a technology utilized in 5G where multiple antennas are adopted to increase the performance of wireless networks by reducing multipath fading and channel interference. Beamforming is utilized in a massive MIMO system to maximize each user's received signal power while minimizing the interference signal power from the other users [36]. It sends signals to wireless terminals in a concentrated and directional manner, which can comprehensively improve the signal quality received by wireless terminals and increase throughput. Another advantages include enhanced energy efficiency, improved spectral efficiency, increased system security, and applicability for mmWave bands [36].

Non-orthogonal Multiple Access (NOMA) is another technique utilized to improve the spectral efficiency with power-domain multiplexing [37]. NOMA allows the same resource to be allocated to different users by increasing receiver complexity to make use of the differences between various users. Therefore, multiple users with different power levels can share one beam in mmWave-NOMA systems [38]. The users in the same beam have strong channel correlation whereas users in different beams have little correlation [39]. The problem of deciding on users in the same beam can be transformed to cluster users with similar features.

Various metrics can to measure the similarity between users. For instance, [38], The channel information metrics are utilized as feature sets to decide which cluster the user should belong to. In [40], a cosine similarity is applied as the metric to evaluate similarity between users. In [41] and [32], clustering is performed based on users' locations.

2.5 Machine Learning Techniques

New technologies proposed in 5G aim to solve the existing obstacles in wireless scenarios and satisfy the user's service quality requirements. Traditional approaches, on the other hand, face increasing obstacles as communication circumstances become more complicated. The advantage of machine learning to cope with a large volume of data and complex environment makes it an outstanding tool for solving problems in wireless networks [42]. ML has been applied in various scenarios, including resource management in the Media Access Control (MAC) layer, networking and mobility management in the network layer, and localization in the application layer [43].

Supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning are the four primary types of machine learning techniques [44]. Supervised learning takes labeled input in training to execute a classification or a regression task. Unsupervised learning focuses on finding the similarities between a group of unlabeled data. Semi-supervised learning utilizes both labeled and unlabeled data. Reinforcement learning emphasizes that the agent learns through direct interaction with the environment without the need for replicable supervisory signals or full modeling of the surrounding environment, and thus has a different paradigm compared to other computational methods.

In this thesis, we will introduce several machine learning technologies that are applied in our work. The methods include feed-forward neural network, convolutional neural network, long-short term memory neural network, and deep reinforcement learning.

2.5.1 Clustering

Clustering is one kind of unsupervised technique that converts a finite unlabeled data collection into a finite and discrete set of hidden data structures [45]. Clustering algorithms divide data into groups based on their similarity and dissimilarity. Those measurements include Euclidean distance, Minkowski distance, pearson correlation, cosine similarity, and so on [46]. In this section, we will introduce two well known clustering algorithms.

K-means

K-means is a partitioning relocation clustering method [47]. Since there is a significant number of ways to partition the dataset, it is infeasible to enumerate and compare all possible results. Partitioning relocation clustering utilizes iteration to optimize the structure of clusters.

In K-means, distances are adopted to measure inter- and intracluster relations. Assume we need to form N clusters. To start with, N points are selected randomly as the initial centers of clusters. Then iterative optimization is executed to relocate the points. Two versions of K-means iterative optimization are known [47]. The first version is composed of two steps:

(1) Assign each data point to the nearest cluster center according to distance measurement.

(2) Recompute the cluster centers as the weighted average of each cluster.

Iteration will stop when there is no change of cluster centers.

The second version is based on the impact of the movement of one point from the current cluster to any other cluster. The measurement function could be the sum of the distances between the points in the same cluster and its center point. If the movement increases the sum, the point will stay in the same cluster. If the movement decreases the sum, the point will be assigned to another cluster.

K-means is a popular clustering algorithm and it is easy to implement. However, it has two major drawbacks. The clustering results rely heavily on the initial randomization of cluster centers. Moreover, since cluster center is recomputed as the weighted average, K-means is sensitive to outliers.

DBSCAN

Density-Based Spatial Clustering of Applications with Noise (DBSCAN) is a density-based clustering algorithm that finds all the dense areas of the sample points as clusters [48]. DBSCAN has no requirement for pre-defining the number of clusters. It can deal with clusters with arbitrary shapes.

There are two parameters in this algorithm: the radius of the cluster (*Eps*) and minimum points required inside the cluster (*Minpts*). Several critique ideas in DBSCAN as follows.

- Neighborhood: The neighborhood of a point p concludes all the data points whose distance to p is not larger than the radius of the cluster.
- Core point: The core point refers to a point whose neighborhood has at least *Minpts* data points.

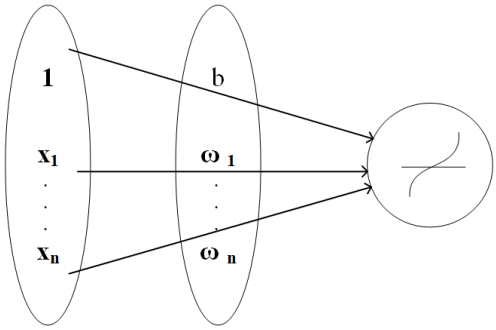


Figure 2.3: Structure of a neuron.

- Boundary point: A point that is not a core point but is within the neighborhood of a core point.
- Noise point: A point that is neither a core point nor a boundary point.
- Directly-reachable: If p is the core point and q is in the neighborhood of p, then p is directly-reachable from q.
- Density-reachable: A point is said to be density-reachable to a core point if it can be connected to it by other directly-reachable core points.

In DBSCAN, a cluster is defined by the set of samples that are connected by the the maximum density-connected sample set derived from the density reachability relationship. In other words, points in the same cluster are mutually reachable to each other. The steps of the algorithm are listed below.

(1) Randomly select a data object point from the dataset.

(2) If the selected data point is a core point, then find all the data points that are density-reachable from this point to form a cluster.

(3) If the selected data point is a boundary point, then select another data point.

(4) Repeat (2) and (3) until all the points are processed.

2.5.2 Neural Network

Neural network is one supervised machine learning technique that is inspired by the human brain. It mimics the way biological neurons transmit signals to each other. A neural network usually takes a large dataset as input, and it will give a predicted result after training and learning from these samples. A neural network concludes layers of nodes, consisting of an input layer, one or more hidden layers, and an output layer. Each node, also known as an artificial neuron, is connected to another node with associated weights and thresholds. The output is determined by whether the sum of the products of the weights and inputs is greater than or equal to a defined threshold [49].

As the basic unit of neural networks, each node can be thought of as its linear regression model, consisting of input data, weights, biases (or thresholds), and outputs. All the inputs are multiplied by their respective weights and sum. After that, the output is passed through an activation function that determines the output. Fig 2.3 shows the structure of a neuron, where $\boldsymbol{x}$ is an input vector, $\boldsymbol{\omega}$ represents the vector of weights, $\boldsymbol{b}$ is the bias and $g(z)$ is the activation function. Therefore, let $z(\boldsymbol{x})$ denotes the pre-activation, the output $o(\boldsymbol{x}; \boldsymbol{b}, \boldsymbol{\omega})$ is as follows

$$o(\boldsymbol{x}; \boldsymbol{b}, \boldsymbol{\omega}) = g(z(\boldsymbol{x})) = g(\boldsymbol{b} + \sum_{i=1}^{n_d} \omega_i \boldsymbol{x_i}) \tag{2.1}$$

where n_d is the total number of data in the dataset.

A widely used activation function named sigmoid activation is as follows:

$$g(z) = \frac{1}{1 + e^{-z}} \tag{2.2}$$

FFNN is the simplest form of neural network. Figure 2.4 shows the structure of a two-layered FFNN. In FFNN, the information is propagated only in one direction [50], from front to back, which represents the input and output, respectively. As shown in the picture, there is no connection within a single layer of the network and usually, neurons are fully connected between adjacent layers. The training process of a FFNN is to find weights that can maximize the accuracy of prediction results. FFNN is trained by back-propagation. The error between predicted result and expected result is propagated from the output layer to the input layer, and the error can be measured by Mean Square Error (MSE) as follows.

$$MSE = \frac{1}{n_d} \sum_{i=1}^{n} (\hat{\boldsymbol{y}} - \boldsymbol{y})^2 \tag{2.3}$$

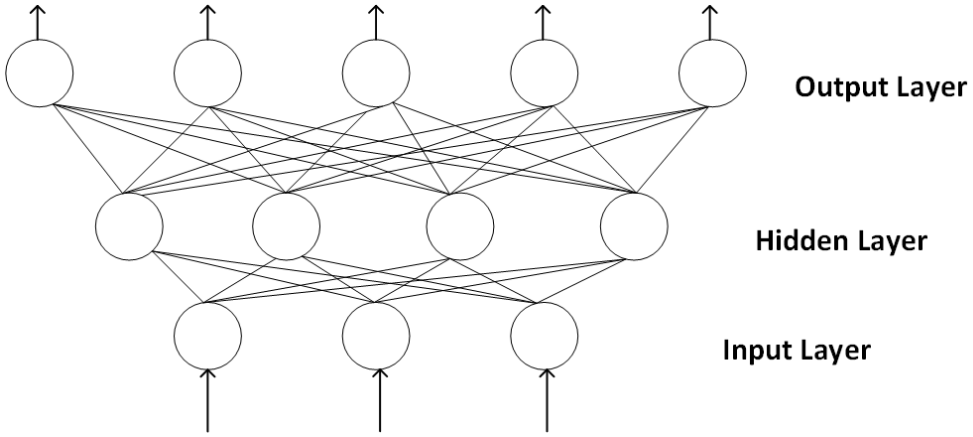


Figure 2.4: Structure of a two-layered feed-forward network

where $\hat{\boldsymbol{y}}$ is the output vector of the neural network, $\boldsymbol{y}$ is the exact data.

2.5.3 Convolutional Neural Network

In the neural networks, there are a large number of parameters, which leads to overfitting easily. Deep neural networks which contain multi-hidden layers are able to extract the abstract information of data through multi-layer high-level features, to obtain better feature robustness.

Convolutional Neural Network (CNN) is one of the most popular deep neural networks. The strength of CNN is to reduce the number of parameters in neural networks [51], which makes CNN suitable for image processing. CNN is composed of convolutional layer, pooling layer, and fully connected layer as shown in Fig 2.5. The number of convolutional layer and pooling layer be arbitrarily increased according to the actual situation. The convolutional layer is responsible for extracting local features in the image. And the pooling layer can reduce the parameter magnitude. Moreover, the fully connected layer is similar to the traditional neural network part as described in Fig 2.4, applied to output the desired results.

In the convolutional layer, a convolution kernel, also known as a filter, is adopted to filter each small area of the image to get the eigenvalues of these small areas. Different filters can bring distinct effects, such as edge detection, gaussian blur, and sharpen. Since each filter only needs to extract the local features and the integration of features is at a higher level, the number of connections is reduced. Moreover, operating convolution with

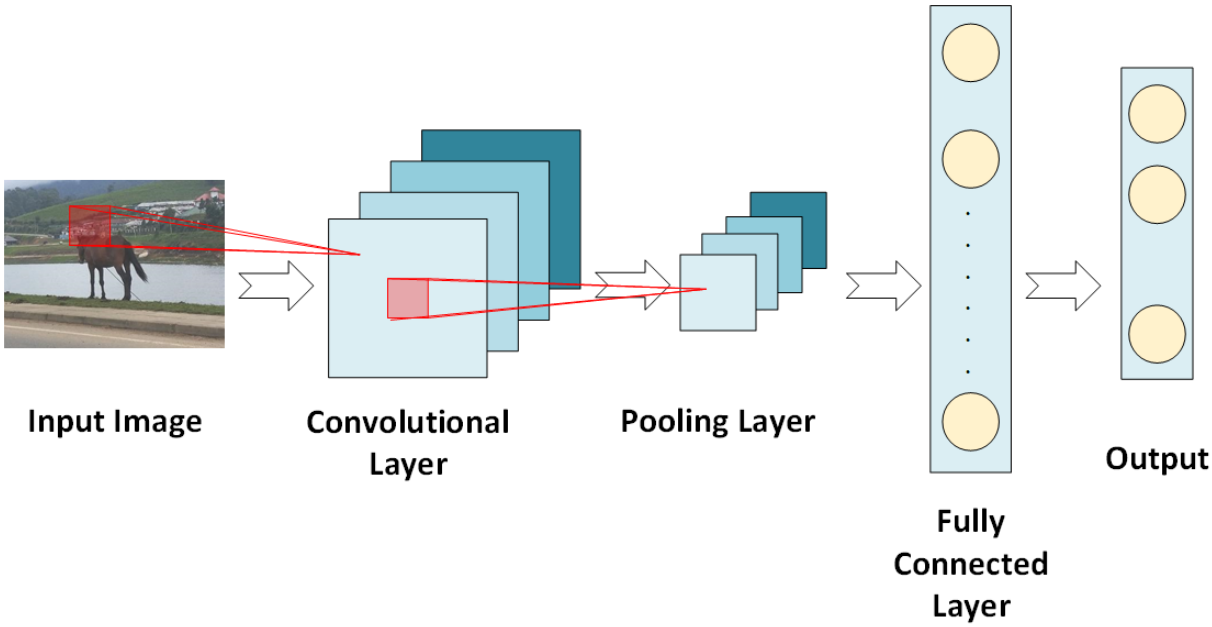


Figure 2.5: Structure of a simple CNN

the same convolution kernel enables weight sharing, which reduces the training parameters of the network, making the neural network structure simpler and more adaptable. The output feature matrixes are then sent to the pooling layer. The main idea of pooling is down-sampling in order to reduce the complexity for further layers [51]. It not only greatly reduces the amount of computation, but also effectively avoids overfitting. Finally, the output is flattened into a single one-dimensional vector, which is fed into a FFNN.

2.5.4 Long-short Term Memory Neural Network

Long-short Term Memory (LSTM) Neural Network is one type of Recurrent Neural Networks (RNNs). RNNs build a memory of time-series events through self-feedback connections between higher and lower layers [52]. Unfortunately, as the gap between the relevant information and the current prediction information becomes larger, RNNs loses its ability to connect such distant information. LSTM is developed to minimize long-term reliance by regulating the flow of information through internal processes known as gates. Gate is composed of a sigmoid neural network layer and a dot-multiply operation and allows information to pass selectively. The sigmoid function outputs a number between 0 and

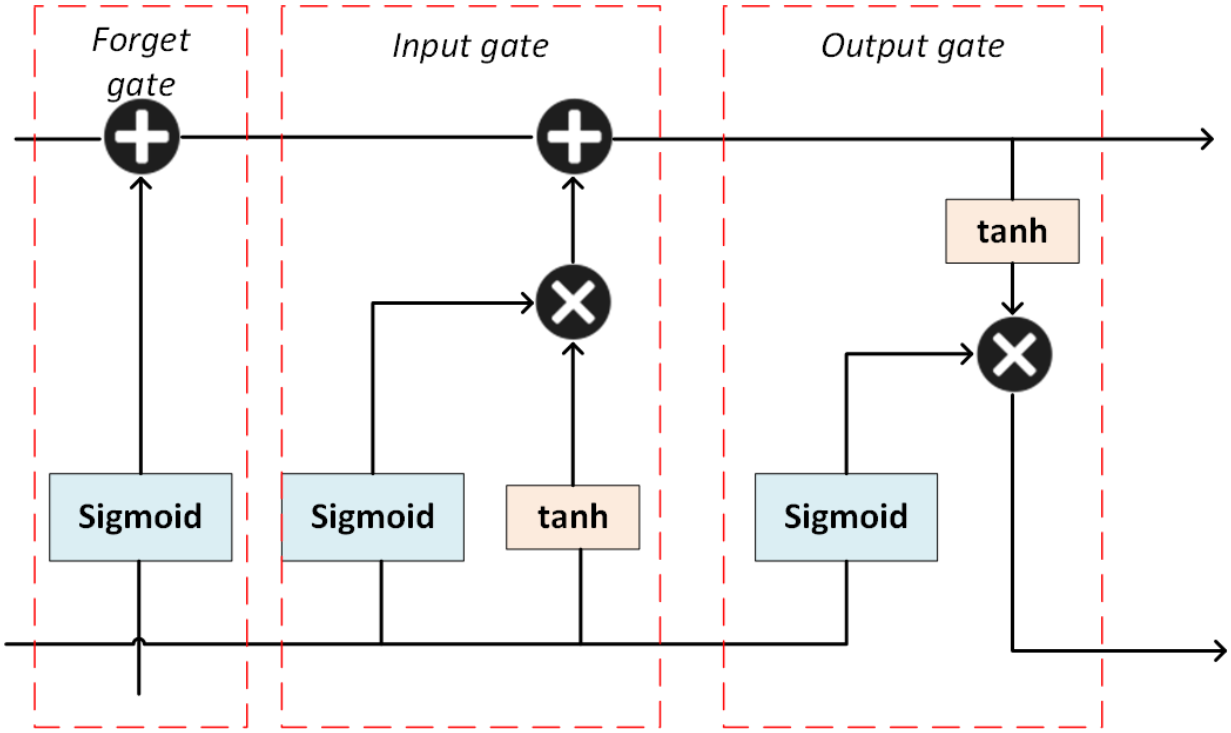


Figure 2.6: The structure of LSTM

1, indicating how much of each component can pass through. 0 means the information can not pass through, and 1 means all information can pass through. Input gate, output gate and forget gate are three main elements in LSTM neural network. The architecture of LSTM is shown in Fig 2.6. The first step in an LSTM is to decide which information to discard and the decision is executed by forget gate. The next step is to decide the new information that needs to be stored and it consists of two phases. A sigmoid layer called the input threshold layer decides which values to update. A tanh layer creates a new candidate vector that will be added to this state. Then the old cell stage is updated to a new cell stage by multiplying the old cell state by the output of the forget gate and adding the product to the output of the input gate. Finally, the output gate decides what the new hidden state should be.

LSTM is suitable for sequential learning applications for its ability to categorize, analyze, and forecast time series with unknown time gaps between critical occurrences [53].

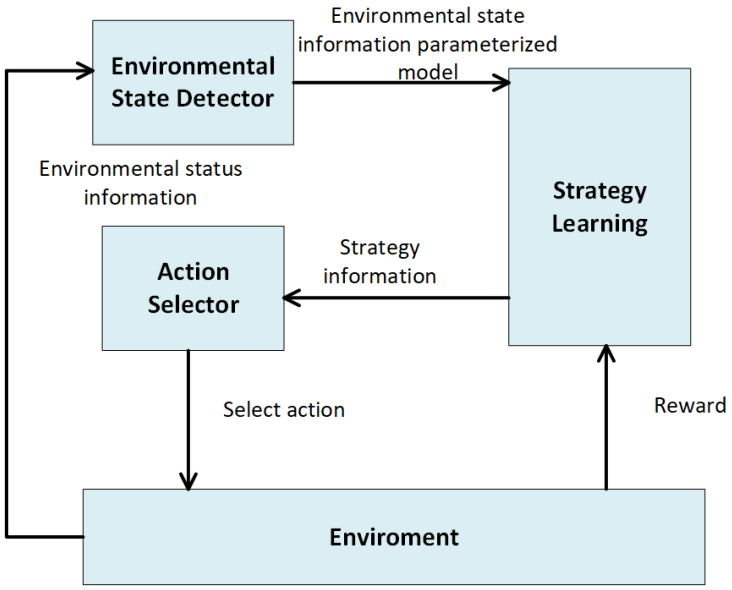


Figure 2.7: Standard reinforcement learning model

2.5.5 Reinforcement Learning

RL is the area of machine learning that deals with sequential decision-making [54]. An agent needs to interact with the environment, collect information over time, and learn the best solution by optimizing the reward [55].

Fig 2.7 presents a standard RL model. Four features are included in a RL mode: state, action, reward and policy. At each time slot t, the current state is denoted by s_t and the agent selects an action a_t following a policy $\pi(a_t|s_t)$. As a result, the agent receives a reward r_{t+1}, the environment transits to next state s_{t+1}. The RL problem can be formulated as a MDP, which means the reward is only determined by current state and current action. RL algorithms are used for robotics, gaming, and navigation in diverse areas [56].

The goal of the agent is to learn an optimal policy Π that can maximize the total expected discounted reward over time. The expected discounted reward and the optimal policy can be defined as follows:

$$V^{\Pi}(s_t) = \sum_{i=0}^{+\infty} \gamma^i r_{t+i}, \tag{2.4}$$

$$\Pi^*(s_t) = \underset{a_t}{\text{argmax}}[r_{t+1} + \gamma V^{\Pi}(s_{t+1})], \tag{2.5}$$

where γ represents discount factor and $0 \leq \gamma \leq 1$.

However, the agent will not have perfect knowledge of the environment, which means the knowledge of r_{t+1} and s_{t+1} is lacking. Therefore, Q-function is defined as an alternative to learning the optimal function.

$$Q(s_t, a_t) = r_{t+1} + \gamma \max_{a'} Q(s_{t+1}, a') \tag{2.6}$$

$$V^*(s_t) = \max_{a'} Q(s_t, a') \tag{2.7}$$

Since there is a close relationship between the largest total expected discounted reward V^* and Q-value function. The learning of V^* can be transformed into the learning of Q-value function. The general iterative method in Q-learning is as follows:

$$Q(s_t, a_t) = (1 - \kappa)Q(s_t, a_t) + \alpha(r_{t+1} + \gamma \max_{a'} Q(s_{t+1}, a')) \tag{2.8}$$

where κ is a function of state-action pairs. During the training process, $Q(s_t, a_t)$ will converge to the optimal Q-value $Q^*(s_t, a_t)$.

2.5.6 Deep Reinforcement Learning

Traditional RL is limited to situations where the action space and sample space are small and generally discrete. For example, Q-table is utilized in Q-learning to store Q-value for each state-action pair. The update of Q-value in every iteration leads to a long convergence time which restricts its applications in large state-action spaces problems. However, realistic tasks often have large state spaces and continuous action spaces. Therefore, DRL is introduced to execute in high-dimensional space by using a deep neural network to estimate Q-values [54].

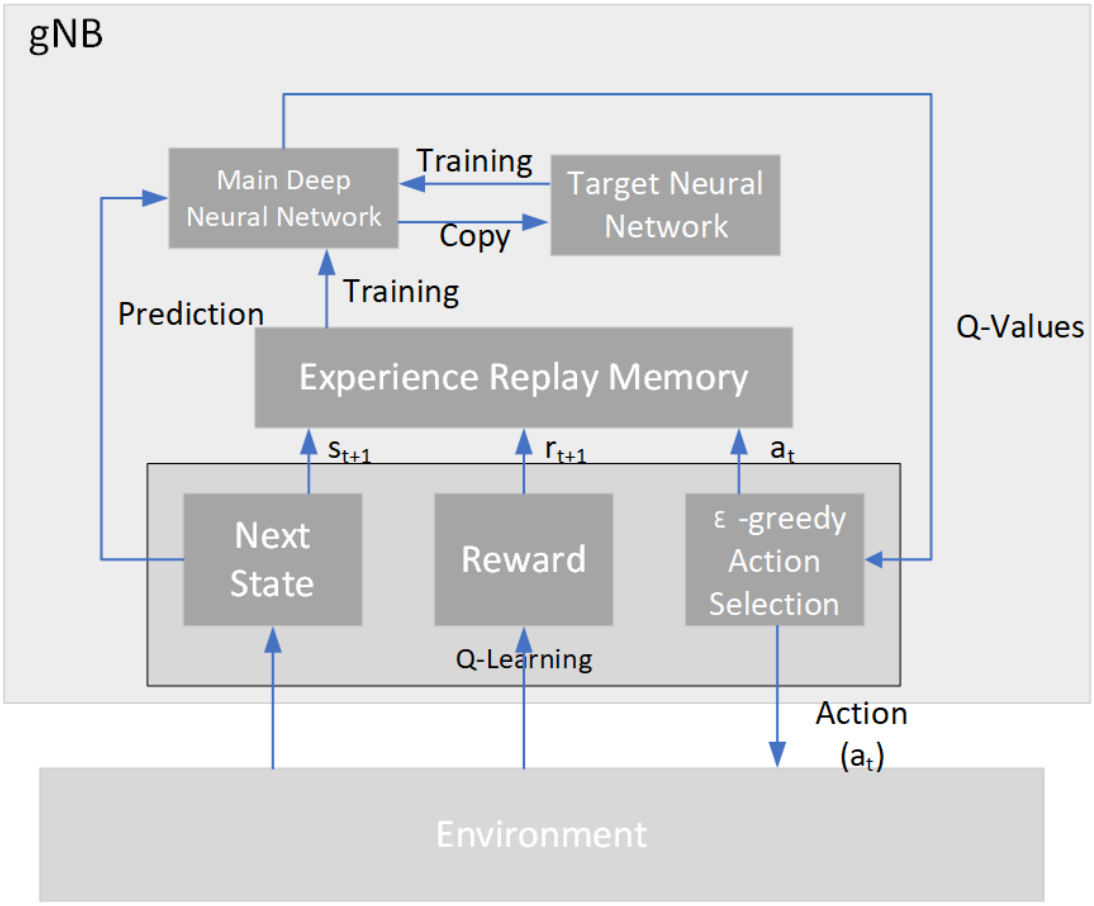


Figure 2.8: Architecture of deep reinforcement learning method

The architecture of a DRL network is shown in Fig 2.8. There are two important elements in DRL: experience replay memory and target neural network [57]. An experience replay memory is applied to collect experience. An experience is composed of four tuples, including s_t, a_t, s_{t+1} and r_{t+1}. The collected experiences are put into the experience replay first, and then a batch is randomly selected from the experience replay for training the network which will predict Q-values. This treatment breaks the correlation between experiences and makes them independent of each other. The existing Q-value is used to calculate the target value of the network, and a slowly updated network named target

neural network is used to provide this Q-value. This improves the stability and convergence of training.

Then, a parameterized value function $Q(s, a; \theta)$ is applied, where θ represents parameters that define the Q-values. Gradient descent algorithm is used to train the main network.

$$L_{DQN} = E(r_t + \gamma \max_a Q(s_{t+1}, a, \theta') - Q(s_t, a_t, \theta)), \tag{2.9}$$

where θ and θ' represent the weight of main and target networks, respectively. $E()$ is the error function. $Q(s^t, a^t, \theta)$ shows main network will predict current Q-values. Then the updated network weight will be copied to the target network after several training.

Chapter 3

Literature Review

In recent decades, applying machine learning to improve the performance of wireless networks has expanded dramatically. To develop a joint beam management and resource allocation scheme, localization, user clustering, and resource allocation are three important parts. User clustering can group UE with similar features to be served within one beam. And RRM is executed in each beam. Furthermore, the accuracy of localization can have a significant influence on clustering performance.

In this section, we will first focus on clustering-based network management works since clustering is one of the main methods we adopt to improve network performance. Then RL based resource allocation works will be illustrated. At last, machine learning and vision-aided localization works are introduced.

3.1 Clustering-based Network Management

Authors of [38] developed a power allocation policy to maximize the sum rate in a mmWave-NOMA system, and a K-means-based clustering algorithm is applied to reduce the computational complexity and group the users with the normalized channel direction. An online K-means-based user clustering model is proposed to handle the placement problem for the newly-arriving users. The problem can be regarded as a classification task based on the formed classification model from existing users' channel information. The simulations show that the proposed scheme can improve the performance of the MmWave network in sum rate compared to traditional user clustering techniques.

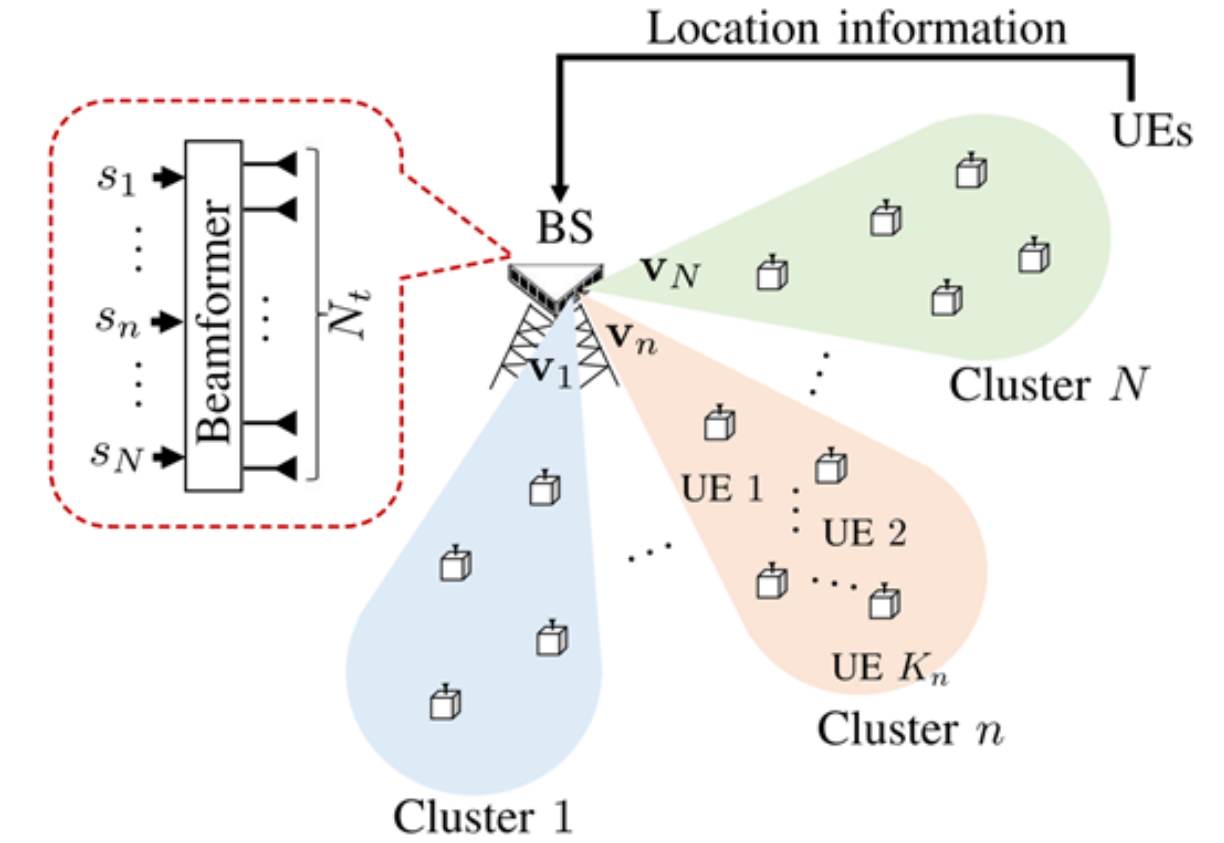


Figure 3.1: System model of [41]

In [40], the authors considered a single-cell mmWave-NOMA system and proposed a hierarchical user clustering technique aiming to maximize the sum rate by automatically identifying the optimal number of clusters. The advantage of the proposed method compared to K-means is that the number of clusters is not pre-required. A cosine similarity is applied as the metric to evaluate channel vector similarity, which means it can specify users that can be clustered together. The simulations show that the proposed algorithm optimises the sum rate of the system while preserving the minimal QoS for all users, without the requirement for the number of clusters as an initial parameter when compared to K-means.

In [41], the authors proposed two schemes of location-aided power allocation optimization to maximize the sum rate of the network. The authors considered a multi user and multiple cluster scenario in a mmWave-NOMA network shown in Fig 3.1. The clustering

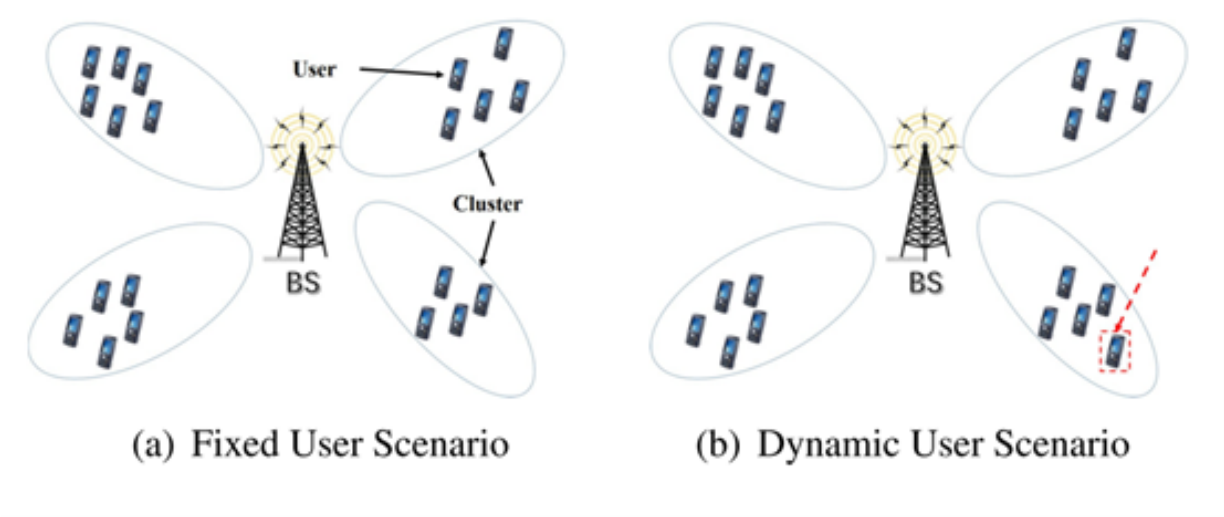


Figure 3.2: System model of [58]

scheme can accelerate the speed of UE cluster assignment. As shown in Fig 3.1, UEs in the same cluster are ordered based on their distances to BS. The whole power for BS is constant and the closest UE is regarded as the primary user. The first location-aided power allocation scheme is based on traditional NOMA. The BS allocates more power to more remote users and less power to closer users. The second location-aided power allocation scheme is based on QoS requirement. The simulation results present that the second scheme outperforms the first scheme significantly in both the total sum rate and the per cluster sum rate.

In [58], the authors firstly developed an Expectation Maximization (EM) based algorithm for fixed users to maximize the sum rate. Then they developed an online EM based clustering algorithm to expand the application scenario to dynamic users. A NOMA-based mobile cellular network is considered for both the fixed user scenario and the dynamic user scenario as shown in Fig 3.2. Based on the EM based scheme for fixed users, the online EM based clustering technique adopts the information of the newly-arriving user to update the adequate statistics. The proposed algorithms both achieve a significant improvement compared with two traditional techniques in sum rate. Moreover, the online EM based scheme can considerably reduce computation complexity.

In [32], the authors considers a MmWave network with two Base Stations (BSs) and multiple dynamic users. User types include URLLC users and eMBB users. They apply DBSCAN for clustering and DRL for resource allocation. DBSCAN can automatically

find the optimal number of clusters instead of pre-defining the number of clusters. The proposed scheme is compared with a K-means-based clustering and priority-based proportional fairness-based resource allocation scheme [38]. The results show that the proposed algorithm outperforms the baseline in latency even though URLLC users have priority accessing radio resources. In addition, the proposed algorithm presents a higher sum rate and lower packet loss rate.

However, for all the works described above, BS is assumed to be able to observe the exact location of UEs, then they implement the clustering and resource allocation accordingly. Unfortunately, this is an unrealistic assumption in practice, and the localization uncertainty is unavoidable due to obstacles, interference, and so on. The performance of wireless networks will degrade due to distorted location information. Different than previous studies, this thesis includes the localization uncertainty of UEs and adopts clustering optimization methods that can mitigate the localization uncertainty.

3.2 Reinforcement Learning-based Network Management

The authors of [59] proposed a Q-learning-based spectrum and power allocation algorithm to increase the capacity of the system for D2D communication in a cellular network. Usually, the interference caused by underlay working mode can be reduced with prior knowledge of the traffic model. The authors employed Q-learning for spectrum and power allocation when prior knowledge of traffic model is not gained. The authors compared the proposed method with two schemes: random allocation and maximum transmitted power allocation. The simulation results prove that the proposed scheme achieves the highest average channel capacity and the interference is successfully reduced.

In [60], the authors improved the throughput by investigating a RL based resources allocation algorithm to control Time Division Duplex (TDD) configuration. Different from previous work, the proposed technique can take the future network status into account and thus avoid traffic congestion. TDD changes the configuration uplink/downlink ratio for achieving communication. Q-learning is applied to choose the optimal resource control policy in each action policy interval. The proposed algorithm is compared with the conventional method and produces a higher throughput and less packet loss rate.

In [61], the authors maximized the overall capacity of a multi-cell scenario by proposing a DRL-based power allocation scheme. A multi-cell OFDM cellular dense network is considered as shown in Fig 3.3. The problem of power allocation is regarded as a multi-objective non-convex optimization problem and conventional methods cost a significant

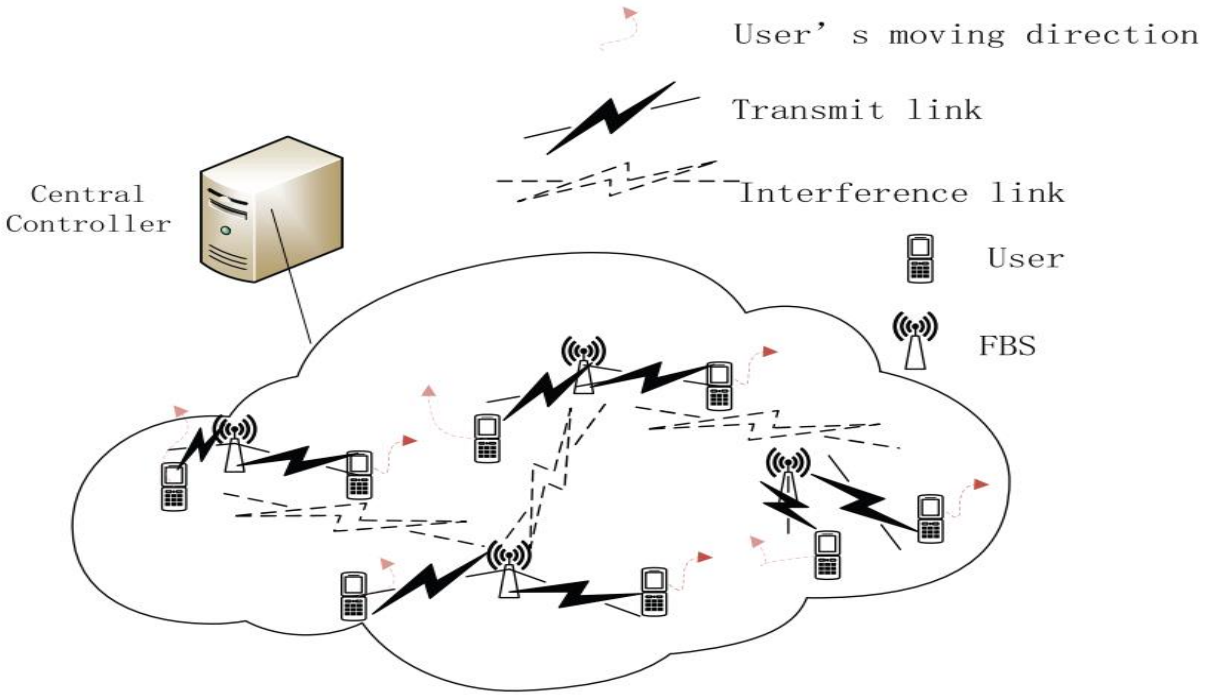


Figure 3.3: System model of [61]

amount of time in training. The authors adopt a deep Q-full-connected-network to reduce the convergence time. The proposed solution is compared to water-filling [61] and Q-learning power allocation and the results demonstrate that the proposed solution can improve the convergence speed considerably.

The authors of [62] applied a neural network to train partial beams for beam alignment. The proposed method reduces training overhead and requires no prior knowledge such as user location. An uplink multi-user mmWave massive MIMO communication system is considered. Fig 3.4 presents the framework of the neural network. The input of the neural network is the stack of received signal vectors at the BS. The output is the predicted beam distribution vector. The proposed method is compared with three conventional beam alignment schemes in spectral efficiency. Results demonstrate that the proposed method is able to improve spectral efficiency.

For the works mentioned above, RL based algorithms are introduced in wireless networks to improve system performance in different areas in RRM stated in 2.4. [59] aims to develop a spectrum and power allocation algorithm by utilizing Q-learning. [60] opti-

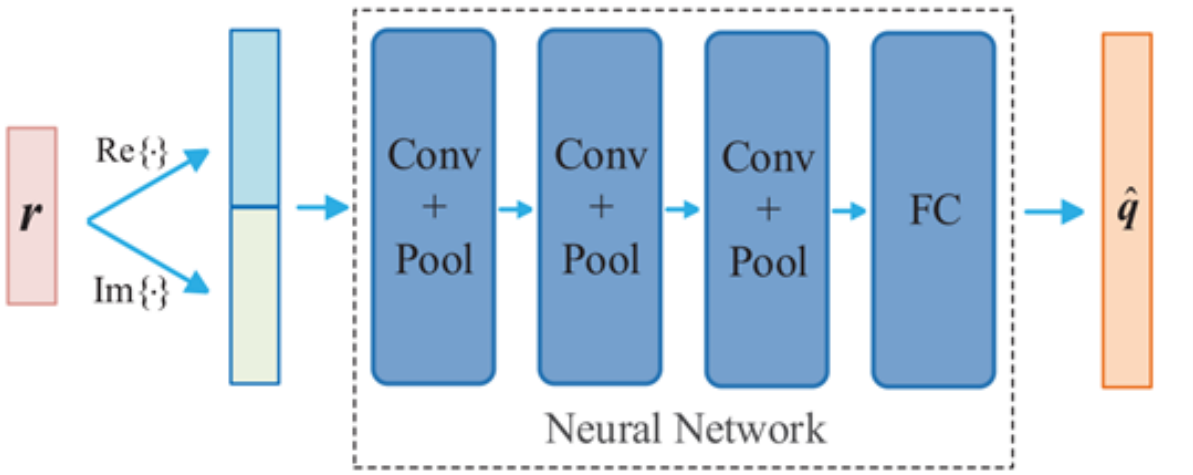


Figure 3.4: Framework of neural network in [62]

mize TDD uplink and downlink multiplexing by Q-learning. [61] improves power allocation scheme by applying DRL. [62] completes the task of beam alignment by adopting deep learning. In addition, RL is utilized for other purposes to improve the sum rate of wireless networks.

[63] considers an aerial base station-assisted terrestrial network. Q-learning is applied to optimize the placement of Unmanned Aerial Vehicles (UAVs) in order to aid ground stations in providing better service to mobile customers. Unlike previous works, this work considers the mobility of users and compensates for the QoS loss. The random walk model is used for all users. The input to Q-learning is the distribution of BSs and the output is the optimal placement of UAVs. Simulation results show that the proposed method achieves a higher QoS with moving users.

In [64], Q-learning is adopted to adjust the downtilt angles of all the ground BSs to guarantee robust wireless connectivity and mobility support for cellular-connected UAVs. The wireless connectivity means good throughput for the ground users. And the efficient mobility support for UAVs is achieved by maximizing the received signal quality. The proposed method is compared to a situation where the downtilt angles are fixed. The results show that the proposed method can reduce the number of handovers and maintain good connectivity to UAVs and ground UEs.

Deep Q-learning is applied in [65] to solve the challenge of packet transmission efficiency in cognitive radio-based IoT. The strategy to transmit packets of different buffers through

multiple channels is optimized to achieve higher throughput. The proposed technique is compared with three algorithms: strategy iteration algorithm, W learning algorithm, and random selection algorithm. Results show that the proposed algorithm outperforms the last two algorithms. Though it has less throughput than strategy iteration algorithm, the algorithm complexity is largely reduced. In addition, the proposed algorithm requires no prior information.

3.3 Localization in Wireless Networks

3.3.1 Machine Learning-based Localization

Traditional localization techniques applies trilateration and triangulation methods based on RF signals to implement simple geometric calculation for localization [66], such as Time of Arrival (ToA) positioning method [67] and Angle of Arrival (AoA) based positioning techniques [68]. However, the application of these methods are limited to LOS scenarios. Recent success of ML in a variety of domains has demonstrated that highly complex problems may be effectively approached with ML. This section will introduce localization works using FFNN, CNN, LSTM and DRL.

In [69], two FFNNs with back propagation are trained for two different features, namely ToA and RSS, as an ensemble model with the unique neural network weights chosen according to the training errors. Let $\hat{\boldsymbol{y}}$, $\hat{\boldsymbol{y}}_1$ and $\hat{\boldsymbol{y}}_2$ representing the final output, output for the first FFNN and output for the second FFNN, respectively.

$$\hat{\boldsymbol{y}} = \omega_1 \hat{\boldsymbol{y}}_1 + \omega_2 \hat{\boldsymbol{y}}_2 \tag{3.1}$$

where ω_1 and ω_2 represent the weights of the two neural networks. They are defined by the rate of individual training error over total training error.

The simulation results show that both ToA method and RSS method lead to larger localization errors than the proposed model. Moreover, the proposed model achieves better generalization capability and stability than a single network.

[70] generates a smooth trajectory consisting of predicted positions using sequential measurements of record samples in a cellular system. The entire region is split into small cells. The cells are regarded as grid structures and then filled with features from the measurement records. As a result, images with height and width similar to the grid structure are created. The images are then sent into CNN to estimate a score for every potential

location based on the extracted spatial features. Then it uses the windowed scores as input to a multi-layer LSTM to learn local spatial features from each measurement record. Finally, the extracted features are put into regression to construct the trajectory.

In [71], the authors focus on cooperative localization for vehicles. DRL is adopted to formulate a scheduling problem to improve the localization accuracy. The authors aim to minimize the number of times performing localization and achieve localization under positioning error bound for all vehicles in the network. They proposed two DRL-based solutions, namely deep Q-network and policy gradient. Policy gradient optimises the policy by enhancing it in the direction of the total reward gradient with regard to the policy parameters [72]. A random policy and a greedy policy are adopted for comparison. Simulation results show that both methods outperform random and greedy policies in terms of required numbers of measurements [71]. Moreover, the two proposed algorithms can reduce the positioning error bound significantly.

3.3.2 Vision Aided-Localization

RGB images can be used directly in localization, especially for pedestrian localization. [73] proposes an algorithm to detect pedestrians in still images utilizing covariance matrices as object descriptors. [74] employs Recurrent Convolutional Neural Network (R-CNN) to identify pedestrians and improves the mean average precision by more than 30%.

Recently, vision information is introduced in wireless systems and the combination provides people with informative data. Vision information and wireless information can compensate for each other's shortcomings [9].

In [75], the authors improve the localization and tracking of individuals with the combination of RGB images and wireless signals emitted by cell phones. This kind of data is referred to as RGB-W data. When the environment becomes crowded, the performance of RGB-based localization decreases due to occlusion. However, with the RSS and mac ID provided by wireless data, the proposed method can keep track of individuals even if they are blocked. The RGB-W based localization is compared with RGB-based localization and W-based localization. The results show that the proposed method outperforms the other two methods in both recall and precision.

In [76], the authors applied the depth images coming from a depth camera to accomplish the localization and navigation for indoor robots. The localization of the robot is done by a vector map with many line segments representing walls. Based on the robot pose, the authors compute a set of line segments that is likely to be observed by the robot. Then the localization can be completed by comparing the line segments in the set with

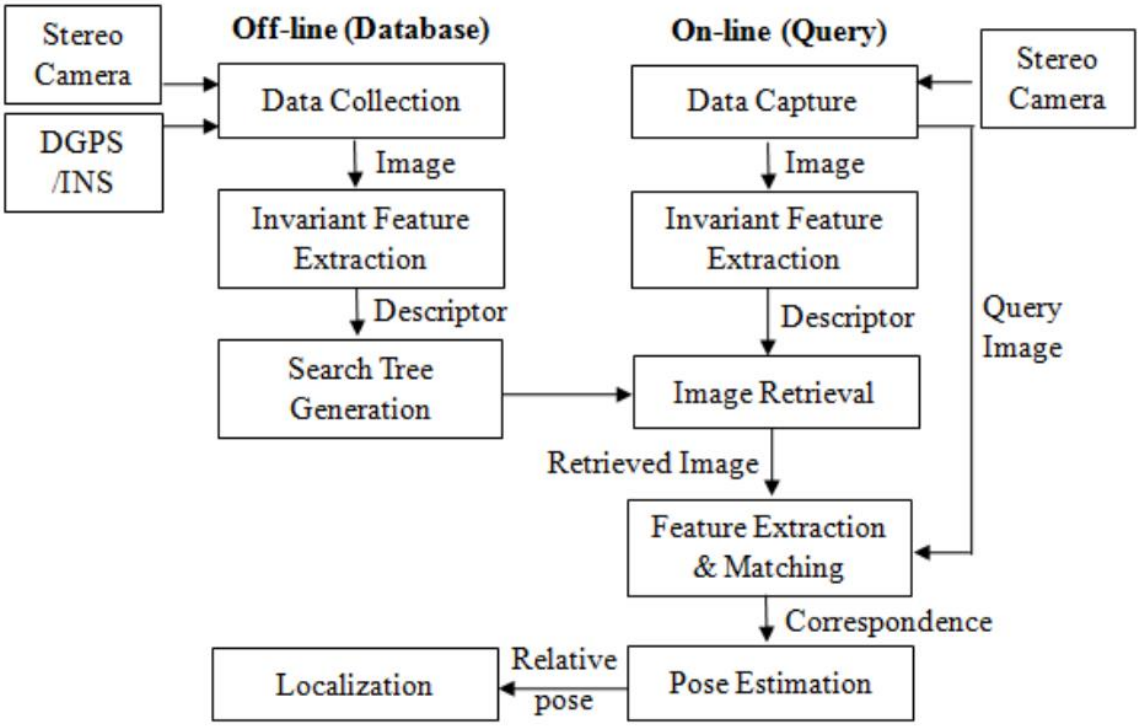


Figure 3.5: Diagram of visual localization algorithm [25]

the vector map. The authors compared the proposed method with the other three localization techniques and checked its long-run performance. The simulation results show that the proposed technique has higher localization accuracy and a lower failure rate. Furthermore, the long run trial test proves the ability of the proposed technique in autonomous navigation.

Authors of [25] proposed an improved localization algorithm for Unmanned Ground Vehicles (UGVs) basing on GPS, INS, and visual localization, where the visual information is used to compensate for the instability of GPS. Fig. 3.5 shows the diagram of visual localization. A vocabulary tree-based structure is generated after feature extraction to store collected data, which is the database. Hierarchical k-means is applied to cluster extracted features to form the tree-based database. When a query image comes, image retrieval is performed to find the matched image. Lastly, position estimation is performed to get the position. The results illustrate that the visual localization estimates the position accurately in when there are lots of trees and buildings wheres GPS aided INS mode

performs well in open areas. Moreover, the switch between the two modes is smooth.

Vision information is also combined with machine learning to replace GPS. In [26], the authors proposed two localization schemes based on deep learning and landmark detection with the assistance of images to replace GPS. The first method applied Faster R-CNN to detect landmarks from an input image. The landmarks features are put into a FFNN for training and predicting. The output of the FFNN is the location and compass orientation. The second method utilized CNN to train all the input data and the output is the same as the first method. Another CNN-based algorithm named AlexNet is used for comparison. Simulation results show that Faster R-CNN outperforms the other two CNN-based techniques in localization accuracy. However, the cost is the decrease in compass accuracy.

3.4 Overview

The aim of the thesis is to develop schemes for joint beam management and resource allocation. Related works are included in this chapter. The way to adopt beam management in this thesis is by clustering. As a result, clustering-based network management related works are firstly introduced in 3.1. Apart from clustering, some works focus on other areas to improve the performance of wireless networks, such as radio resource allocation and power control. Since DRL is applied in this thesis for resource allocation. We include related works that focus on RL model for network management in 3.2. At last, a vision-aided outdoor localization is combined with network management in this thesis. As a result, we introduce some works based on ML and visual information in 3.3.2.

All the works stated in 3.1 and 3.2 assumes accurate location is gained by the BS, which is an ideal situation. This work considers network management based on distorted location and proposes schemes which can mitigate the network performance limit brought by uncertain data. Moreover, this work has combined outdoor localization with network management when GPS performance is limited in congested metropolitan areas.

Chapter 4

Methodology

In this thesis, we have developed two different clustering-based beam management and deep reinforcement learning-based radio resource allocation schemes which can handle both localization uncertainty and various QoS requirements. The position information for clustering comes from real datasets. Before clustering, we adopt data from two datasets for localization to offer both exact location and distorted location for our simulations which demonstrate the effect of the proposed schemes. The first dataset named Raymobtime s009 is found at [77]. LIDAR images and raytracing data are included. The second dataset [78] contains mobile communication system measurements and satellite images from a driving test. In this thesis, we have improved the localization accuracy of [79] by extracting more features from satellite images.

The overall system model of the work is illustrated in Fig 4.1. First, the original dataset is given. In the first scheme, the features provided by the original dataset are used as input directly to a CNN that predicts the user locations. In the second scheme, we extract the image features of each episode in the dataset. Then, these features are used as input for

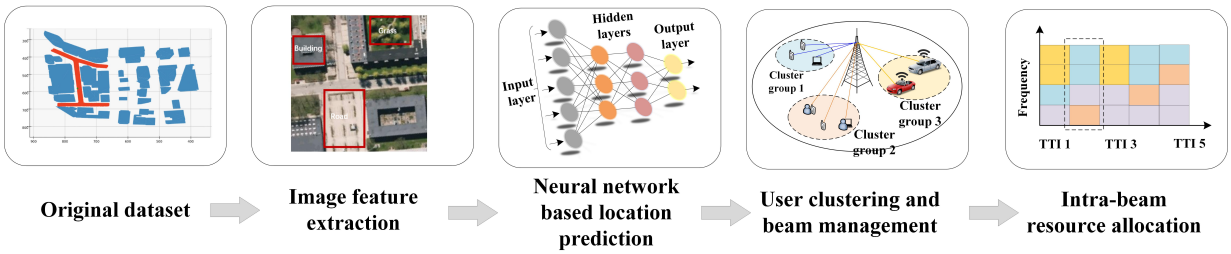


Figure 4.1: Overall System Architectures

a FFNN, which predicts the user locations. Next, considering the localization uncertainty, we propose a clustering method to form cluster groups for beam management. Power is allocated equally to formed clusters. Finally, we apply DRL for the intra-beam resource allocation of each beam.

In this chapter, two datasets for localization are firstly introduced, with a summary of the two works [80] and [79] that we use for localization. Then we will describe the proposed techniques including the system model, vision-aided localization, clustering methods, and resource allocation algorithm.

4.1 Datasets for Localization

4.1.1 Raymobtime Dataset

Raymobtime [77] is collected as a realistic dataset for simulating wireless communications. The methodology utilized ray-tracing, a credible simulation strategy for 5G transmission channels, to simulate MIMO channels. Then a traffic simulator named Simulation of Urban Mobility depicted the mobility of vehicles and pedestrians. At last, the ray-tracing simulator and traffic simulator are combined to produce the final dataset. In detail, Raymobtime generated the data that maps onto Rosslyn, an urban neighborhood in Arlington, VA. Open Street Map is utilized for the collection of outdoor scenarios and Cadmapper is adopted to map the information into 3-dimensional. The scenario should be sufficiently comprehensive to increase ray-tracing accuracy. The 3D scenario is imported into Wireless InSite to simulate the propagation of the signal. The whole studied area is around 300m by 900m and the distribution of Rosslyn map is shown in Fig 4.2. The blue places represent buildings while the red place is the location of vehicles. Moreover, the green point stands for the BS.

The dataset is composed of episodes and scenes as shown in Fig 4.3. The episode contains several scenes and it includes the information for all its scenes. For Raymobtime s009, there are 2000 episodes in total with a time interval of 30s. An episode contains one scene with 10 receivers in each scene. The receivers are a mix of moving automobiles, trucks, and buses, each with its own set of speed limits.

Table 4.1 presents the information that Raymobtime s009 provides. The ray-tracing data such as received power is used for the prediction of location with the assistance of CNN. The predicted location is regarded as Location with Error whereas the location information on the right side of the table is regarded as Exact Location.

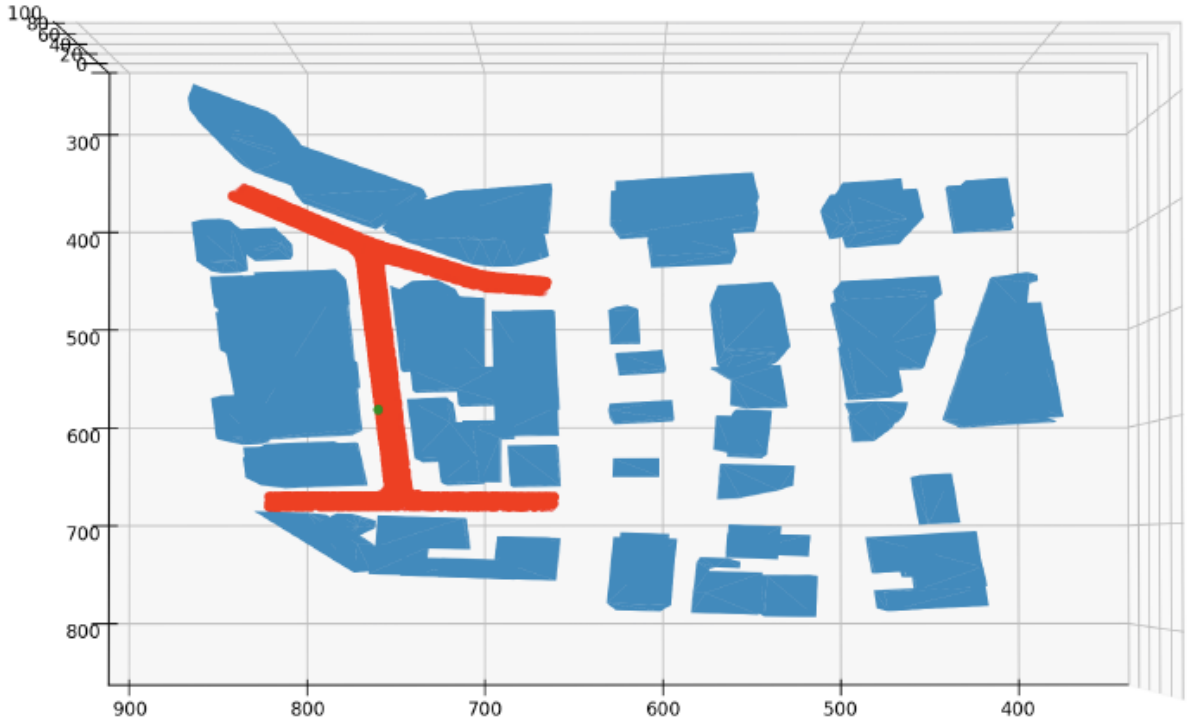


Figure 4.2: Rosslyn map [80]

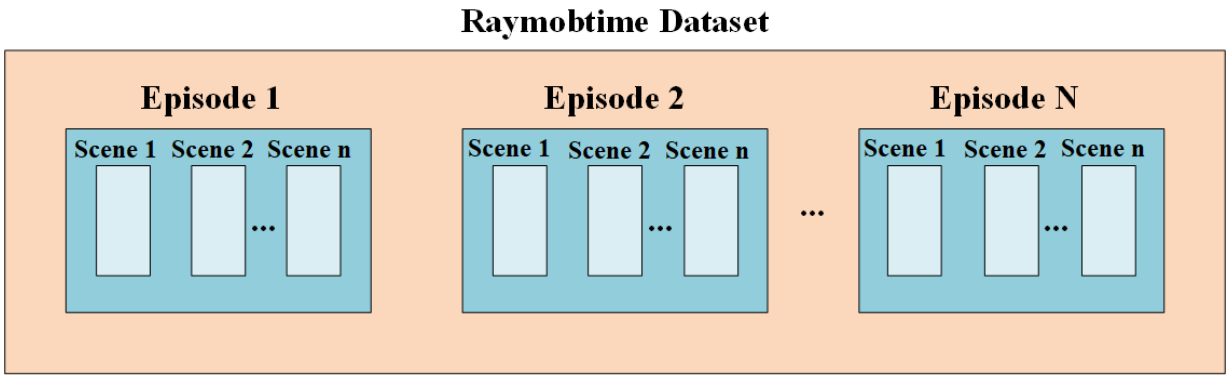


Figure 4.3: Structure of Raymobtime.

Table 4.1: Information Provided by Raymobtime

Ray-tracing Data	**User Information**
Received Power	Valid/Invalid Channel Flag
Time of Arrival	Episode ID
Elevation Angle of Arrival	Scene ID
Azimuth Angle of Arrival	Vehicle Array ID
Elevation Angle of Departure	Vehicle Name
Azimuth Angle of Departure	LOS Flag
Gain Phase	X Co-ordinate
Ray Phase	Y Co-ordinate
LOS Flag	Z Co-ordinate

In [80], the authors plot the value of RSS to find the correlation between location and RSS. Based on RSS, regional classification is adopted where data is divided into three regions. Data with strong RSS is sorted to one region. For the other two regions, the contour of lidar images is used by comparing the similarity between a test image and a set of training images. In particular, Structure Similarity Index Matrices (SSIM) is applied to produce a score indicating the similarity between two images. Each region has its own CNN network. The input to CNN network is a matrix with 7 rows and 25 columns as follows. The columns represent the measurement from each antenna while the rows represent distinct features.

$$\begin{bmatrix} \text{AOA\_el\_Ant0} & \text{AOA\_el\_Ant1} & \ldots\ldots & \text{AOA\_el\_Ant24} \\ \text{AOA\_az\_Ant0} & \text{AOA\_az\_Ant1} & \ldots\ldots & \text{AOA\_az\_Ant24} \\ \text{AOD\_el\_Ant0} & \text{AOD\_el\_Ant1} & \ldots\ldots & \text{AOD\_el\_Ant24} \\ \text{AOD\_az\_Ant0} & \text{AOD\_az\_Ant1} & \ldots\ldots & \text{AOD\_az\_Ant24} \\ \text{RSS\_Ant0} & \text{RSS\_Ant1} & \ldots\ldots & \text{RSS\_Ant24} \\ \text{gainPhase\_Ant0} & \text{gainPhase\_Ant1} & \ldots\ldots & \text{gainPhase\_Ant24} \\ \text{ToA\_Ant0} & \text{ToA\_Ant1} & \ldots\ldots & \text{ToA\_Ant24} \\ \text{RayPhase\_Ant0} & \text{RayPhase\_Ant1} & \ldots\ldots & \text{RayPhase\_Ant24} \end{bmatrix}$$

4.1.2 Mobile Communication System Measurements and Satellite Images Dataset

The second dataset is based on genuine measurements from a communications system located at the University of Denmark [78]. In this dataset, one mobile vehicle drives over a

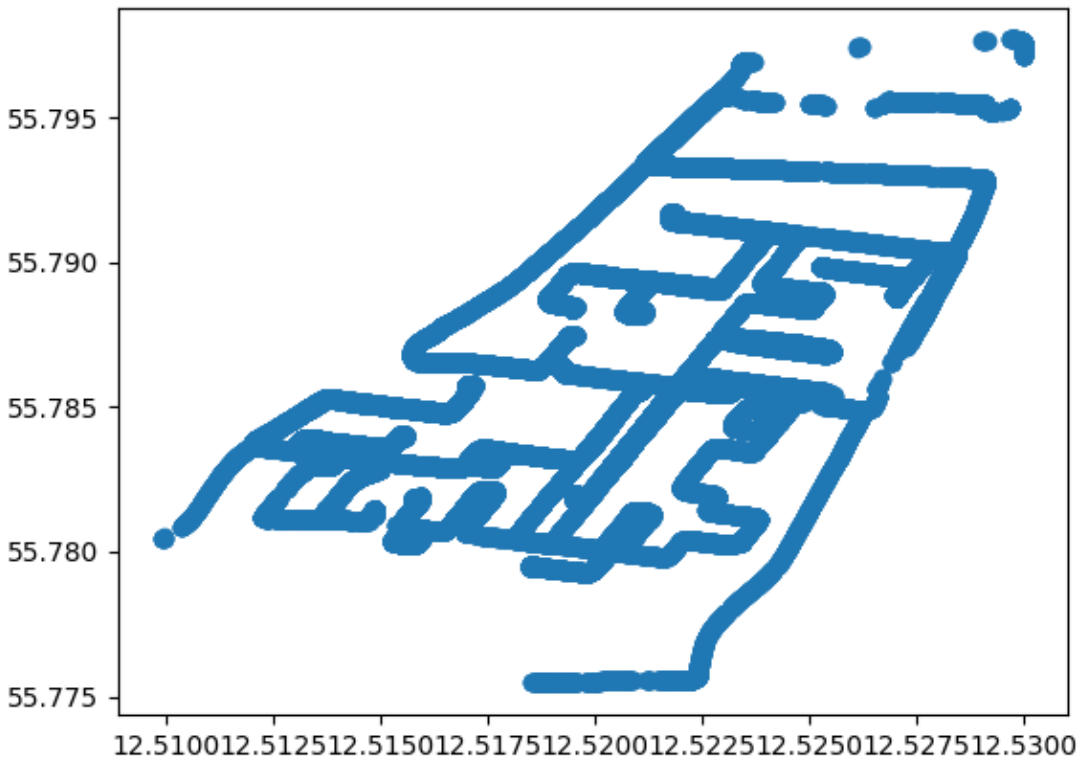


Figure 4.4: Location distribution of mobile communication system measurements and satellite images dataset

region of 2.4 km by 1.25 km and provides communication system measurements, including Signal to Interference and Noise Ratio (SINR), Reference Signal Received Power (RSRP), Reference Signal Received Quality (RSRQ) and Received signal strength indication (RSSI). The information is a part of the input to FFNN aiming to predict the locations. Each sample contains one satellite image that shows the environment around the vehicle. Moreover, this dataset provides the exact location in the form of longitude and latitude. Fig 4.4 shows the distribution of the location. The x-axis is longitude and the y-axis is latitude.

In [79], the authors also employ regional classification to divide the whole studied area into four regions equally. By plotting the RSSI measurements over the localization area,

they find that the top right of the studied area has the strongest RSSI. Similar to [80], the contour of satellite images is utilized for image differencing through SSIM. Each region has its specific FFNN and the input to the FFNN is as follows.

$$[SINR \quad RSRP \quad RSRQ \quad RSSI]$$

In this thesis, we improve the localization of [79] by utilizing the pixel characteristic-based features extracted from satellite images. The proposed method is illustrated in 4.2.2 in detail.

4.2 System Model and Proposed Techniques

This thesis proposes two clustering-based beam management and deep reinforcement learning-based radio resource allocation schemes under localization uncertainty. In the first scheme, a UK-means-based clustering and DRL-based resource allocation algorithm for radio resource allocation and beam management is proposed. We summarize the first scheme as UK-DRL.

In the second scheme, a vision-aided localization technique is firstly developed to reduce localization uncertainty. Then a UK-medoids-based clustering and DRL-based resource allocation algorithm is proposed. We summarize the second scheme as UKM-DRL.

To start with, the system model including the network model and channel model is described. Then the vision-aided localization solution is illustrated. Two adopted clustering methods to handle localization uncertainty are introduced following. At last, the DRL used in the resource allocation phase is depicted.

4.2.1 System Model

Network Model

We consider a gNB serving a group of users with different QoS requirements. Two types of users are included in this a heterogeneous mmWave network: URLLC users and eMBB users. The UE set is denoted by $\boldsymbol{U}$ and each UE is represented by U_i. The gNB will observe the location information of users, and perform clustering to group the users. Based on observed locations, gNB will produce a certain number of UE clusters, and each cluster will be served by one wide beam denoted by b_j. Here the inter-beam interference can

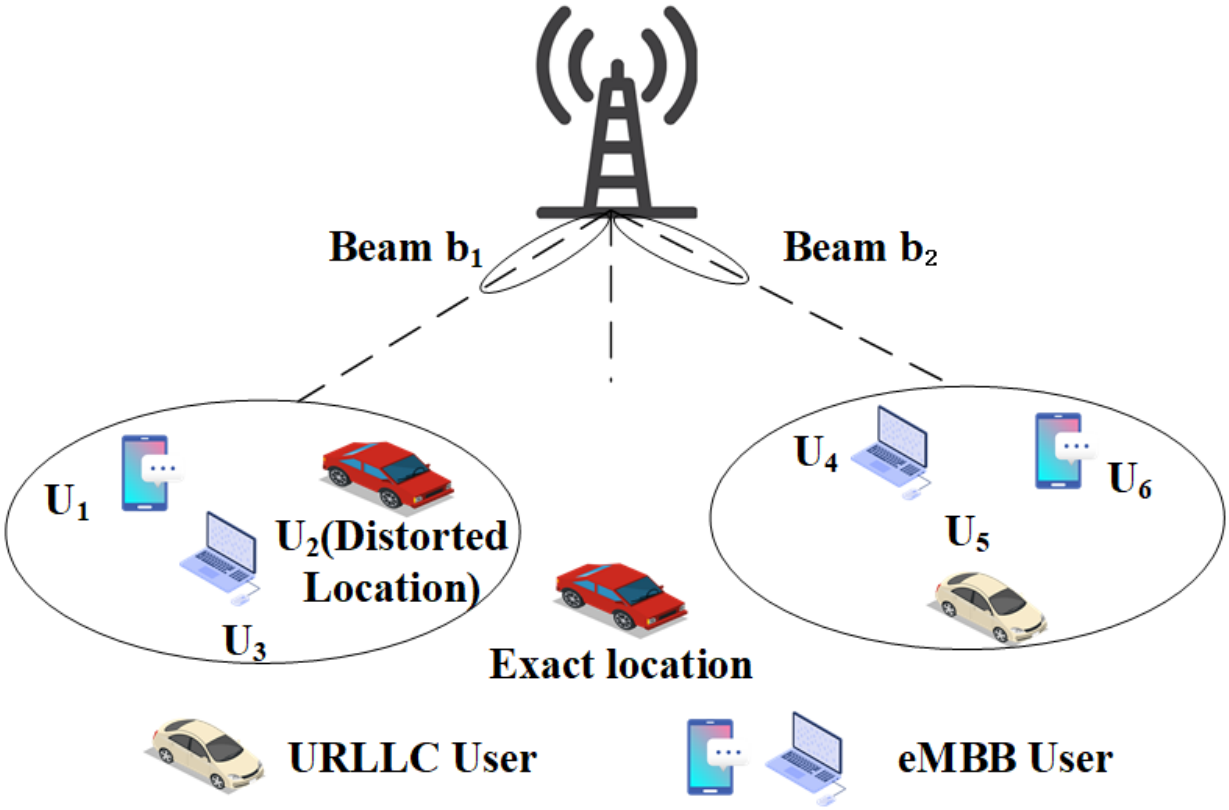


Figure 4.5: Network model

be neglected due to OFDMA technique. In each beam, the radio resource is allocated to distribute the RB groups to the attached users.

As shown in Fig. 4.5, two clusters are formed to cover 6 UEs. We assume the gNB receives a piece of distorted location information of the red car U_2, and two beams are formed accordingly. Therefore, due to the wrong beam direction, U_2 is covered by neither beam b_1 nor b_2 in practice, which may result in a longer delay. Since U_2 is not allocated to any beam, when there is a transmission request for U_2, it will be ignored. The packet can only be transmitted after another clustering is executed and one beam covers U_2 in reality. Moreover, since the former request has no response, when the traffic simulator generates a new request, it will be discarded. This scenario explains how localization error can decrease the network performance of mmWave networks, which is often characterized by a lower throughput and higher latency.

Channel Model

Here we assume the gNB is equipped with N_t antennas. According to [81], the mmWave channel between the gNB and the u^{th} UE can be represented by

$$\boldsymbol{H}_u = \sqrt{\frac{N_t}{M_u}} \sum_{l=1}^{K_u} \alpha_{u,l} \boldsymbol{a}_u(\Theta_{u,l}) \boldsymbol{a}_{gNB}(\varphi_{u,l}) \tag{4.1}$$

where K_u denotes the number of paths between gNB and UE, $\alpha_{u,l}$ is the complex gain of l^{th} path. $\Theta_{u,l}$ denotes the AoA and $\varphi_{u,l}$ denotes the Angle of Departure (AoD). Moreover, $\boldsymbol{a}_u(\Theta_{u,l})$ and $\boldsymbol{a}_{gNB}(\varphi_{u,l})$ are the antenna array response vectors of the gNB and u^{th} UE respectively.

If a Uniform Linear Array (ULA) is considered, $\boldsymbol{a}_{gNB}(\varphi_{u,l})$ can be represented by

$$\boldsymbol{a}_{gNB}(\varphi) = \frac{1}{\sqrt{N_t}}[1, e^{j2\pi \frac{d}{\lambda} \sin(\varphi)}, ..., e^{j(N_t-1)2\pi \frac{d}{\lambda} \sin(\varphi)}]^T \tag{4.2}$$

where d denotes the distance between gNB's antennas and λ denotes the wavelength. The array response at the user $\boldsymbol{a}_u(\theta_{u,l})$ can be written in a similar way.

Since the mm-wave channel can be modeled using a single LOS path model [38]. The channel model vector between the gNB and the u^{th} user can be modeled as:

$$\mathbf{h}_u = \boldsymbol{a}_u(\theta_{u,l}) \frac{\alpha_{u,l}}{\sqrt{L}(1 + d_u^{\xi})} \tag{4.3}$$

where d_u represents the distance between gNB and the user, ξ represents path loss component

4.2.2 Vision-aided Localization

For the mobile communication system measurements and satellite images dataset described in section 4.1.2, pixel characteristic-based features are extracted from satellite images to improve the accuracy of localization.

In digital imaging, the domain of the image is partitioned into X rows and Y columns during the sampling process. A pixel represents the area of the intersection of one row and one column, which is considered the smallest element in one picture, then the resolution of the image is $X \times Y$. In the RGB color model, each pixel is defined by three values representing red, green, and blue ranging from 0 to 255.

(a) (b)

Figure 4.6: Two examples of provided satellite images

The satellite images used in the localization process are RGB images with the same resolution of 256×256, and each image is correlated to a location sample. Fig 4.6 shows two examples of proposed satellite images related to two locations. Each location has its geomorphic features. The grass area in Fig 4.6a is smaller than the grass area in Fig 4.6b.

Therefore, we count the number of pixels that represent one specific object to compute the proportion of this object in the image. It is assumed that a combination set of the proportions of the same objects can establish a relationship between the location. For example, in Fig 4.6a, if the vehicle moves north, the proportion of grass will increase. Since there are some common objects in the satellite images such as grass and buildings. We can use this pixel characteristic-based feature to characterise one image. In this work, we selected three objects. As shown in Fig. 4.7, the three select objects are grass, gray building, and road. Then we count the total number of pixels representing three objects respectively as three new features, which will be used to improve the localization accuracy. Since the color of a pixel is determined by its RGB values, we can classify a pixel into grass, gray building, or road based on its RGB values.

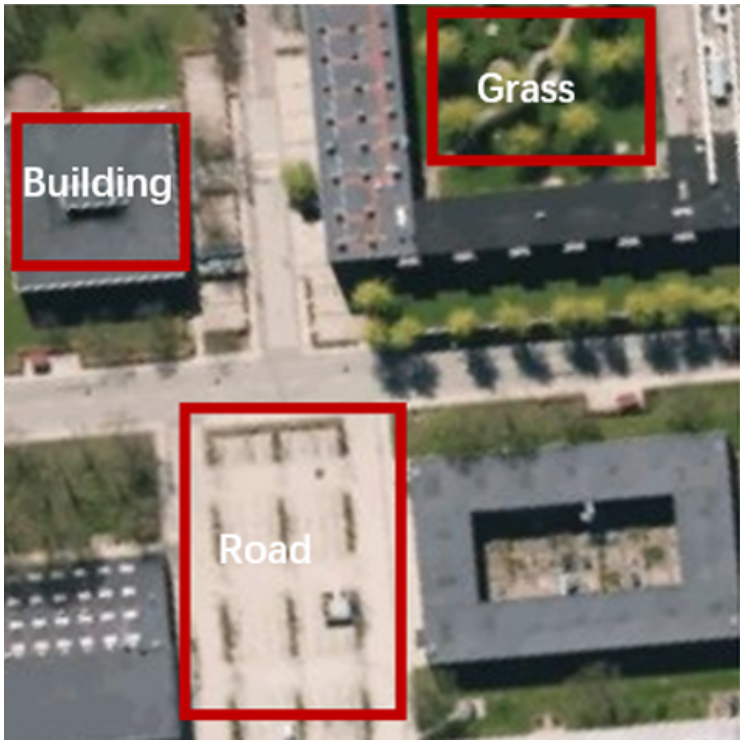


Figure 4.7: Three selected objects for characterising an image

Table 4.2: Mean value of Color Characteristics

	Grass	Building	Road
$\|R-G\|$	12.14	5.41	15.43
$\|R-B\|$	21.77	8.74	27.31
$\|G-B\|$	31.30	3.39	11.87
$\|2G-R-B\|$	42.13	3.01	4.38

To distinguish three objects, we adopted the method from [82]. We selected four color characteristics to observe their mean values, where the color-based characteristics are operations between three basic color values. As shown in Table 4.2, the mean values of $|G-B|$ for the three categories are significantly different. The values for grass, gray building, and road are 31.30, 3.29, and 11.87, respectively. Moreover, the mean values of the other three pixel characteristic-based features are slightly different for the three objects. For $|R-G|$, the mean value of gray building is much smaller than that of the other objects. However,

Table 4.3: Range of Three Categories

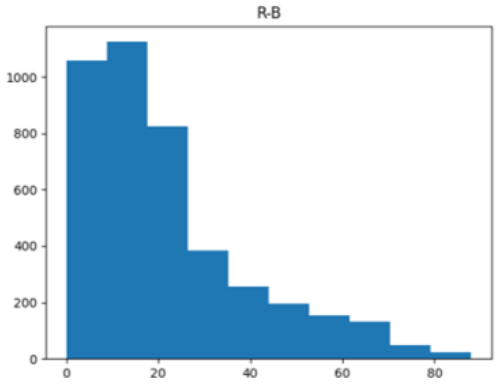
	Grass	Building	Road
$\|G-B\|$	[0,80]	[0,6]	(6,20]
$\|2G-R-B\|$	(12,85]	[1,8]	[0,12]

the difference between grass and road is not obvious. The situation is the same for $|R-B|$. For $|2G-R-B|$, the mean value of grass is much larger than the other two objects. However, there is a little gap between gray building and road. To better decide the value range for three categories, the histograms of these four values for three categories are plotted.

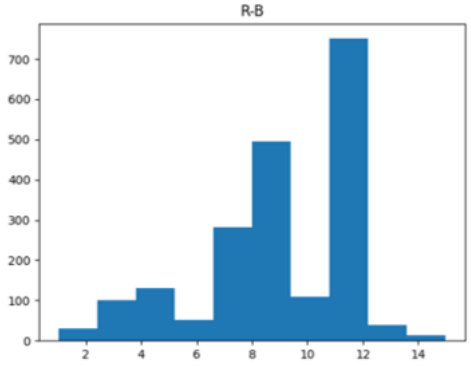
Fig 4.8 - 4.11 shows the histograms of $|R-B|$, $|R-G|$, $|G-B|$ and $|2G-R-B|$ for three objects, respectively. The value range of grass for all four features are wide and include the range of building and road. Moreover, the value range of $|G-B|$ for building and road is completely separated whereas there is part of overlap areas for the other three features. As a result, $|G-B|$ feature is selected to distinguish between building and road. To distinguish grass from building and road, part of the value range should be discarded due to overlap issues. To minimize the reduced number of pixels representing grass, the combination of $|G-B|$ and $|2G-R-B|$ is chosen to distinguish grass.

After observing the histogram of these two values for three categories, we decide the value range for three categories as shown in Table 4.3. The value range is able to include most parts of pixels representing the three selected objects.

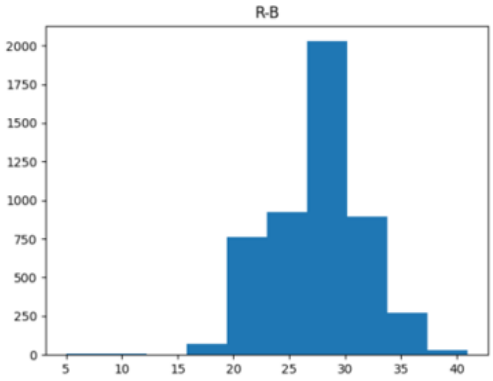
Finally, we deploy 7 features as input of FFNN, including SINR, RSRP, RSRQ, RSSI, the number of pixels representing grass, gray building and road. Then the FFNN will be trained to predict the user locations based on the input features [79].



(a) Histogram of $|R - B|$ for Grass

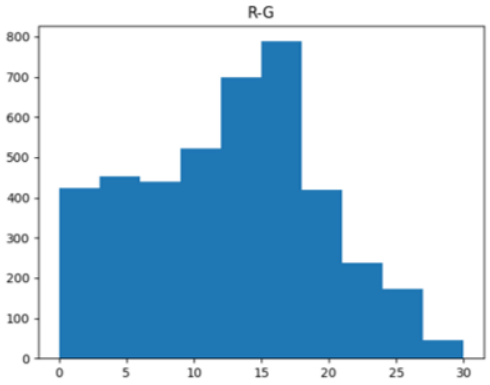


(b) Histogram of $|R - B|$ for Building

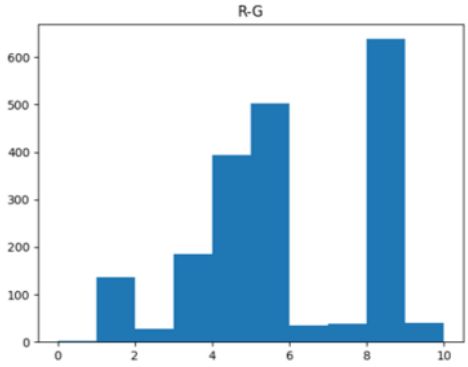


(c) Histogram of $|R - B|$ for Road

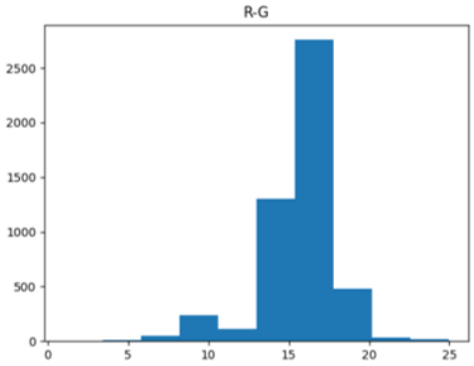
Figure 4.8: Histograms of $|R - B|$ for three objects



(a) Histogram of $|R - G|$ for Grass

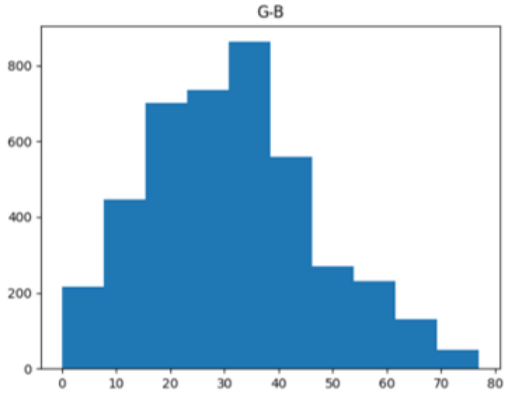


(b) Histogram of $|R - G|$ for Building

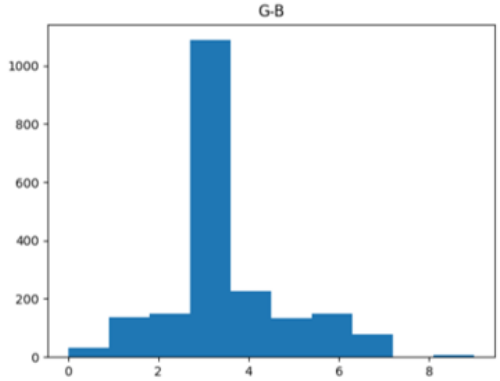


(c) Histogram of $|R - G|$ for Road

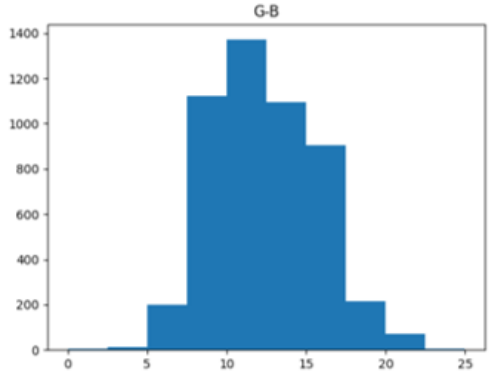
Figure 4.9: Histograms of $|R - G|$ for three objects



(a) Histogram of $|G - B|$ for Grass

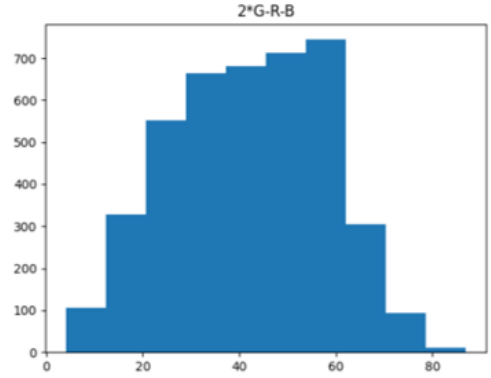


(b) Histogram of $|G - B|$ for Building

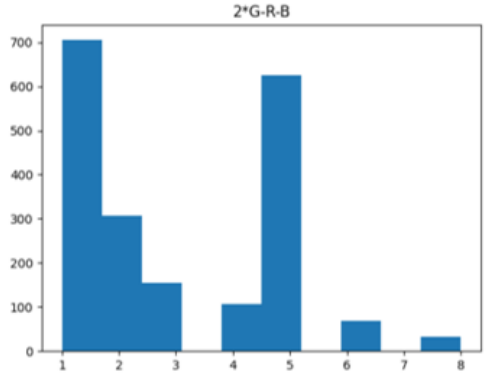


(c) Histogram of $|G - B|$ for Road

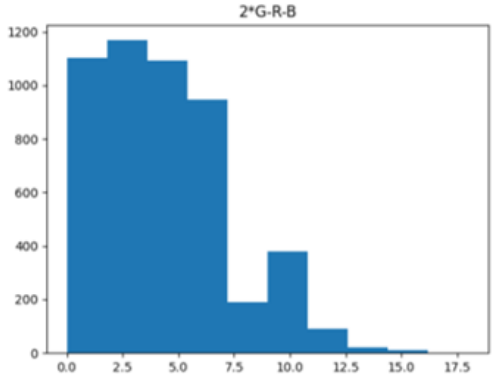
Figure 4.10: Histograms of $|G - B|$ for three objects



(a) Histogram of $|2G - R - B|$ for Grass



(b) Histogram of $|2G - R - B|$ for Building



(c) Histogram of $|2G - R - B|$ for Road

Figure 4.11: Histograms of $|2G - R - B|$ for three objects

.

4.2.3 UK-means-based User Clustering

To handle localization uncertainty in user clustering, we adopt the UK-means for clustering [83]. UK-means is an advanced K-means technique in which the PDF is considered to reduce the influence of localization uncertainty. The clustering results will be transformed into a set of beams with users they serve. Then the DRL is adopted in each beam to allocate physical resource blocks.

The classical K-means algorithm is composed of 4 steps:
Step 1: Assign initial centers of clusters.
Step 2: Compute the distance between the data and centers, and assign them to the cluster with minimum distance.
Step 3: Recalculate cluster centers.
Step 4: Repeat 2-3 until the algorithm converges.

The stages in the UK-means algorithm are the same as in the K-means algorithm. However, UK-means differs in that it uses different methods to determine distance and adopts PDF to update cluster centers. UK-means enables us to get the most similar clustering results with distorted position information compared to the results with exact position information.

Let us assume we need to form N clusters, and the j^{th} cluster is denoted by C_j. The center of the cluster is denoted by $\boldsymbol{c_j}$. For each user position $\boldsymbol{p_i}$, the uncertainty is represented by a known PDF $f(\boldsymbol{p_i})$.

In K-means, the Euclidean distance between data point $\boldsymbol{p_i}$ and cluster center $\boldsymbol{c_j}$ is calculated by:

$$||\boldsymbol{p_i} - \boldsymbol{c_j}|| = \sqrt{\sum_{i=1}^{D} |p_i - c_j|^2}, \tag{4.4}$$

where D represents the dimension of data. In this paper, we consider two-dimensional space.

The center point update formula is as follows:

$$\boldsymbol{c_j} = \frac{1}{C_j} \sum_{\boldsymbol{p_i} \in C_j} \boldsymbol{p_i} \tag{4.5}$$

where $|C_j|$ represents the number of users that are assigned to cluster $|C_j|$.

The two formulas applied in UK-means are as follows [83]:

$$E(||\boldsymbol{p_i} - \boldsymbol{c_j}||^2) = \int ||\boldsymbol{p_i} - \boldsymbol{c_j}||^2 f(\boldsymbol{p_i}) d\boldsymbol{p_i}, \tag{4.6}$$

$$\boldsymbol{c_j} = \frac{1}{|C_j|} \sum_{i \in C_j} \int \boldsymbol{p_i} f(\boldsymbol{p_i}) d\boldsymbol{p_i}. \tag{4.7}$$

In equations (4.6) and (4.7), the distance is an expected distance between the cluster center and uncertain data. It is computed from the double integral over the uncertainty region. The center of one cluster is computed as the mean of the expected values over the PDF of the uncertain data in the same cluster.

In this thesis, the uncertainty region is assumed to be specified as a circle with center $\boldsymbol{x_i}$ and radius R. A node position has an identical chance of being dispersed in any location inside the circle, implying that it is uniformly distributed within the circle. To make the computation easier, the polar coordinate system is applied. The PDF for a uniformly distributed circle is:

$$f(r_d, \theta) = \frac{r_d}{\pi {R_d}^2} \tag{4.8}$$

where $r_d \in [0, R_d]$ and $\theta \in [0, 2\pi]$.

In this way, equation (4.6) can be rewritten as:

$$E(||\boldsymbol{p_i} - \boldsymbol{c_j}||^2) = \int_0^{R_d} \int_0^{2\pi} ||\boldsymbol{p_i} - \boldsymbol{c_j}||^2 f(r_d, \theta) r_d dr_d d\theta \tag{4.9}$$

For a data point $\boldsymbol{x_i}$, its Cartesian coordinate is (x_i, y_i). Assume the coordinate of center $\boldsymbol{c_j}$ is (c_x^j, c_y^j), then:

$$\begin{aligned} ||\boldsymbol{x_i} - \boldsymbol{c_j}||^2 &= (x_i + r_d \cos\theta - c_x^j)^2 + (y_i + r_d \sin\theta - c_x^j)^2 \\ &= 2r_d \sin\theta (x_i - c_x^j) + 2r_d \cos\theta (y_i - c_y^j) \\ &+ (x_i - c_x^j)^2 + (y_i - c_y^j)^2 + {r_d}^2, \end{aligned} \tag{4.10}$$

where $(x_i + r\cos\theta, y_i + r\sin\theta)$ denotes the distorted location. The error $r\cos\theta$ and $r\sin\theta$ is represented in the form of cylindrical coordinate system.

4.2.4 UK-medoids-based User Clustering

Although UK-means is able to give a clustering result more similar to the resulting structure of the real case clustering when distorted positions are applied. The way that UK-means defines cluster centers may result in a loss of accuracy [84]. Developed from K-means, UK-means computes the cluster centers as deterministic objects utilizing the expected value of PDF. Moreover, the distance and cluster centers are updated each iteration, leading to an efficiency issue. UK-medoids, an advanced K-medoids-based clustering technique, is developed to improve the shortcomings of UK-means. It chooses actual data points as cluster centers instead of computing the mean value of clusters. Meanwhile, an uncertain distance function that can better describe the real distance between two uncertain objects is adopted in UK-medoids.

Assuming there are M positions to form N clusters. As stated in 4.2.3, the j^{th} cluster is denoted by C_j and the center of each cluster is denoted by $\boldsymbol{c_j}$. For each user position $\boldsymbol{p}$, the uncertainty is represented by a known PDF $f(\boldsymbol{p_m})$. We assume there are two uncertain objects $\boldsymbol{p_m}$ and $\boldsymbol{p_n}$, and the uncertain region is R_m and R_n, respectively. Then the uncertain distance $\delta(\boldsymbol{p_m}, \boldsymbol{p_n})$ is computed by:

$$\delta(\boldsymbol{p_m}, \boldsymbol{p_n}) = \int_{\boldsymbol{p_m} \in R_m} \int_{\boldsymbol{p_n} \in R_n} dist(\boldsymbol{p_m}, \boldsymbol{p_n}) f(\boldsymbol{p_m}) f(\boldsymbol{p_n}) d\boldsymbol{p_m} d\boldsymbol{p_n}, \tag{4.11}$$

where $dist(\boldsymbol{p_m}, \boldsymbol{p_n})$ is a distance measure between $\boldsymbol{p_m}$ and $\boldsymbol{p_n}$. The uncertain distance is the expected value of the distance measurement between two uncertain objects.

Then, the center of a cluster is computed as the data point in the cluster with the least average dissimilarity to all other points in the cluster. The update method of a cluster center is:

$$\boldsymbol{c_j} = \arg\min_{\boldsymbol{p} \in C_j} \sum\nolimits_{p' \in C_j} \delta(\boldsymbol{p}, \boldsymbol{p'}) \tag{4.12}$$

As stated in 4.2.3, the uncertainty region is specified as a circle whose center is the data point with radius R_d for every data. The PDF is the same as equation (4.8).

In this way, equation (4.11) can be rewritten as:

$$\delta(\boldsymbol{p_m}, \boldsymbol{p_n}) = \int_0^{R_d} \int_0^{2\pi} \int_0^{R_d} \int_0^{2\pi} dist(\boldsymbol{p_m}, \boldsymbol{p_n}) f(r_d, \theta)^2 r_d^2 dr_{d1} d\theta_1 dr_{d2} d\theta_2 \tag{4.13}$$

which is the expansion of the previous equation (4.11) from the rectangular coordinate system to the polar coordinate system. In this thesis, we use Euclidean distance as our distance measurement. The Euclidean distance between (x_m, y_m) and (x_n, y_n) is as follows.

$$dist(\boldsymbol{p_m}, \boldsymbol{p_n}) = \sqrt{(x_m - x_n)^2 + (y_m - y_n)^2} \tag{4.14}$$

Algorithm 4.1 UK-Medoids

Input: A set of uncertain location $\boldsymbol{P} = \{\boldsymbol{p_1}, ..., \boldsymbol{p_M}\}$.
Output: A set of clusters $\boldsymbol{C} = \{\boldsymbol{C_1}, ..., \boldsymbol{C_N}\}$.

```
Computing uncertain distance δ(p_m, p_n), ∀p_m, p_n ∈ P
Selecting initial cluster centers c = {c_1, ..., c_N}.
while c ≠ c' do
    c' ← c
    c ← ∅
    C = {C_1, ..., C_n} ← {∅, ..., ∅}.
    for Each object p ∈ P do
        Allocating each object to the nearest cluster center according to the uncertain distance.
    end for
    for Each cluster C_j ∈ C do
        Recomputing the cluster center c_j.
        c ← c ∪ c_j
    end for
end while
```

The UK-medoids algorithm is summarized by Algorithm 4.1. First, all the uncertain distances between each two-objects pair are computed. The distances only need to be computed once during the clustering algorithm, which will reduce the computation complexity compared to UK-means. Then, the initial cluster centers are selected, and it comes to the main loop. The main loop of UK-medoids contains two phases: the first phase is to assign objects to a cluster based on the minimum uncertain distance, and the second phase is to recompute cluster centers according to equation (4.12). The loop continues until there is no change in the cluster centers.

4.2.5 Deep Reinforcement Learning-based Radio Resource Allocation

After the user clustering, each mmWave beam is assumed to cover a cluster of users within a specific angle. The number of clusters and the angle of the beam is pre-defined. As aforementioned, the clustering technique is UK-means for the first work and UK-medoids for the second work. Then the DRL algorithm is applied in each beam to allocate radio resources to intra-beam UEs, aiming to maintain high QoS requirements for both URLLC and eMBB users. The MDP of DRL is defined as follows:

1) Actions: In j^{th} beam, the action is denoted by allocating k^{th} RB group to u^{th} user, which is defined by

$$a_k^j = U_u^j \tag{4.15}$$

2) States: States are specified as the Channel Quality Indicator (CQI) feedback provided by the user. As a result, the state of k^{th} RB group in the j^{th} beam can be expressed as

$$s_k^j = q_k^j \tag{4.16}$$

where q_k^j is the CQI of k^{th} RB group in the j^{th} beam.

3) Reward: The reward function needs to consider the QoS requirements for both URLLC and eMBB users. URLLC users require high reliability and low latency, while eMBB users desire a high data rate. The reward function of k^{th} RB group in the j^{th} beam can be written as:

$$r_{j,k} = \begin{cases} \mathrm{sigm}(\frac{\mathrm{S_k}}{\mathrm{S^{QoS}}}), & eMBB\ user \\ \mathrm{sigm}(\frac{\mathrm{S_k}}{\mathrm{S^{QoS}}} \frac{\mathrm{T^{QoS}}}{\mathrm{T_u^q}}), & URLLC\ user \end{cases} \tag{4.17}$$

where S_k is the SINR of the link allocation of k^{th} RB group to a user in the j^{th} beam, S^{QoS} is the SINR requirement of eMBB users, T^{QoS} is the latency requirement of URLLC users, and T_u^q is the queuing delay. $sigm$ means the sigmoid function, which will keep the reward between the interval [0,1]:

$$\mathrm{sigm(x)} = \frac{1}{1 + \mathrm{e^{-x}}} \tag{4.18}$$

To identify the type of users, we adopt the Quality Class Indicator (QCI) of users. When QCI is equal to 1, it stands for a URLLC user. When QCI is equal to 2, it refers to a eMBB user.

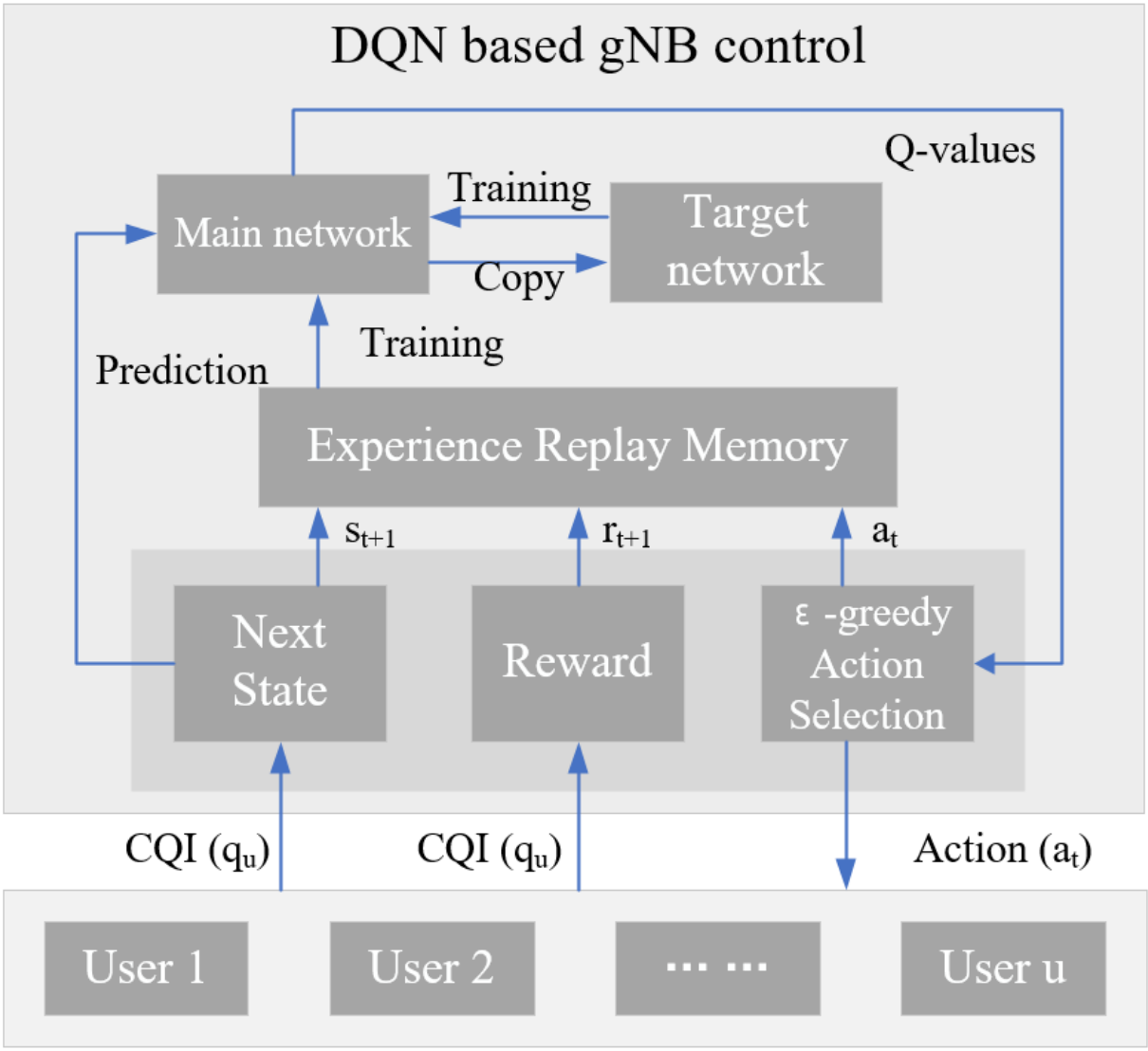


Figure 4.12: Diagram of LSTM-based Deep Q-learning.

In equation (4.17), the reward of eMBB users depends on the SINR ratio, which means a higher throughput will lead to a higher reward. On the other hand, the URLLC reward consists of two terms, in which the $\frac{T^{QoS}}{T_u^q}$ indicates that a lower delay will bring a higher reward, and $\frac{S_i}{S^{QoS}}$ indicates the requirement for higher reliability.

Fig. 4.12 shows the diagram of DRL-based intra-beam radio resource allocation. In the t^{th} time interval, the current state and action are represented by s_t and a_t, respectively. The agent will first select an action based on ϵ-greedy policy, and the actions are implemented in the wireless environment. In ϵ-greedy policy, when the agent makes a decision, there is a small positive probability $\epsilon < 1$ to randomly select an unknown action, leaving $1 - \epsilon$ probability to select the action with the greatest value among the actions that have already been performed.

After an action is taken, the next state s_{t+1} and reward r_{t+1} are calculated with the assist of CQI and SINR from the users. The s_t, a_t, s_{t+1} and r_{t+1} forms an experience tuple, which is saved in the experience pool. The experience will be further used to train the neural network which will predict the Q-values. Lastly, the agent will determine the next action a_{t+1} based on predicted Q-values.

In this thesis, an LSTM network is applied for Q-values prediction. As a special recurrent neural network, LSTM can better capture the long-term data dependencies than traditional neural networks, which makes it a favorable choice to handle the complicated wireless environment. A pack of experience tuples $< s_t, a_t, s_{t+1}, r_{t+1} >$ are sampled from the experience pool after several iterations, which will be used to train the main network by gradient descent algorithm:

$$L_\omega = Er(r_t + \gamma \max_a Q(s_{t+1}, a, \omega') - Q(s_t, a_t, \omega)) \tag{4.19}$$

where ω and ω' represent the weight of main and target LSTM networks, respectively. $Er()$ indicates the error function. $r^t + \gamma \max_a Q(s_{t+1}, a, \omega')$ shows that target network will predict the target Q-values, and $Q(s_t, a_t, \omega)$ shows main network will produce current Q-values. Then the network weight will be copied to the target network after several training iterations. Such a late update of the target network will provide a stable reference for the main network.

Note that since this work focuses on the effect of localization and clustering that can handle localization uncertainty, we apply the same algorithm , DRL, for resource allocation phase. In the following simulations, we will firstly use different clustering methods to prove the proposed schemes can mitigate localization error and improve the sum rate. Then we will show the impact of the value of localization error.

The steps of the proposed UK-DRL algorithm are presented in Algorithm 4.2. The steps of UKM-DRL can be rewritten in the same way.

Algorithm 4.2 UK-DRL

Initialization: Wireless and learning parameters.
for TTI t = 1 to T **do**
    Updating clustering queue and get location of users
    Performing UK-means using equation (4.5) and (4.10). Generating beams based on clustering results.
    for Each beam $b_j \in \boldsymbol{B}$ **do**
        Selecting an action a_t in each beam based on ϵ-greedy algorithm.
        for User $U_u \in b_j$ **do**
            Perform the action assigned by gNB.
            Send uplink report (i.e. SINR) to gNB.
        end for
        Compute the reward r_{t+1} in equation (4.17).
        Store the experience tuple in the experience replay memory.
        Sample randomly from experience replay memory every Γ TTIs and apply equation (4.19).
        Predict the Q-values for next action selection.
    end for
end for

Chapter 5

Results

This chapter will evaluate the performance of two proposed techniques, which are UK-DRL and UKM-DRL. The aim of the simulations is to prove the proposed schemes can mitigate the negative influence brought by localization uncertainty. The evaluation metrics are throughput and latency. Moreover, in the second technique, we will present the advantage of the proposed vision-aided localization solution in terms of localization accuracy.

Since two different datasets are used in two schemes, respectively. And the simulation settings are slightly different. The effects of UK-DRL and UKM-DRL are shown in separated sections. For each technique, simulation settings are depicted first followed by detailed simulation results.

5.1 UK-means-based Clustering and Deep Reinforcement Learning-based Radio Resource Allocation

5.1.1 Simulation Settings

The simulation includes one gNB and 6 users. The radius of the gNB is 160 meters. The traffic of users is generated following Poisson distribution with λ inter-arrival time and a packet size of 32 bytes. During the simulation, the number and angle of beams are predefined. Three beams with 20° angle are formed. The QoS requirement of latency T^{QoS} is equal to 1 ms, and the QoS requirement for SINR S^{QoS} is set to 15 dB [85]. Other settings of DRL are shown in Table 5.1. Too frequent clustering will have a negative impact on

RRM. As a result, clustering is performed every 100 TTIs. The total simulation time is 1400 TTIs. The simulation is repeated 5 times, and the confidence interval is 95%.

The exact location is provided by Raymobtime dataset [77] and the distorted location is provided by the predicted location from [80]. The error between the predicted location and the exact location is measured by Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE). In our simulation, RMSE is assumed to be the radius of the circle in the PDF.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n_d}(\hat{\boldsymbol{y}} - \boldsymbol{y})^2}{n_d}} \tag{5.1}$$

$$MAE = \frac{\sum_{i=1}^{n_d}|\hat{\boldsymbol{y}} - \boldsymbol{y}|}{n_d} \tag{5.2}$$

where n is the total number of data, $\hat{y}$ and y represent predicted data and exact data, respectively.

Moreover, the location information is presented in 3D format. Since the value of the z-axis represents the height of objects, the value is small compared to the value of the x-axis and y-axis. The location information applied in our simulation is converted to 2D by discarding the value of the z-axis. The total number of predicted location data, i.e, the number of data divided to the testing set in [80] is 2938, which is 35% of the total number of data. The distributions of the test data and the corresponding predicted value are shown in Fig 5.1 and Fig 5.2.

The mobility of users reflect the mobility model in the dataset. 60 of them are selected to represent the trajectory of 6 UEs in 10 time periods. In the simulation, the location is updated every 10 TTIs, which means the users remain in the same place for 10 TTIs. The distribution of selected location information is shown in Fig 5.3 where the red point is gNB and blue points represent users.

Table 5.1: Simulation Parameters

Physical Layer	
Bandwidth	20 MHz
Subcarrier spacing	15 kHz
TTI size	$\frac{1}{7}$ ms
Network Model	
Number of URLLC per cluster	1
Number of eMBB per cluster	1
Number of clusters	3
Radius of cell	160m
Beam angle	20°
N_t	1024
Q-Learning	
Learning rate	0.5
Discount factor	0.9
Exploration probability	0.1
LSTM	
Size of input layer	1
Number of hidden units	20
Size of output layer	24
Size of mini-batch	20
Size of replay memory	60
Training Interval	60
Copy Interval	120
Simulation Parameters	
Simulation time	0.2s
Number of TTIs in Every Run	1400
Traffic Load	[1 2 3 4] Mbps
Number of Clusters to Form	3
Number of Runs	5
Confidence Interval	95%

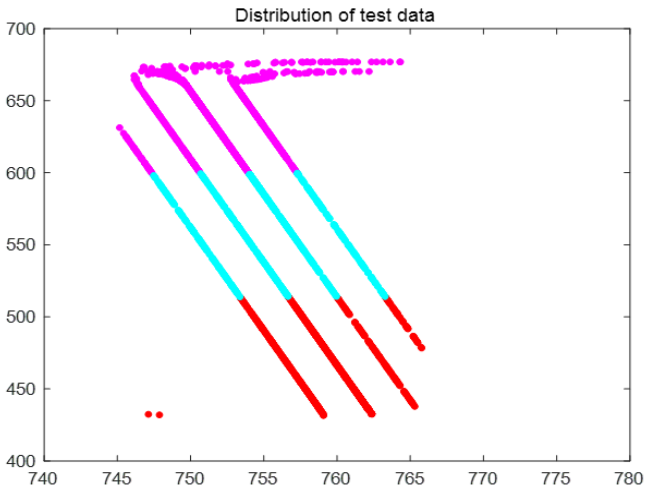


Figure 5.1: Distribution of test data

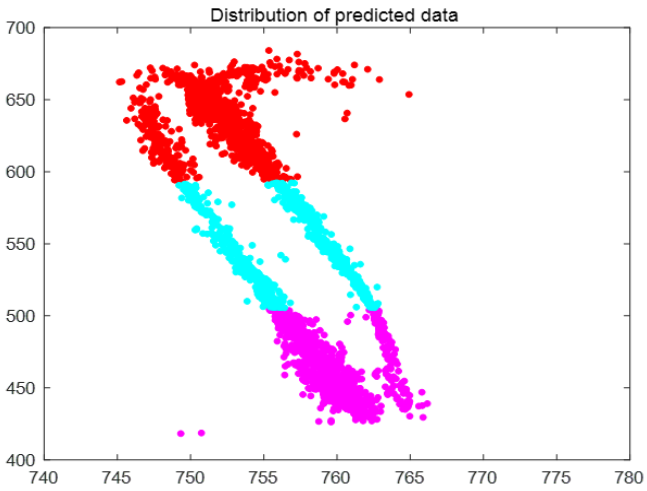


Figure 5.2: Distribution of predicted data

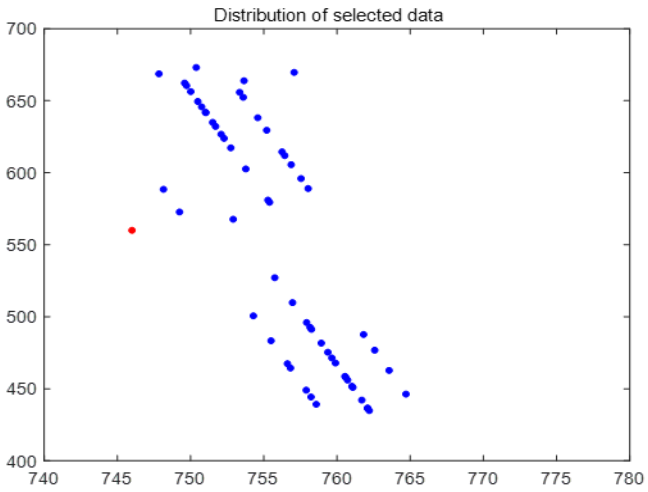


Figure 5.3: Distribution of selected data

Three scenarios are defined to examine the performance of the proposed technique:

- Scenario 1: In this scenario, we adapt the predicted position given by [80] as locations with error. We applied K-means to cluster users with location error, and DRL for resource allocation. The scenario is named by K-DRL+Location with Error.
- Scenario 2: Compared with scenario 1, the only difference in scenario 2 is that we use UK-means to group the users with localization error. The scenario is named by UK-DRL+Location with Error.
- Scenario 3: In this scenario, we use the exact location provided by Raymobtime s009 dataset for clustering, which is considered an optimal baseline. This scenario is named by K-DRL+Exact Location.

The aim of the simulation is to prove the superiority of UK-DRL over K-DRL when location error appears. The number of clusters and beam angle need to be designed to not cover the whole region. There is no need to handle location uncertainty if enough beams are formed and cover the whole region all the time. The results of coverage rate are shown in the next section. In the following results for coverage rate, it is shown that for any number of beams, UK-means achieves higher coverage rate than K-means when distorted data is applied. We let K-means form three clusters because the difference of coverage rate is the largest when beam number is set to three. In this way, we can have more obvious results on sum rate and latency.

5.1.2 Results

Results for Coverage Rate

After clustering, one beam is responsible for serving one cluster, then it can be easily found whether one UE is covered by a beam. The coverage rate is computed by dividing the number of covered users by the total number of users. In our experiment, we firstly compare the number of beams to achieve full coverage and the average coverage rate under different beam angles when the number of beams is fixed to 4.

Fig. 5.4 shows the coverage rate under a different number of beams when the beam angle is 30°. In order to achieve full coverage, the third scenario needs 6 beams due to the exact locations. When distorted locations are applied, the UK-means approach requires 7 beams. However, the K-means-based method needs 9 beams to fully cover all the UEs.

On the other hand, we investigate the coverage rate against different beam angles in Fig. 5.5. The beam number is fixed to 4 beams. Scenario 3 still has the best coverage rate because of the exact location. When distorted locations are applied, the localization error will decrease the coverage rate. Compared with K-means, UK-means has an improvement of 7% when the beam angle is 20°. Fig. 5.4 and 5.5 demonstrate that we can achieve a better coverage rate with the same number of beams by using UK-means method when localization error exists.

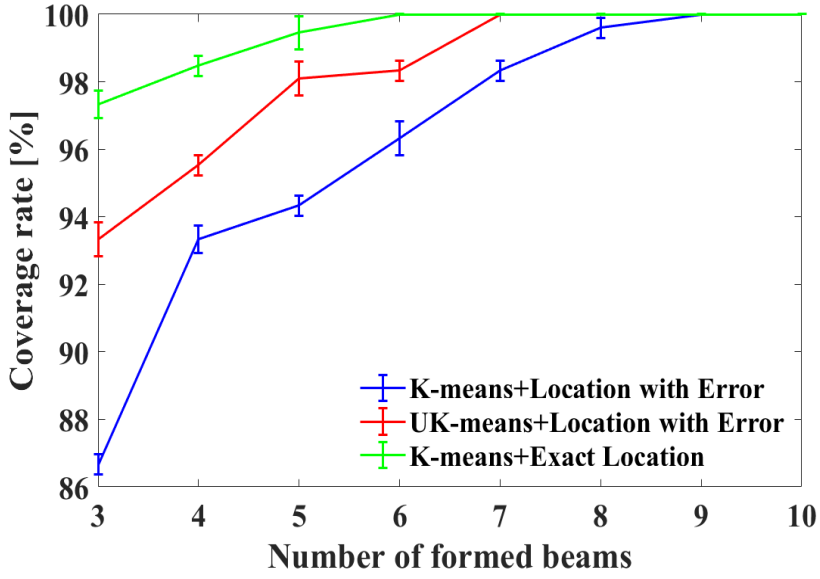


Figure 5.4: Coverage rate comparison under varying number of beams

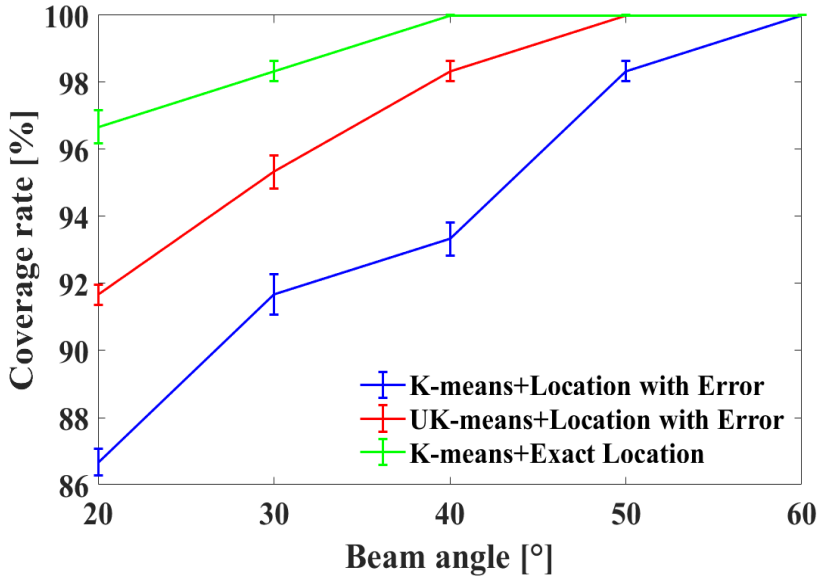


Figure 5.5: Coverage rate comparison under varying beam angle

Results for Resource Allocation

In this section, we display the displayed simulation results of the three scenarios under different traffic loads. The network performance is evaluated by the total data rate and average delay.

Fig. 5.6 and Fig. 5.7 present the sum rate and delay under different traffic loads. As shown in the plots, scenario 3 has the best performance. In scenario 3, since the gNB has the exact position of UEs and the beams are formed according to exact locations, the gNB can cover most of the UEs, leading to the highest sum rate and lowest latency. Moreover, when location error exists, UK-means proves its superiority over K-means in both sum rate and delay. In scenarios 1 and 2, gNB only has the knowledge of location with error, and the distorted positions are used to identify the direction of beams, which will degrade the system performance. However, our proposed UK-DRL can mitigate the effect of localization uncertainty by using the PDF, which will provide a more accurate clustering result, and consequently it achieves a better network performance.

When the traffic load is 1 Mbps, UK-DRL has a 138.6% improvement in sum rate and 60.6% improvement in latency. When the traffic load is 2 Mbps, UK-DRL has a 148.5% improvement in sum rate and 42.2% improvement in latency. When the traffic load is

3 Mbps, UK-DRL has a 156.1% improvement in sum rate and 62.1% improvement in latency. When the traffic load is 4 Mbps, UK-DRL has a 172.4% improvement in sum rate and 64.1% improvement in latency.

However, as the traffic load increases, it will stress all three algorithms. Another interesting point is that for each scenario, the delay has not changed much as the traffic load grows. The reason is that we assume users to move every 10 TTIs. For the UE that is not covered, after at least 10 TTIs, it will move to a new place and is very likely to be covered.

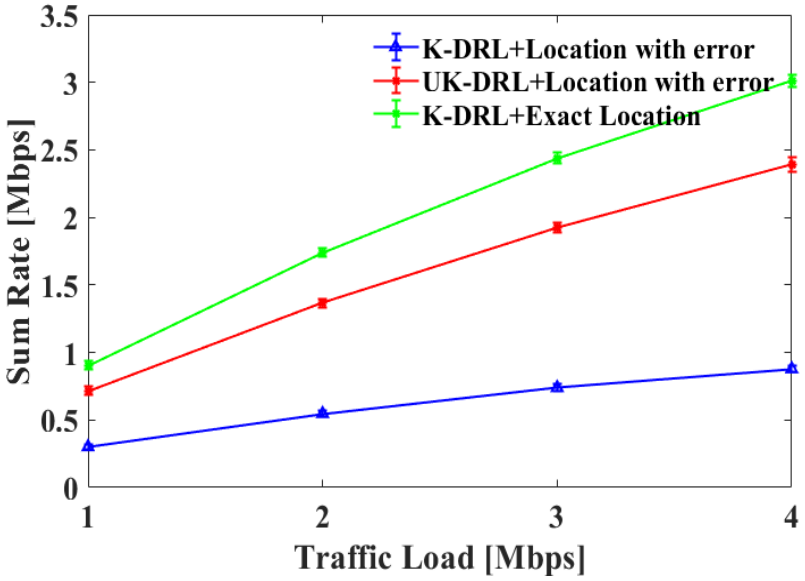


Figure 5.6: Sum rate comparison under varying traffic loads

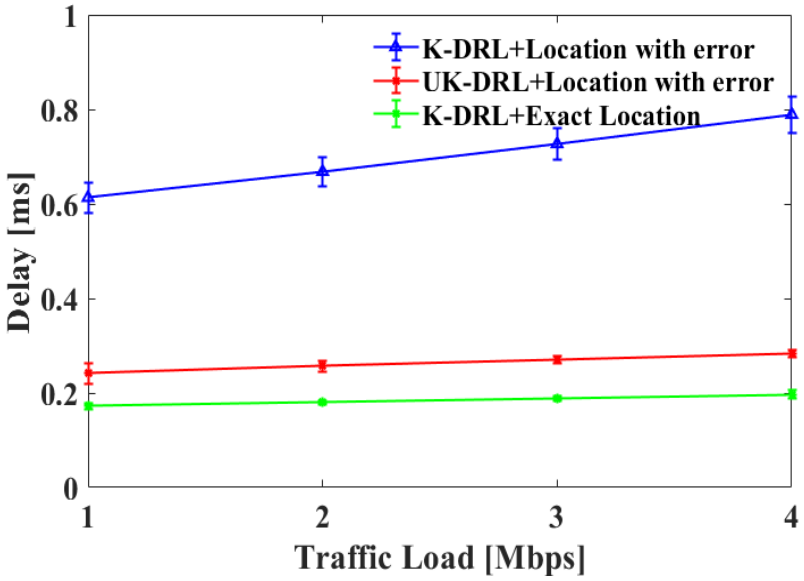


Figure 5.7: Delay comparison under varying traffic loads

5.2 UK-medoids based Clustering and Deep reinforcement learning based radio resource allocation

5.2.1 Simulation Settings

Localization settings and dataset

The mobile communication system measurements and satellite images dataset [78] is divided into 4 regions and each region has its own FFNN to predict the UE locations. The FFNN architecture is $128 \times 64 \times 32 \times 16 \times 2$. Sigmoid activation functions are employed for hidden layers, whereas linear activation functions are used for output layers. Other settings of FFNN including epochs, batch size and dropout are shown in Table 5.2.

Table 5.2: Simulation settings of FFNN

	Region 0	**Region 1**	**Region 2**	**Region 3**
Epochs	5000	3000	3000	5000
Batch Size	10	10	10	10
Dropout	0.2	N/A	0.2	N/A

Resource Allocation Settings

We perform the simulation on MATLAB 5G Toolbox. The mmWave network consists of 1 gNB and 6 UEs, and the clustering is implemented every 10 TTIs based on the UE trajectories. The traffic of users follows Poisson distribution and the packet size is 32 bytes. Three beams with 60° angle are predefined. The total simulation time is 1400 TTIs, and each TTI is 1/7 ms. The simulation is repeated 5 times, and the confidence interval is 95%. The detailed information of parameters' settings are shown in Table 5.3.

Table 5.3: Simulation Parameters

Physical Layer	
Bandwidth	20 MHz
Subcarrier spacing	15 kHz
TTI size	$\frac{1}{7}$ ms
Network Model	
Number of URLLC per cluster	1
Number of eMBB per cluster	1
Number of clusters	3
Beam angle	60°
Q-Learning	
Learning rate	0.5
Discount factor	0.9
Exploration probability	0.1
LSTM	
Size of input layer	1
Number of hidden units	20
Size of output layer	24
Size of mini-batch	20
Size of replay memory	60
Training Interval	60
Copy Interval	120

In the same way stated in section 5.1.1, for users' positions, the location provided by the dataset [78] is considered as exact location, while the localization results provided by [79] are considered as location with error for comparison. From thousands of predicted locations, 60 pieces of location information are chosen to apply in the simulation.

Moreover, the location information is given by longitude and latitude. Therefore, a transformation is required to project location information to the plane coordinate system. Miller projection is applied to complete the conversion.

$$x = \Lambda - \Lambda_0, \tag{5.3}$$

$$y = \frac{5}{4}\ln[\tan(\frac{1}{4}ß + \frac{2}{5}Œ)]. \tag{5.4}$$

Here, x and y are the plane coordinates of a projected point, Λ is the longitude of a point on the globe, Λ_0 is central longitude used for the projection, and ϕ is the latitude of the point on the globe.

We include four scenarios in the simulations as follows:

- Scenario 1: We applied K-means for clustering and DRL for resource allocation. The scenario is summarized by K-DRL+Location with Error.
- Scenario 2: Compared with scenario 1, the only difference is that we use UK-means for clustering. It updates centers as the expected value over PDF and the distance between objects are also the expected distance. The scenario is named by UK-DRL+Location with Error.
- Scenario 3: We deploy our proposed UK-medoids based clustering method and DRL for resource allocation. The scenario is named by UKM-DRL+Location with Error.
- Scenario 4: In scenario 4, we use the exact location and K-means for clustering, which is considered as an optimal baseline. This scenario is named by K-DRL+Exact Location.

Localization and coverage results

The localization accuracy is measured by RMSE and MAE. The computation of RMSE and MAE follows equation (5.1) and (5.2), respectively.

As this dataset provides the location information in longitude and latitude format, equation (5.5) and (5.6)is applied to calculate the distance between two points. Then RMSE and MAE are computed accordingly.

$$\Delta d = 2R_E \cdot atan2(\sqrt{a}, \sqrt{1-a}), \tag{5.5}$$

$$a = \sin^2(\Delta_1/2) + \cos(Œ_1)\cos(Œ_2)\sin^2(\Delta_2). \tag{5.6}$$

where R_E is the radius of the earth which is equal to 6371Km. Δ_1 and Δ_2 represent the difference between two latitudes and longitudes, respectively. ϕ_1 and ϕ_2 represent the latitude of the first point and the second point. $atan2$ is the arctangent function.

The results of the two measurements are shown in Table 5.4. With the added new features, the proposed method reduces the RMSE to less than 4 meters in each region. For the total localization results, RMSE is reduced from 17.11 meters to 3.62 meters, MAE is reduced from 7.83 meters to 1.51 meters. The main reason is that our vision-aided new features lead to higher accuracy in prediction results. In our simulation, we select the location randomly from the whole testing dataset, RMSE is regarded as the radius of the circle for localization uncertainty. And we will examine the network performance under both errors.

Table 5.4: Measurements of Localization Results

Results	Method in [79]		Proposed method	
	RMSE	MAE	RMSE	MAE
Region 1	11.68	6.03	3.82	1.78
Region 2	23.11	10.68	3.52	1.21
Region 3	13.68	8.85	1.38	0.97
Region 4	16.72	6.12	3.75	1.54
Total	17.11	7.83	3.62	1.51

After clustering, one beam is employed to serve each cluster, then it can be easily found whether one UE is covered by a beam. Then coverage rate is computed by dividing the number of covered users by the total number of users. Fig. 5.8 shows the coverage rate with the different number of beams when the beam angle is set to 60°. Scenarios 3 and 4 need 5 beams to achieve full coverage while scenario 1 and scenario 2 need 6 beams. Since the exact location is applied, the coverage rate of scenario 4 is the highest. With the distorted location, UK-medoids shows its strength when compared with UK-means. Meanwhile, K-means has the lowest coverage rate. Fig. 5.9 presents the average coverage rate under different beam angles when the number of beams is 5. It reveals the same conclusion as Fig. 5.8. One can observe that when beam angle is 20°, UK-medoids has an improvement of 7.7% than UK-means.

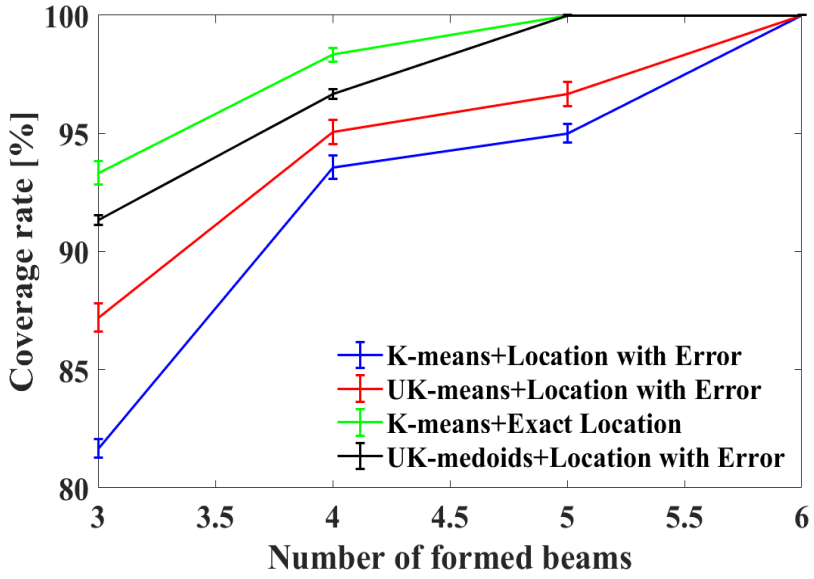


Figure 5.8: Coverage rate comparison under varying beam angle.

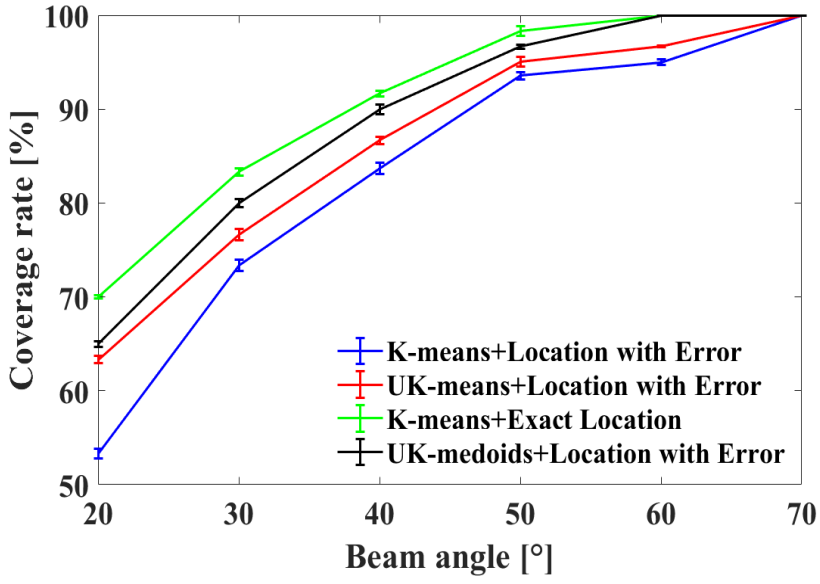


Figure 5.9: Coverage rate comparison under varying beam angle

Resource Allocation

In this section, we compare the network performance under different traffic loads and algorithms. As aforementioned, the error is 17.11 meters by using the localization method from [79], and the 3.62 meters error is achieved by the proposed method. We present the results under both types of errors to better illustrate the advantage of the proposed UKM-DRL method.

Fig. 5.10 shows the sum rate of different scenarios under various traffic loads when localization error is 17.11 meters and 3.62 meters. Under both larger error and smaller error, scenario 4 achieves the highest throughput as our optimal baseline. The reason is that exact locations allow beams to cover most users without any uncertainty. Under both localization errors, UKM-DRL outperforms UK-DRL and K-DRL. K-DRL produces the worst outcome when localization uncertainty is considered due to the lack of ability to deal with uncertainty. Moreover, scenarios 1 to 3 can all benefit from reducing the localization error, which is indicated by a higher sum rate. Scenario 4 remains the same performance because the location information applied is exact location. It further demonstrates the capability of our proposed vision-aided localization method.

For the larger error, UKM-DRL has an improvement of 57.3% and 302.6% over UK-DRL and K-DRL when the traffic load is 1 Mbps, respectively. When the traffic load is 2 Mbps, UKM-DRL has an improvement of 26.9% and 216.5% over UK-DRL and K-DRL, respectively. When the traffic load is 3 Mbps, UKM-DRL has an improvement of 43.9% and 232.6% over UK-DRL and K-DRL, respectively. When the traffic load is 4 Mbps, UKM-DRL has an improvement of 57.9% and 245.5% over UK-DRL and K-DRL, respectively. A larger localization error greatly limits the performance of the wireless networks. Even UKM-DRL has the largest improvement when traffic load is 4 Mbps, there is still an enhancement of 124.9% when exact location is utilized.

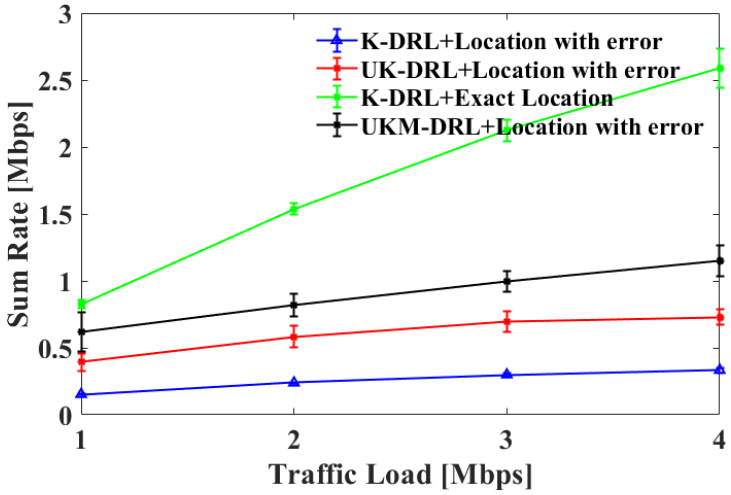
For the smaller error, UKM-DRL has an improvement of 4.2% and 108.8% over UK-DRL and K-DRL when the traffic load is 1 Mbps, respectively. When the traffic load is 2 Mbps, UKM-DRL has an improvement of 11.6% and 136.7% over UK-DRL and K-DRL, respectively. When the traffic load is 3 Mbps, UKM-DRL has an improvement of 29.3% and 184.5% over UK-DRL and K-DRL, respectively. When the traffic load is 4 Mbps, UKM-DRL has an improvement of 17.2% and 231% over UK-DRL and K-DRL, respectively. As the traffic load increases, the improvement effect of UKM-DRL first increases and then decreases when compared with UK-DRL. In the simulation, when one UE has a service request, the ID of this UE is recorded. If the request has no response, a new coming request for the same UE will not be recorded. Therefore, a larger traffic load increases the frequency of generating service requests for the same user in a short period. The ID of

UEs' service list is likely to be fully occupied in both UKM-DRL and UK-DRL, leading to a decrease in the improvement of UKM-DRL over UK-DRL. The problem could be solved by increasing the number of users to lower the frequency of generating service requests for the same user in a short period.

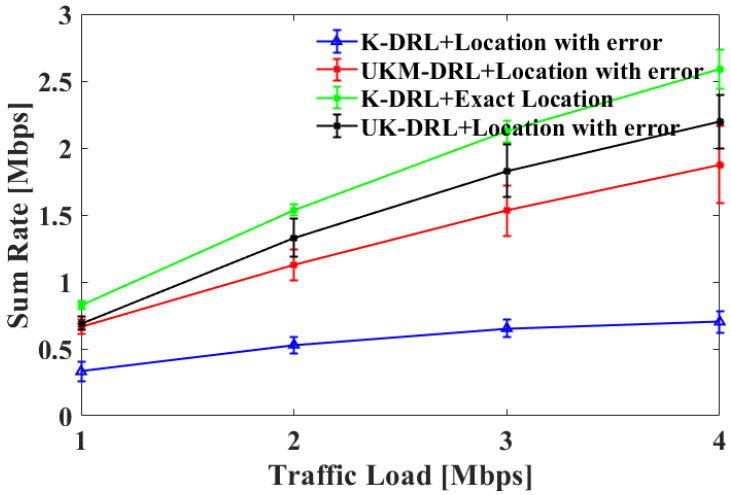
We also investigate in the effect of the value of localization uncertainty. Comparing the sum rate under larger error and smaller error when traffic load is 4 Mbps, UKM-DRL achieves a 90.8% higher sum rate under smaller error. UK-DRL achieves a 157.1% higher sum rate under smaller error. K-DRL achieves a 110.6% higher sum rate under smaller error.

In conclusion, the value of localization uncertainty has a greater impact on the network performance produced by the algorithms than the value of the traffic load. Reducing localization error can improve the performance of UKM-DRL, UK-DRL and K-DRL. UKM-DRL has a critical benefit under large errors.

Similarly, Fig. 5.11 presents the delay under different traffic loads when the localization error is 17.11 meters and 3.62 meters. The result is consistent with Fig. 5.10. The optimal baseline presents the lowest latency. With distorted location, UKM-DRL has lower latency than UK-DRL and K-DRL. K-DRL has the longest latency because it has no ability to mitigate localization uncertainty. When the error is 3.62 meters, the latency of both UK-DRL and UKM-DRL are very close to the optical baseline. But it is still obvious that UK-DRL produces larger latency than UKM-DRL. Compared with the results in Fig. 5.11(a), the latency of scenarios 1 to 3 is greatly reduced.

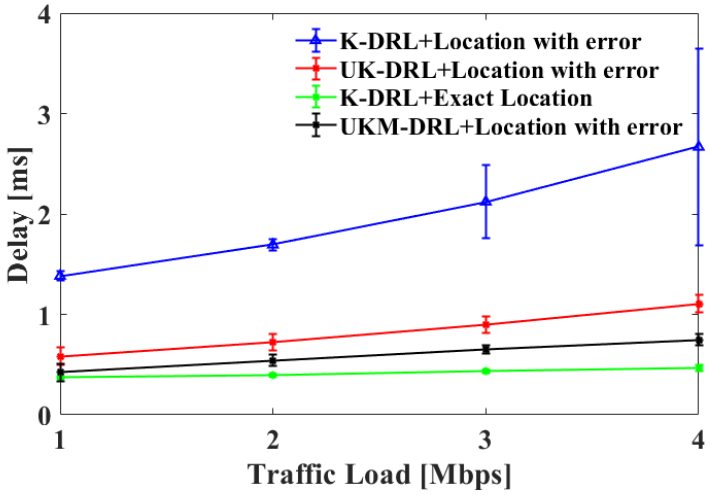


(a) Sum rate comparison when error is 17.11m

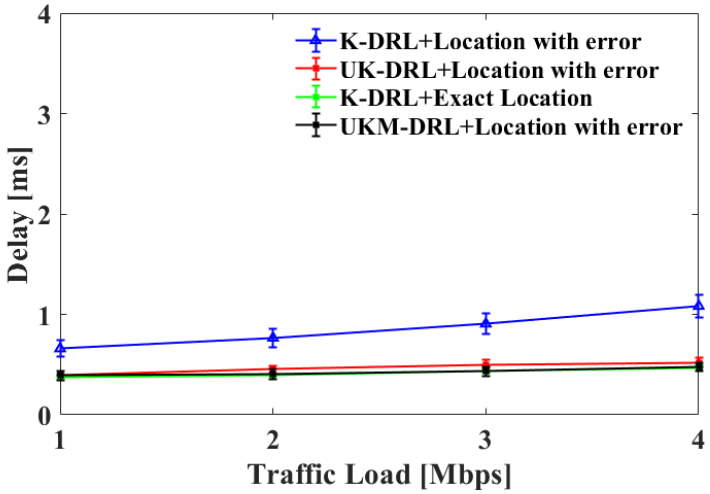


(b) Sum rate comparison when error is 3.62m

Figure 5.10: Sum rate comparison under various traffic loads



(a) Delay comparison when error is 17.11m



(b) Delay comparison when error is 3.62m

Figure 5.11: Delay comparison under various traffic loads

For the larger error, UKM-DRL has an improvement of 27.5% and 69.5% over UK-DRL and K-DRL when traffic load is equal to 1 Mbps, respectively. When the traffic load is 2 Mbps, UKM-DRL has an improvement of 22.3% and 66.7% over UK-DRL and K-DRL, respectively. When the traffic load is 3 Mbps, UKM-DRL has an improvement of 26.8% and 67.9% over UK-DRL and K-DRL, respectively. When the traffic load is 4 Mbps, UKM-DRL has an improvement of 32.4% and 72% over UK-DRL and K-DRL, respectively.

For the smaller error, the average latency of UKM-DRL and UK-DRL is nearly the same when traffic load is equal to 1 Mbps. The average latency of UKM-DRL is 0,3924 ms whereas the average latency of UK-DRL is 0.3930 ms. However, the improvement of UKM-DRL over K-DRL is significant, which is 40.5%. When the traffic load is 2 Mbps, UKM-DRL has an improvement of 7.3% and 43.7% over UK-DRL and K-DRL, respectively. When the traffic load is 3 Mbps, UKM-DRL has an improvement of 17% and 51.4% over UK-DRL and K-DRL, respectively. When the traffic load is 4 Mbps, UKM-DRL has an improvement of 7.7% and 55.8% over UK-DRL and K-DRL, respectively. Same as aforementioned, when comparison to UK-DRL, the beneficial impact of UKM-DRL increases at first and diminishes thereafter.

Comparing the latency under larger error and smaller error when traffic load is 4 Mbps, UKM-DRL achieves a 56.1% lower delay under smaller error. UK-DRL achieves a 113.2% lower delay under smaller error. K-DRL achieves a 146.4% lower delay under smaller error.

Chapter 6

Conclusion and Future Work

6.1 Conclusion

Highly directional beams are tailored in mmWave to compensate for the considerable propagation loss based on users' location. However, location uncertainty is unavoidable in certain settings and thus leading to localization errors. Meanwhile, the heterogeneous service requirements in 5G networks necessitate an efficient RRM method to suit a variety of user needs.

This thesis focuses on solving the problem of localization uncertainty and satisfying QoS requirements of URLLC users and eMBB users in joint beam management and radio resource allocation. The impact of localization error is mitigated via two different clustering techniques and then DRL is applied to allocate radio resource and meet the requirements of two different types of users. The first proposed algorithm is UK-DRL. The second proposed algorithm is UKM-DRL. Two datasets with true measurements from the real world are used for localization to demonstrate the superiority of the proposed algorithms, respectively. The first simulation utilizes Raymobtime s009 dataset [77]. The second simulation utilizes the Mobile Communication System and Satellite Images dataset [78]. The information provided by these two datasets is applied for location prediction. The predicted location is regarded as the distorted location while exact location information is provided by the two datasets. Distorted location is applied in the simulations to present the effect of the proposed techniques. Furthermore, the results are compared with the simulation results using the exact location.

For the Raymobtime s009 dataset, the predicted location is from the work of [80] where CNN is applied. The RMSE between predicted data and exact data is 8 meters, which

is regarded as a parameter for PDF in UK-means. UK-means is applied to the distorted location afterward. Then DRL is adopted in each cluster to assign radio resource aiming to satisfy the QoS requirements of both URLLC users and eMBB users. UK-DRL is compared with K-DRL. Three scenarios are designed to show the excellence of UK-DRL. The results show that when location error exists, UK-means proves its superiority over K-means in both sum rate and delay. When the traffic load is 4 Mbps, UK-DRL has a 172.4% improvement in sum rate and 64.1% improvement in latency.

For the Mobile Communication System and Satellite Images dataset, the predicted location comes from an improved version of [79] where FFNN is applied for localization. Based on [79], more pixel characteristic-based features are extracted from satellite images. With the increased number of input features, the RMSE between predicted location and exact location is reduced from 17.11 meters to 3.62 meters, and MAE is reduced from 7.83 meters to 1.51 meters. A new clustering method called UK-medoids is adopted in clustering and DRL is applied in resource allocation. The second proposed algorithm is named UKM-DRL and it is compared with UK-DRL and K-DRL. The network performances under both larger error and smaller error are investigated. The results show that UKM-DRL outperforms UK-DRL and K-DRL, especially under the larger error. For example, when the traffic load is 4 Mbps, UKM-DRL has an improvement of 57.9% in sum rate and 32.4% in latency over UK-DRL. Moreover, it has an improvement of 245.5% in sum rate and 72% in latency over K-DRL.

In conclusion, the results have demonstrated the promising capability of the two proposed techniques to handle localization uncertainty issue. Moreover, the attempt to introduce vision information in localization has been proved to be successful in localization error reduction. The application of two datasets instead of simulated data proves that the proposed techniques have the potential to be implemented in the real world. However, the possibility to process satellite images real-time is a problem in practical application. Limited to the found dataset, one location data needs to be paired with one image, causing the requirement to take images continuously. But the idea of extracting geomorphic features for accuracy improvement has been proved to be feasible.

6.2 Future Work

The future direction can be divided into two aspects: clustering algorithm and radio resource allocation algorithm.

In this thesis, the number of gNB and active users is relatively small. More complicated network scenarios will be considered with a larger number of gNB and active users. When

the number of gNB increases, the user-cell association will be more complicated. We need to consider the service area of gNBs. Users will move into the coverage area of different gNBs, and the clustering needs to be performed more frequently. Different mobility models will be included to guarantee the robustness of clustering methods that deal with localization uncertainty. In the simulations, 60 UEs' locations are selected from the predicted location to represent the trajectories of users, which means the movement of users follows certain trajectories repeatedly.

On the other hand, the comparatively slow convergence period of ML algorithms sabotages their handiness in highly dynamic wireless networks. Therefore, novel techniques are needed for faster convergence. Apart from the time cost, ML algorithms that applied in highly dynamic wireless networks should be able to be adapt to large datasets and have the ability to update learning parameters with the rapidly changing environment.

Bibliography

[1] J.-S. Sheu, W.-H. Sheen, and T.-X. Guo, "On the design of downlink multiuser transmission for a beam-group division 5g system," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 8, pp. 7191–7203, 2018.

[2] C. V. N. Index, "Global mobile data traffic forecast update, 2016–2021," *white paper*, vol. 7, 2017.

[3] V. W. Wong, R. Schober, D. W. K. Ng, and L.-C. Wang, *Key technologies for 5G wireless systems*. Cambridge university press, 2017.

[4] Y. Niu, Y. Li, D. Jin, L. Su, and A. V. Vasilakos, "A survey of millimeter wave communications (mmwave) for 5g: opportunities and challenges," *Wireless networks*, vol. 21, no. 8, pp. 2657–2676, 2015.

[5] C.-X. Wang, F. Haider, X. Gao, X.-H. You, Y. Yang, D. Yuan, H. M. Aggoune, H. Haas, S. Fletcher, and E. Hepsaydir, "Cellular architecture and key technologies for 5g wireless communication networks," *IEEE communications magazine*, vol. 52, no. 2, pp. 122–130, 2014.

[6] J. Cui, Z. Ding, and P. Fan, "The application of machine learning in mmwave-noma systems," in *2018 IEEE 87th Vehicular Technology Conference (VTC Spring)*, pp. 1–6, IEEE, 2018.

[7] S. Omar and C. Rizos, "Incorporating gps into wireless networks: Issues and challenges," in *The 6th International Symposium on Satellite Navigation Technology Including Mobile Positioning & Location Services*, 2003.

[8] D. Burghal, A. T. Ravi, V. Rao, A. A. Alghafis, and A. F. Molisch, "A comprehensive survey of machine learning based localization with wireless signals," *arXiv preprint arXiv:2012.11171*, 2020.

[9] T. Nishio, Y. Koda, J. Park, M. Bennis, and K. Doppler, "When wireless communications meet computer vision in beyond 5g," *IEEE Communications Standards Magazine*, vol. 5, no. 2, pp. 76–83, 2021.

[10] M. Shafi, A. F. Molisch, P. J. Smith, T. Haustein, P. Zhu, P. De Silva, F. Tufvesson, A. Benjebbour, and G. Wunder, "5g: A tutorial overview of standards, trials, challenges, deployment, and practice," *IEEE journal on selected areas in communications*, vol. 35, no. 6, pp. 1201–1221, 2017.

[11] M. Elsayed and M. Erol-Kantarci, "Ai-enabled future wireless networks: Challenges, opportunities, and open issues," *IEEE Vehicular Technology Magazine*, vol. 14, no. 3, pp. 70–77, 2019.

[12] H. Zhou, M. Elsayed, and M. Erol-Kantarci, "Ran resource slicing in 5g using multi-agent correlated q-learning," in *in Proceedings of 2021 IEEE conference on PIMRC*, pp. 1–6, Jun. 2021.

[13] K. Arulkumaran, M. P. Deisenroth, M. Brundage, and A. A. Bharath, "A brief survey of deep reinforcement learning," *arXiv preprint arXiv:1708.05866*, 2017.

[14] Y. Yao, H. Zhou, and M. Erol-Kantarci, "Deep reinforcement learning-based radio resource allocation and beam management under location uncertainty in 5g mmwave networks," *arXiv preprint arXiv:2204.10984*, 2022.

[15] Y. Yao, H. Zhou, and M. Erol-Kantarci, "Joint vision-aided sensing and communications using deep reinforcement learning-based beam management in 6g networks," *submitted to 2022 IEEE Globecom*, pp. 1–6, May. 2022.

[16] A. U. Gawas, "An overview on evolution of mobile wireless communication networks: 1g-6g," *International Journal on Recent and Innovation Trends in Computing and Communication*, vol. 3, no. 5, pp. 3130–3133, 2015.

[17] A. Gupta and R. K. Jha, "A survey of 5g network: Architecture and emerging technologies," *IEEE access*, vol. 3, pp. 1206–1232, 2015.

[18] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive mimo for next generation wireless systems," *IEEE communications magazine*, vol. 52, no. 2, pp. 186–195, 2014.

[19] A. El-Rabbany, *Introduction to GPS: the global positioning system*. Artech house, 2002.

[20] J. Parthasarathy, "Positioning and navigation system using gps," *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science*, vol. 36, no. Part 6, pp. 208–212, 2006.

[21] S. Vatansever and I. Butun, "A broad overview of gps fundamentals: Now and future," in *2017 IEEE 7th Annual Computing and Communication Workshop and Conference (CCWC)*, pp. 1–6, Ieee, 2017.

[22] R. Bill, C. Cap, M. Kofahl, and T. Mundt, "Indoor and outdoor positioning in mobile environments a review and some investigations on wlan positioning," *Geographic Information Sciences*, vol. 10, no. 2, pp. 91–98, 2004.

[23] D. V. Le, N. Meratnia, and P. J. Havinga, "Unsupervised deep feature learning to reduce the collection of fingerprints for indoor localization using deep belief networks," in *2018 International conference on indoor positioning and indoor navigation (IPIN)*, pp. 1–7, IEEE, 2018.

[24] F. Dou, J. Lu, Z. Wang, X. Xiao, J. Bi, and C.-H. Huang, "Top-down indoor localization with wi-fi fingerprints using deep q-network," in *2018 IEEE 15th International Conference on Mobile Ad Hoc and Sensor Systems (MASS)*, pp. 166–174, IEEE, 2018.

[25] J.-H. Choi, Y.-W. Park, J.-B. Song, and I.-S. Kweon, "Localization using gps and vision aided ins with an image database and a network of a ground-based reference station in outdoor environments," *International Journal of Control, Automation and Systems*, vol. 9, no. 4, pp. 716–725, 2011.

[26] S. Nilwong, D. Hossain, S.-i. Kaneko, and G. Capi, "Deep learning-based landmark detection for mobile robot outdoor localization," *Machines*, vol. 7, no. 2, p. 25, 2019.

[27] Y.-N. R. Li, B. Gao, X. Zhang, and K. Huang, "Beam management in millimeter-wave communications for 5g and beyond," *IEEE Access*, vol. 8, pp. 13282–13293, 2020.

[28] Y. Heng, J. G. Andrews, J. Mo, V. Va, A. Ali, B. L. Ng, and J. C. Zhang, "Six key challenges for beam management in 5.5 g and 6g systems," *IEEE Communications Magazine*, vol. 59, no. 7, pp. 74–79, 2021.

[29] S. Manap, K. Dimyati, M. N. Hindia, M. S. A. Talip, and R. Tafazolli, "Survey of radio resource management in 5g heterogeneous networks," *IEEE Access*, vol. 8, pp. 131202–131223, 2020.

[30] H. Yin and S. Alamouti, "Ofdma: A broadband wireless access technology," in *2006 IEEE sarnoff symposium*, pp. 1–4, IEEE, 2006.

[31] Y. Cai, Z. Qin, F. Cui, G. Y. Li, and J. A. McCann, "Modulation and multiple access for 5g networks," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 1, pp. 629–646, 2017.

[32] M. Elsayed and M. Erol-Kantarci, "Radio resource and beam management in 5g mmwave using clustering and deep reinforcement learning," in *GLOBECOM 2020-2020 IEEE Global Communications Conference*, pp. 1–6, IEEE, 2020.

[33] J. Pérez-Romero, J. Sánchez-González, R. Agustí, B. Lorenzo, and S. Glisic, "Power-efficient resource allocation in a heterogeneous network with cellular and d2d capabilities," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 11, pp. 9272–9286, 2016.

[34] H. Munir, S. A. Hassan, H. Pervaiz, Q. Ni, and L. Musavian, "Energy efficient resource allocation in 5g hybrid heterogeneous networks: A game theoretic approach," in *2016 IEEE 84th vehicular technology conference (VTC-Fall)*, pp. 1–5, IEEE, 2016.

[35] R. Q. Hu and Y. Qian, "An energy efficient and spectrum efficient wireless heterogeneous network framework for 5g systems," *IEEE Communications Magazine*, vol. 52, no. 5, pp. 94–101, 2014.

[36] E. Ali, M. Ismail, R. Nordin, and N. F. Abdulah, "Beamforming techniques for massive mimo systems in 5g: overview, classification, and trends for future research," *Frontiers of Information Technology & Electronic Engineering*, vol. 18, no. 6, pp. 753–772, 2017.

[37] Z. Chen, Z. Ding, X. Dai, and R. Zhang, "An optimization perspective of the superiority of noma compared to conventional oma," *IEEE Transactions on Signal Processing*, vol. 65, no. 19, pp. 5191–5202, 2017.

[38] J. Cui, Z. Ding, P. Fan, and N. Al-Dhahir, "Unsupervised machine learning-based user clustering in millimeter-wave-noma systems," *IEEE Transactions on Wireless Communications*, vol. 17, no. 11, pp. 7425–7440, 2018.

[39] Y. Liu, Z. Ding, M. Elkashlan, and J. Yuan, "Nonorthogonal multiple access in large-scale underlay cognitive radio networks," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 12, pp. 10152–10157, 2016.

[40] D. Marasinghe, N. Jayaweera, N. Rajatheva, and M. Latva-Aho, "Hierarchical user clustering for mmwave-noma systems," in *2020 2nd 6G Wireless Summit (6G SUMMIT)*, pp. 1–5, IEEE, 2020.

[41] I. Orikumhi, H.-K. Jwa, J.-H. Na, and S. Kim, "Location-aided user clustering and power allocation for noma in 5g mmwave networks," in *2020 International Conference on Information and Communication Technology Convergence (ICTC)*, pp. 264–268, IEEE, 2020.

[42] M. Chen, U. Challita, W. Saad, C. Yin, and M. Debbah, "Machine learning for wireless networks with artificial intelligence: A tutorial on neural networks," *arXiv preprint arXiv:1710.02913*, vol. 9, 2017.

[43] Y. Sun, M. Peng, Y. Zhou, Y. Huang, and S. Mao, "Application of machine learning in wireless networks: Key techniques and open issues," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 4, pp. 3072–3108, 2019.

[44] J. A. Anderson, *An introduction to neural networks*. MIT press, 1995.

[45] A. Baraldi and E. Alpaydin, "Constructive feedforward art clustering networks. i," *IEEE transactions on neural networks*, vol. 13, no. 3, pp. 645–661, 2002.

[46] R. Xu and D. Wunsch, "Survey of clustering algorithms," *IEEE Transactions on neural networks*, vol. 16, no. 3, pp. 645–678, 2005.

[47] P. Berkhin, "A survey of clustering data mining techniques," in *Grouping multidimensional data*, pp. 25–71, Springer, 2006.

[48] M. Parimala, D. Lopez, and N. Senthilkumar, "A survey on density based clustering algorithms for mining large spatial databases," *International Journal of Advanced Science and Technology*, vol. 31, no. 1, pp. 59–66, 2011.

[49] M. A. Nielsen, *Neural networks and deep learning*, vol. 25. Determination press San Francisco, CA, USA, 2015.

[50] O. I. Abiodun, A. Jantan, A. E. Omolara, K. V. Dada, N. A. Mohamed, and H. Arshad, "State-of-the-art in artificial neural network applications: A survey," *Heliyon*, vol. 4, no. 11, p. e00938, 2018.

[51] S. Albawi, T. A. Mohammed, and S. Al-Zawi, "Understanding of a convolutional neural network," in *2017 international conference on engineering and technology (ICET)*, pp. 1–6, Ieee, 2017.

[52] R. C. Staudemeyer and E. R. Morris, "Understanding lstm–a tutorial into long short-term memory recurrent neural networks," *arXiv preprint arXiv:1909.09586*, 2019.

[53] M. Chen, U. Challita, W. Saad, C. Yin, and M. Debbah, "Machine learning for wireless networks with artificial intelligence: A tutorial on neural networks," *arXiv preprint arXiv:1710.02913*, vol. 9, 2017.

[54] V. François-Lavet, P. Henderson, R. Islam, M. G. Bellemare, and J. Pineau, "An introduction to deep reinforcement learning," *arXiv preprint arXiv:1811.12560*, 2018.

[55] H. Zhou and M. Erol-Kantarci, "Knowledge transfer based radio and computation resource allocation for 5g ran slicing," in *in Proceedings of 2022 IEEE conference on CCNC*, pp. 617–623, Jan. 2022.

[56] L. P. Kaelbling, M. L. Littman, and A. W. Moore, "Reinforcement learning: A survey," *Journal of artificial intelligence research*, vol. 4, pp. 237–285, 1996.

[57] H. Zhang, H. Zhou, and M. Erol-Kantarci, "Team learning-based resource allocation for open radio access network (o-ran)," in *arXiv:2201.07385*, pp. 1–6, Jan. 2022.

[58] J. Ren, Z. Wang, M. Xu, F. Fang, and Z. Ding, "Unsupervised user clustering in non-orthogonal multiple access," in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3332–3336, IEEE, 2019.

[59] Y. Luo, Z. Shi, X. Zhou, Q. Liu, and Q. Yi, "Dynamic resource allocations based on q-learning for d2d communication in cellular networks," in *2014 11th international computer conference on wavelet actiev media technology and information processing (ICCWAMTIP)*, pp. 385–388, IEEE, 2014.

[60] Y. Zhou, F. Tang, Y. Kawamoto, and N. Kato, "Reinforcement learning-based radio resource control in 5g vehicular network," *IEEE Wireless Communications Letters*, vol. 9, no. 5, pp. 611–614, 2019.

[61] Y. Zhang, C. Kang, T. Ma, Y. Teng, and D. Guo, "Power allocation in multi-cell networks using deep reinforcement learning," in *2018 IEEE 88th Vehicular Technology Conference (VTC-Fall)*, pp. 1–6, IEEE, 2018.

[62] W. Ma, C. Qi, and G. Y. Li, "Machine learning for beam alignment in millimeter wave massive mimo," *IEEE Wireless Communications Letters*, vol. 9, no. 6, pp. 875–878, 2020.

[63] R. Ghanavi, E. Kalantari, M. Sabbaghian, H. Yanikomeroglu, and A. Yongacoglu, "Efficient 3d aerial base station placement considering users mobility by reinforcement learning," in *2018 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1–6, IEEE, 2018.

[64] M. M. U. Chowdhury, W. Saad, and I. Güvenç, "Mobility management for cellular-connected uavs: A learning-based approach," in *2020 IEEE International Conference on Communications Workshops (ICC Workshops)*, pp. 1–6, IEEE, 2020.

[65] J. Zhu, Y. Song, D. Jiang, and H. Song, "A new deep-q-learning-based transmission scheduling mechanism for the cognitive internet of things," *IEEE Internet of Things Journal*, vol. 5, no. 4, pp. 2375–2385, 2017.

[66] D. Burghal, A. T. Ravi, V. Rao, A. A. Alghafis, and A. F. Molisch, "A comprehensive survey of machine learning based localization with wireless signals," *arXiv preprint arXiv:2012.11171*, 2020.

[67] T. S. Rappaport, J. H. Reed, and B. D. Woerner, "Position location using wireless communications on highways of the future," *IEEE communications Magazine*, vol. 34, no. 10, pp. 33–41, 1996.

[68] C. D. McGillem and T. S. Rappaport, "Infra-red location system for navigation of autonomous vehicles," in *Proceedings. 1988 IEEE International Conference on Robotics and Automation*, pp. 1236–1238, IEEE, 1988.

[69] H. Ge, F. Jiang, and Z. Zhang, "A hybrid localization algorithm of rss and toa based on an ensembled neural network," in *2019 IEEE 8th Joint International Information Technology and Artificial Intelligence Conference (ITAIC)*, pp. 1280–1284, IEEE, 2019.

[70] Y. Zhang, Y. Xiao, K. Zhao, and W. Rao, "Deeploc: deep neural network-based telco localization," in *Proceedings of the 16th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services*, pp. 258–267, 2019.

[71] B. Peng, G. Seco-Granados, E. Steinmetz, M. Fröhle, and H. Wymeersch, "Decentralized scheduling for cooperative localization with deep reinforcement learning," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 5, pp. 4295–4305, 2019.

[72] R. S. Sutton, D. McAllester, S. Singh, and Y. Mansour, "Policy gradient methods for reinforcement learning with function approximation," *Advances in neural information processing systems*, vol. 12, 1999.

[73] O. Tuzel, F. Porikli, and P. Meer, "Pedestrian detection via classification on riemannian manifolds," *IEEE transactions on pattern analysis and machine intelligence*, vol. 30, no. 10, pp. 1713–1727, 2008.

[74] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 580–587, 2014.

[75] A. Alahi, A. Haque, and L. Fei-Fei, "Rgb-w: When vision meets wireless," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 3289–3297, 2015.

[76] J. Biswas and M. Veloso, "Depth camera based indoor mobile robot localization and navigation," in *2012 IEEE International Conference on Robotics and Automation*, pp. 1697–1702, IEEE, 2012.

[77] A. Klautau, P. Batista, N. González-Prelcic, Y. Wang, and R. W. Heath, "5g mimo data for machine learning: Application to beam-selection using deep learning," in *2018 Information Theory and Applications Workshop (ITA)*, pp. 1–9, IEEE, 2018.

[78] J. Thrane and H. L. Christiansen, "Mobile communication system measurements and satellite images," 2019.

[79] H. Mukhtar and M. Erol-Kantarci, "Satellite image and received signal-based outdoor localization using deep neural networks," in *2021 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, pp. 1–6, IEEE, 2021.

[80] H. Mukhtar and M. Erol-Kantarci, "Machine learning-enabled localization in 5g using lidar and rss data," in *2021 IEEE Symposium on Computers and Communications (ISCC)*, pp. 1–6, IEEE, 2021.

[81] A. Alkhateeb, G. Leus, and R. W. Heath, "Limited feedback hybrid precoding for multi-user millimeter wave systems," *IEEE transactions on wireless communications*, vol. 14, no. 11, pp. 6481–6494, 2015.

[82] S. Chowdhury, B. Verma, M. Tom, and M. Zhang, "Pixel characteristics based feature extraction approach for roadside object detection," in *2015 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–8, IEEE, 2015.

[83] M. Chau, R. Cheng, B. Kao, and J. Ng, "Uncertain data mining: An example in clustering location data," in *Pacific-Asia conference on knowledge discovery and data mining*, pp. 199–204, Springer, 2006.

[84] F. Gullo, G. Ponti, and A. Tagarelli, "Clustering uncertain data via k-medoids," in *International Conference on Scalable Uncertainty Management*, pp. 229–242, Springer, 2008.

[85] N. I. B. Hamid, N. Salele, M. T. Harouna, and R. Muhammad, "Analysis of lte radio parameters in different environments and transmission modes," in *2013 International Conference on Electrical Information and Communication Technology (EICT)*, pp. 1–6, IEEE, 2014.