UNIVERSITY OF LEEDS

Handover Techniques for High-Speed Rail Wireless Communications

Siling Wang

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

University of Leeds

Faculty of Engineering

School of Electronic and Electrical Engineering

February 2024

Declaration

The candidate confirms that the work submitted is her own and that appropriate credit has been given where reference has been made to the work of others. The material contained in the chapters of this thesis has been previously published in research articles written entirely by the author of this thesis who also appears as the lead author in all published papers. The research has been supervised and guided by Dr. Li Zhang who appears as the co-author in all the published articles. All the materials included in this document are of the author's intellectual ownership.

The work in Chapter 4 of the thesis has appeared in publication as follows:

- S. Wang and L. Zhang, "A Multi-Criteria Handover Scheme for Distributed Antenna System Based High-Speed Rail Wireless Communications," *2021 24th International Symposium on Wireless Personal Multimedia Communications (WPMC)*, Okayama, Japan, 2021, pp. 1-6.

The work in Chapter 5 of the thesis has appeared in publication as follows:

- S. Wang and L. Zhang, "Q-learning based Handover Algorithm for High-Speed Rail Wireless Communications," *2023 IEEE Wireless Communications and Networking Conference (WCNC)*, Glasgow, United Kingdom, 2023, pp. 1-6.

The work in Chapter 6 of the thesis is based on the following submitted revised paper.

- S. Wang and L. Zhang, ”Support Vector Machine based Handover Scheme for Heterogeneous Ultra Dense Network of High-Speed Railway,” in *IET Communications*, Submitted.

Dr Li Zhang is the co-author for all the publications listed above. These publications have been written under her supervision, benefiting from excellent technical advice and editorial, patient guidance and valuable feedback.

Acknowledgements

First and foremost, I would like to express my sincere appreciation and thanks to my supervisor Dr. Li Zhang for all her valuable support during my PhD study from the first day. Thank you for all the knowledge transferred to me and all the confidence you put in me, which motivated my evolution in the PhD. I am extremely thankful to my supervisor for her study guidance, support, and life suggestions especially during the COVID-19 pandemic to motivate me constantly. Although there is much more to learn, I am profoundly grateful for the wealth of knowledge and guidance I have received from my supervisor. Her support has been invaluable, and I am truly thankful for the opportunity to learn under her mentorship.

I would like to express my deepest gratitude to my parents for their endless love and support throughout my academic journey. Their encouragement, love, and sacrifices have been my driving force. Their belief in my abilities has inspired me to persevere and achieve my goals. Words cannot express how grateful I am to them. Thank you from the heart.

My gratitude also goes to my beloved twin sisters for her unconditional love and support that have always been my strength. Their love and support have been my anchor, and I dedicate this achievement to them with immense gratitude.

I also thank all of my friends from the ICaPNet research group for their support and motivation during my PhD study. In addition, the shared moments of challenges and triumphs have made the process memorable and enriched my learning experience. Their insights, discussions and encouragement have been invaluable to me. I am grateful for the intimate friendship and lucky to work with wonderful friends like you.

Abstract

In recent years, the development of high-speed railways (HSRs) has emerged as a transformative force in the realm of transportation, revolutionizing the way people and goods move across regions. HSRs improve the quality of service (QoS) and customer satisfaction on the trip, creating enormous social and economic benefits. As a result of mobility, train users will have to perform handover between base stations (BSs) to maintain service continuity. However, the increased train speed and the trend of dense distribution of BSs in railway systems pose a set of challenges to wireless communication between the train and the ground. Many problems remain unsolved, such as the high probability of handover failure, frequent handovers and a large system overhead caused by group handovers, etc. Therefore, it is of vital significance to improve the handover performance and guarantee communication quality and service demand.

The purpose of this thesis is to investigate the existing literature works and then propose effective handover techniques to address the problems mentioned above in HSR wireless communication systems.

Firstly, the performance of the handover is analyzed, by considering dual-layer and distributed antenna system (DAS) architecture, and an improved handover decision scheme employing multi-criteria is proposed. The handover performance is theoretically analyzed in terms of handover trigger probability, handover success probability and outage probability. The performance obtained via Matlab simulation coincides with the theoretical analysis perfectly and they both show that the proposed scheme can achieve a higher probability of handover success in comparison with the conventional RSRP based handover decision algorithm.

Then, a Q-learning algorithm based handover method is proposed for the HSR network to ensure robust wireless connectivity for train users. By using reinforcement learning, HO decisions are dynamically optimized in response to changing environments, achieving more adaptive and efficient handover decisions, and consequently improving overall network performance and user experience. The proposed framework leverages the reference signal received power (RSRP), signal-to-interference-plus-noise ratio (SINR) and time of stay (ToS) to provide effective HO rules for seamless connectivity while considering HO signaling overhead. Furthermore, as demonstrated by the simulation results, the Q-learning based handover method outperforms the conventional approach and existing algorithms in terms of throughput, HO number, HO success probability and HO latency.

Finally, due to the presence of a massive number of mobile devices, we study the handover management of an ultra-dense Network in HSR. An intelligent HO method based on Kalman filter and machine learning algorithms for 5G heterogeneous ultra-dense network of HSR is proposed. This HO method adopts a two-step methodology. In the first step, it uses a Kalman filter to predict the future signal quality of the neighboring BSs. In the second step, the HO decision can be optimized by utilizing support vector machine (SVM) algorithm to select the target cell. The proposed SVM based handover method determines the optimal hyper-plane using labeled training data to classify the available target cells by jointly considering multiple HO criteria. In addition, the analysis and simulation results show that the proposed algorithm can improve the HO performance in terms of the number of HOs, handover failure probability, user mean throughput and handover latency compared to other machine learning approaches and existing methods in the literature.

Keywords: high mobility, high-speed railway wireless communication, handover, 5G, machine learning, reinforcement learning

Contents

List of Abbreviations

3GPP Third Generation Partnership Project

4G Fourth Generation

5G Fifth Generation

5G-R 5G-Railway

AI Artificial Intelligence

ANN Artificial Neural Network

AP Access Point

BBU Building Base band Unit

BER Bit Error Rate

bps Bits Per Second

BS Base Station

CBTC Communication-Based Train Control

CDFs Cumulative Distribution Functions

CDMA Code Division Multiple Access

CU Central Unit

DAS Distributed Antenna System

DC Dual Connectivity

DT Decision Tree

FDD Frequency Division Duplex

FSW Frequency Switch

GAN Generative Adversarial Network

GM Gray Model

GSM-R Global System for Mobile Communications-Railways

H Hysteresis

HetNets Heterogeneous Networks

HO Handover

HSR High Speed Railway

HST High Speed Train

HTP Handover Trigger Probability

IBH Inter-Beam Handover

ICI Inter Cell Interference

IoT Internet of Things

JOD-IBH Jointly Optimized Dynamic Inter-Beam Handover

KNN K-Nearest Neighbors

LCX Leaky Coaxial Cable

LSTM Long Short-Term Memory

LTE-A Long-Term Evolution Advanced

LTE-R Long-Term Evolution for Railway

MC Moving Cell

MCDM Multiple Criteria Decision Making

MDP Markov Decision Process

MFC Moving Frequency Concept

MIMO Multiple Input Multiple Output

ML Machine Learning

MME Mobility Management Entity

mmWave Millimeter Wave

MRs Mobile Relays

OFDM Orthogonal Frequency Division Multiplexing

PDF Probability Density Function

PoA Point of Attachment

QoS Quality of Service

RAU Remote Antenna Unit

RBF Radial Basis Function

RF Radio Frequency

RFIC Radio Frequency Integrated Circuit

RL Reinforcement Learning

RLF Radio Link Failure

RoF Radio over Fiber

RRH Remote Radio Head

RRU Remote Radio Unit

RSRP Reference Signal Received Power

RSRQ Reference Signal Received Quality

RSS Received Signal Strength

RSSI Received Signal Strength Indicator

SAGIN Space-Air-Ground Integrated Network

SAW Simple Additive Weighting

SD	Standard Deviation
SINR	Signal-to-Interference-plus-Noise Ratio
SNR	Signal-to-Noise Ratio
SON	Self Organizing Network
SVM	Support Vector Machine
T2T	Train to Train
TD	Temporal Difference
ToS	Time of Stay
TRS	Train Relay Station
TTT	Time To Trigger
UDN	Ultra Dense Network
UEs	User Equipments
UIC	International Union of Railways

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

HSRs are playing a key role in mass transportation in the world. People's lives and work are becoming increasingly convenient with the development of HSRs which have witnessed a significant evolution in communication systems, progressing from GSM-R (Global System for Mobile Communications-Railways) to the advent of 5G technology. Initially designed to meet the specific communication needs of railways, GSM-R provides a reliable platform for train control, voice communication, and data transmission. However, as the demand for higher data rates, low-latency communication, and the number of more comprehensive applications grows, the railway industry is transitioning towards more advanced solutions. Integrating 4G LTE (Long-Term Evolution) technologies allows for improved broadband connectivity, supports passenger WiFi services and enhances operational efficiency. With the emergence of 5G, HSR communication systems will experience a revolutionary leap. 5G offers faster data speeds, ultra-low latency, and massive device connectivity, paving the way for innovations such as real-time monitoring, predictive maintenance, and enhanced passenger experiences. In addition, the maximum velocity of trains has already reached 575km/h approximately, which was tested by the French National Rail Corporation [21]. With the development of HSR, some of the challenges in HSR wireless communications include signal fading and fluctuations, bandwidth requirements and handover (HO) issues [22]. With the rapid movement of the train, ensuring consistent connectivity and low-

latency communication becomes more challenging through different network cells. The increasing demand for high-data-rate services poses challenges to the available network capacity and bandwidth resources. In the context of 5G-Railway, the deployment of ultra-dense network (UDN) can be considered as a solution to meet the increasing demand for mobile internet services for large numbers of train passengers. However, ultra-dense small cells can result in more severe frequent HO problems. Furthermore, high-speed trains (HSTs) moving through diverse environments such as tunnels, urban areas, and open landscapes will experience signal fading and fluctuations. These fast variations in signal strength can impact communication quality. One of the most significant challenges is the handover issue. Handover is the process of transferring a data session from one serving cell to another one in the network [23]. Handover problems in HSR include high handover failure probability, frequent handovers and group handovers, etc. For instance, handover failure occurs if the user equipment (UE) fails to connect to the target cell [24] which can lead to interrupted data services and temporary loss of connectivity. This can impact both operational communications for train control and passenger services, degrading the overall reliability of the communication network. The increased speed of trains results in frequent handovers between different BSs as the train moves along the railway track. This can lead to increased signaling overhead, handover delay and potential interruption in services. Additionally, the high mobility of trains in HSR environments necessitates swift and efficient handover mechanisms to maintain continuous communication.
Despite the fact that high-speed train is a convenient and fast public transportation, the high mobility of high-speed trains raises several significant issues mentioned above. Therefore, efficient handover management in HSR communication systems is needed to ensure a seamless and reliable experience for train passengers to overcome the handover issues.

1.2 Objective

To ensure service continuity, the train users on the move have to handover between different BSs. Considering the characteristics of the railway communication systems and the high mobility of trains in the HSR system, this might lead to a large number of handovers and high handover failure rate due to short residence time in the overlapping area, and rapid changes in the channel environment. Robust and seamless mobility in the HSR network is a big challenge. Indeed, the aforementioned handover issues can have a serious impact on rail passengers' user experience, especially when dense small cells are deployed to provide coverage for the HSR system. Thus, it is crucial to develop different mobility management approaches that can solve the handover challenges in HSR systems. These approaches entail considering the characteristics and nature of high-speed railway communication systems. Therefore, the objectives of this project can be summarized as follows:

- **Mitigate Frequent Handovers:** Frequent handovers in HSR communication systems are a main issue, and a proper handover strategy is needed to reduce their occurrence. This research aims to optimize handover algorithms to minimize handover events.

- **Decrease Handover Failure Probability:** Evaluate and improve the robustness of handover procedures to decrease failure rates. Investigate factors contributing to handover failures such as signal interference and signal fading, and propose handover solutions to enhance handover success rates.

- **Reduce Handover Latency:** Explore innovative techniques to reduce handover latency in HSR environments. Investigate the impact of high-speed train movement on latency.

- **Improve Network Performance:** An efficient handover technique is needed to improve network performance such as the system throughput in the HSR communication system.

- **Propose Predictive Approaches:** Explore the feasibility and effectiveness of predictive handover strategies. Employ predictive analysis, machine learning (ML), or other advanced technologies to predict handover events and systematically proactively optimize network resource allocation.

  By achieving these objectives, this research aims to contribute significant improvements in handover techniques in HSR communication systems, in order to achieve seamless and robust communications.

1.3 Major Contribution

This thesis addresses various handover problems in HSR wireless communication systems and major contributions are summarized as follows:

1. Design a handover decision-making algorithm that jointly considers multi-criteria in order to have more precise handover decisions. Theoretical analysis was carried out in terms of handover trigger probability, handover success probability and outage probability. The simulation results coincide with the theoretical analysis perfectly and they both show that the proposed scheme can improve the handover performance.

2. Design a reinforcement learning based handover scheme to reduce the number of handovers and handover delay time, in addition to improving the user experience in terms of throughput and handover success probability. The optimal HO decision can be selected considering multiple criteria as a weighted reward along the railway track.

3. Design a machine learning algorithm based handover method. The HO decision making is considered as a classification problem that takes into account available states that they may have in the HSR network. The proposed intelligent handover method significantly reduces the handover number, handover latency and handover failure, in addition to enhancing the user mean throughput when compared with other machine learning based approaches.

1.4 Thesis Layout

This thesis is organized as follows:

Chapter 2 gives a comprehensive introduction to the handover in wireless communications.

Chapter 3 introduces the HSR wireless communication systems and handover challenges in HSR and then reviews the existing works in this area.

Chapter 4 proposes a multi-criteria handover method to enhance the overall handover

decision process. Then, the handover performance is analyzed theoretically and also evaluated by Matlab simulation. The simulation results coincide with the theoretical analysis perfectly. In addition, the proposed method can reduce the handover failure probability and outage probability in the network by improving the handover decision process.
Chapter 5 proposes a Q-learning based handover method. This work can provide an effective strategy for optimal HO decision making which considers multiple criteria as weighted rewards along the railway track, in order to enhance the overall experience in terms of throughput, HO number, HO success probability and HO latency.
Chapter 6 proposes an intelligent handover method. First, the Kalman filter is used to estimate the posteriori of RSRP for the neighboring BSs, relying on the measurement report as its primary input. Second, the machine learning module considers the output of the Kalman filter, SINR and time criteria of neighboring BSs as its input to determine the most suitable target BS for a handover. We evaluated the efficiency of the proposed HO method in terms of the average number of HOs, HO failure probability, user mean throughput and handover latency.
Finally, chapter 7 draws the conclusions of the thesis and remarks on the future direction of the research.

1.5 List of Publications

The work undertaken in this thesis has resulted in the following publications.

1. S. Wang and L. Zhang, ”A Multi-Criteria Handover Scheme for Distributed Antenna System Based High-Speed Rail Wireless Communications,” **Published** in 2021 24th International Symposium on Wireless Personal Multimedia Communications (WPMC). IEEE, 2021. (**Best Paper Award**)

2. S. Wang and L. Zhang, ”Q-learning based Handover Algorithm for High-Speed Rail Wireless Communications,” **Published** in 2023 IEEE Wireless Communications and Networking Conference (WCNC), Glasgow, United Kingdom, 2023, pp. 1-6.

3. S. Wang and L. Zhang, ”Support Vector Machine based Handover Scheme for Heterogeneous Ultra Dense Network of High-Speed Railway,” **Revised version submitted**

to IET Communications.

Chapter 2

Handover in Wireless Communications

2.1 Introduction

This chapter introduces the handover technology in wireless communications. The rest of this chapter is organized as follows. Section 2.2 introduces the handover techniques. In section 2.3, the commonly used handover criteria are explained. Section 2.4 discusses different handover types. Then, different handover performance metrics are presented in section 2.5. Finally, section 2.6 summarises this chapter.

2.2 Handover Technique

Handover is a process that transfers an ongoing call from one cell to another when a user moves through the coverage area of a cellular system. Handover process occurs between two cells. A handover model is shown in Figure 2.1. A classical handover procedure is mainly composed of three steps as follows:

1) **Handover initiation**: the mobile station measures various criteria such as SINR, capacity and RSRP, etc., of different candidate base stations. Subsequently, the mobile station will send the measurement reports to the source base station.

2) **Handover decision**: the source BS transmits the reports to the central unit (CU). Then, the CU employs the decision-making process by considering resource availability and network load to determine the optimal target BS to maintain the connection.

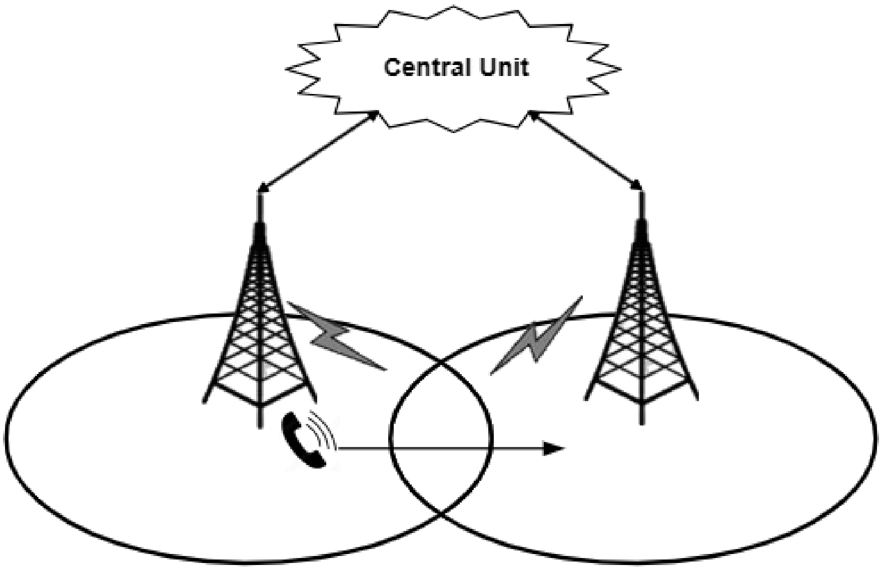

Figure 2.1: Handover model

If a handover is needed, the CU will notify the target BS to prepare for the connection and inform the source BS to send handover commands to the mobile station.

3) **Handover execution**: it is to perform the handover process based on the outcomes of the handover decision-making process. The mobile station establishes a new connection with the selected target base station using resources allocated through operations such as authentication, synchronization, network reconfiguration, etc. Ultimately, the mobile station will send a message to the target BS informing the successful accomplishment of the handover process.

The handover process is illustrated in Figure 2.2.

2.2.1 A1 to A6 Events

In communication systems, there are different events that will occur under different conditions and usage. The A3 event is the conventional HO trigger method in the LTE protocol [25]. The key principle involves the periodic measurement of signals such as RSRP or reference signal received quality (RSRQ) from both the source and target cells. When the received signal strength of the target BS surpasses that of the serving BS by a value denoted as *H*, as represented by equation (2.1), for a duration of time defined as the Time To Trigger (*TTT*), the handover is triggered. If this trigger condition is met within the designated trigger delay *TTT*, the user performs the handover. Fig. 2.3 illustrates the

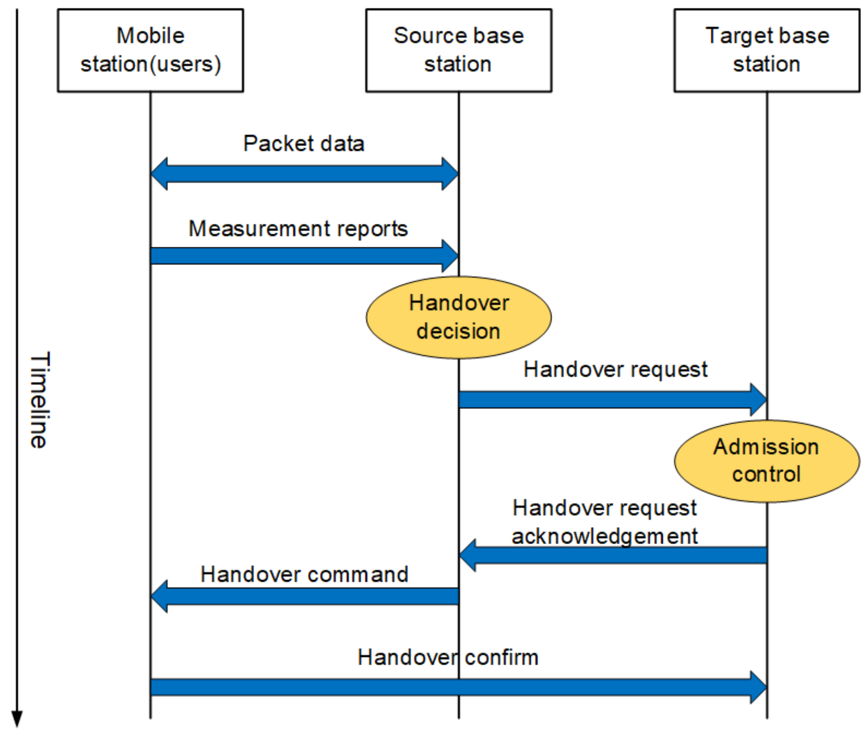

Figure 2.2: Handover process [1]

hysteresis threshold parameter H and handover trigger delay parameter TTT for the A3 event and they are both the handover control parameters.

$$R_t - R_s > H \tag{2.1}$$

Here, R_t and R_S are the received signal strength of the target BS and the serving BS respectively, while H means the hysteresis threshold.

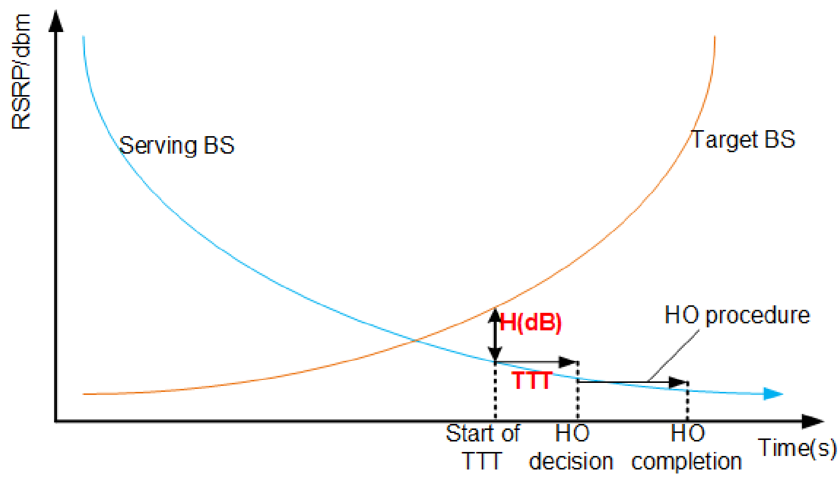

Figure 2.3: Schematic diagram of A3 event algorithm [2]

Parameter *TTT* plays a crucial role in the seamless execution of the handover process. Its primary purpose is to mitigate the occurrence of unnecessary and ping-pong handovers by ensuring that the handover criteria remain stable for a specific time duration. Nonetheless, the *TTT* introduces a delay issue in the handover operation. If the *TTT* duration is too long, it can result in an increased radio link failure (RLF) rate due to the handover delay caused by *TTT*, consequently reducing the overall success probability of handovers. Conversely, an excessively short *TTT* can contribute to the emergence of ping-pong handovers. Thus, paper [26] proposes a fuzzy logic algorithm based approach that aims to dynamically adapt the optimal *TTT* value based on factors such as train speed and cell size, to decrease the RLF while controlling the ping-pong HO simultaneously. In addition, to evade communication disruptions and call drops, the handover process is expected to occur before the train departs from the coverage area of the current serving BS, and a high handover triggering probability is desired.

In addition, the 3GPP standard, No. TS 36.331, introduces various HO events, and their triggering conditions are also explained [27]. Based on this report, A1, A2, A4, A5 and A6 events are summarised as follows:

- **Event A1**: This event is triggered when the signal strength of the serving BS becomes better than a certain threshold. The condition for this event can be represented as

  $$Ms - Hys > Threshold \tag{2.2}$$

  where:
  – Ms is the measurement of the signal received from the serving BS,
  – Hys is the hysteresis parameter relating to A1 event,
  – $Threshold$ is the threshold relating to A1 event.

- **Event A2**: This event is similar to A1, but vice versa. In this case, it is triggered when the signal strength of the serving BS becomes worse than a certain threshold. The condition for this event is expressed as

  $$Ms + Hys < Threshold \tag{2.3}$$

- **Event A4**: This event is triggered when the signal strength received from the target BS becomes better than a certain threshold. The condition for this event can be described as

$$Mt - Hys > Threshold \tag{2.4}$$

where Mt is the measurement of the signal received from the target BS.

- **Event A5**: A5 is a relatively complex event with two thresholds. It is triggered when the signal strength of the serving BS becomes worse than a certain threshold1 and the signal strength of the target BS becomes better than another threshold2. The condition for this event can be defined as:

$$Ms + Hys < Threshold1 \tag{2.5}$$

$$Mt - Hys > Threshold2 \tag{2.6}$$

- **Event A6**: This event occurs when the measurement of the signal received from the target BS becomes greater than that of the signal received from the serving BS. The condition for this event can be depicted as:

$$Mt - Hys > Ms \tag{2.7}$$

Event A6 is similar to event A3, but event A3 can be used for intra and inter-frequency HO whereas event A6 is used only for intra-frequency HO which is explained in section 2.4. Furthermore, we summarize various events and their respective triggering conditions in Table 2.1. Paper [28] has investigated and analyzed various HO

Table 2.1: Events and triggering conditions

Event Type	Description
Event A1	Serving BS becomes better than threshold
Event A2	Serving BS becomes worse than threshold
Event A3	Target BS becomes offset better than serving BS
Event A4	Target BS becomes better than threshold
Event A5	Serving BS becomes worse than threshold1, target BS becomes better than threshold2
Event A6	Similar to A3, but only used in intra-HO

events in terms of HO probability, HO failure probability and probability density function (PDF) of handover success. Their simulation results reveal that the A3 event has the best performance according to the defined indicators in comparison with other events. However, it also shows that increasing speed from low to high speed of 500km/h will increase the handover failure probability by 1.89dB.

Moreover, the choice of threshold value can also greatly impact the performance of the network, and these would typically be optimized based on the specific characteristics of various scenarios. For instance, due to the high speed of users, the handover process needs to be initiated earlier to ensure a seamless transition. This could be achieved by adjusting the threshold values used in the A1-A6 events. However, setting these control parameters is a complex task that needs to consider scenario features and various factors such as the speed of the users, the distance between the BSs and the network load, etc.

2.3 Commonly Used Handover Criteria

The commonly used criteria for handover algorithms in the wireless network are listed as follows:

1) **Reference signal received power (RSRP)**: RSRP is the power of the LTE reference signals which spread over the full bandwidth. It is a key measure of signal level and quality in wireless networks. In cellular networks, when a mobile device moves from one cell to another cell and performs handover, it must measure the signal strength of the neighboring cells. The conventional decision-making process for the HSR network is based on the A3 event handover algorithm [26] [29]. This algorithm makes handover decisions by regularly measuring the RSRP between the source and target cells.

2) **Reference signal received quality (RSRQ)**: RSRQ indicates the quality of the received reference signal. It provides additional information when RSRP is not sufficient to make a reliable handover decision. It is defined as the ratio of RSRP to RSSI,

expressed as

$$RSRQ = N \times \frac{RSRP}{RSSI} \tag{2.8}$$

where N is the number of resource blocks within measurement bandwidth. The received signal strength indicator (RSSI) is the average power of the received signals of the UE in the measurement bandwidth and it is used to indicate the interference condition of the channel. Thus, RSRQ is determined by the strength of reference signals, interference and noise, and thus it is similar to the signal-to-interference-plus-noise ratio of the link [30].

3) **User speed**: it is the widely used criterion to minimize the number of handovers due to high speed moving users in the coverage of cells. Considering this parameter as a criterion for HO decision will improve the HO performance such as less HO number and higher HO success rate, therefore it is a valuable criterion for handovers in HSR wireless communication [31].

4) **Load on target cell**: this criterion is used to measure the available number of resource blocks in the target cell. The use of this criterion can effectively reduce the handover failure. Combining this with RSRP and other parameters can enhance the decision making process, especially in Heterogeneous network (HetNet) [32].

5) **Signal to interference plus noise ratio (SINR)**: this criteria is utilized to measure signal quality and it is defined as the ratio of the signal power to the summation of the average interference power from other cells and the noise power, described as

$$\text{SINR} = \frac{\text{signal power}}{\text{interference power} + \text{noise power}} \tag{2.9}$$

6) **Time of stay (ToS)**: it is the expected time that the users will stay in the coverage area of a BS. When combined with the speed, this criteria can be used to decrease the number of HOs [33].

2.4 Handover Types

Traditionally, in the handover process, users usually tend to connect to the closest cell based on its received signal strength (RSS). Choosing the proper target cell for HO among so many cells is a big challenge. Handover is performed in various ways depending on different factors resulting in different types of HOs. As illustrated in Fig. 2.4, HO can be classified into different types:

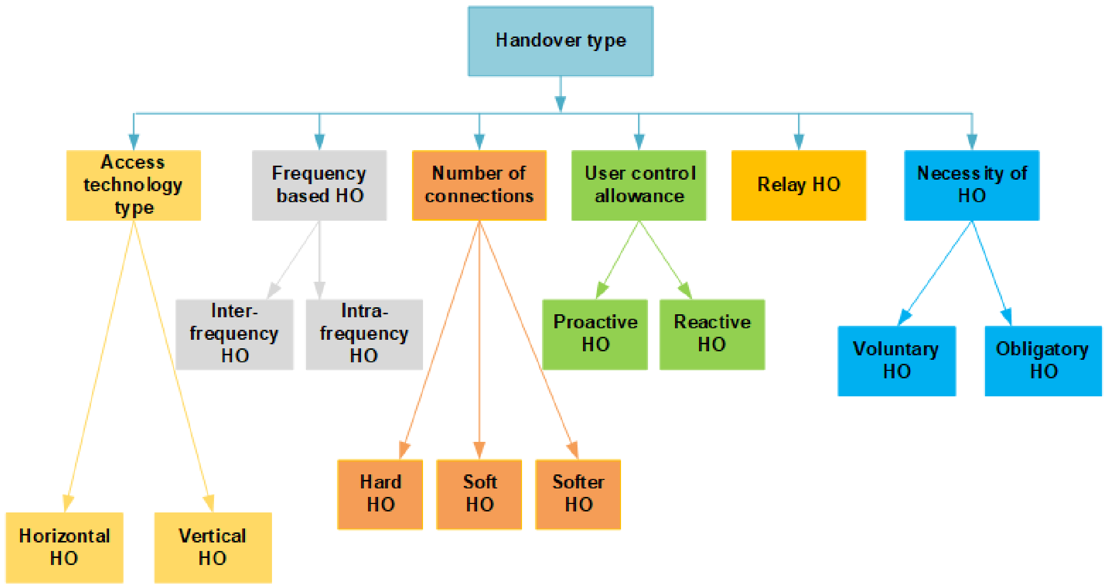

Figure 2.4: Handover types [3]

2.4.1 Access Technologies based Handover

According to the type of access technologies involved, handover can be classified into two types: horizontal and vertical handovers:

a) **Horizontal HO**: usually, this handover happens between two base stations that belong to identical networks such as two geographically neighboring BSs of a 4G cellular network.

b) **Vertical HO**: it occurs between different network interfaces which support various wireless access technologies [34] such as 3GPP and 802.11, etc. Vertical handover can offer numerous advantages. For instance, it is able to seamlessly transit users between different types of subnets while maintaining continuous connectivity and service quality. This technology ensures that users can enjoy uninterrupted services regardless of their location or the network characteristics they are moving into. By merging network features and user demands, vertical handovers provide a smooth and transparent experience, enhancing user satisfaction and optimizing resource utilization. From the network perspective, this can not only have efficient bandwidth usage but also enable effective load balancing, contributing to overall network performance and more reliable service delivery. For specific overviews of vertical handovers, please refer to [35–37].

An example of vertical and horizontal handover in a heterogeneous network is depicted in Fig. 2.5,

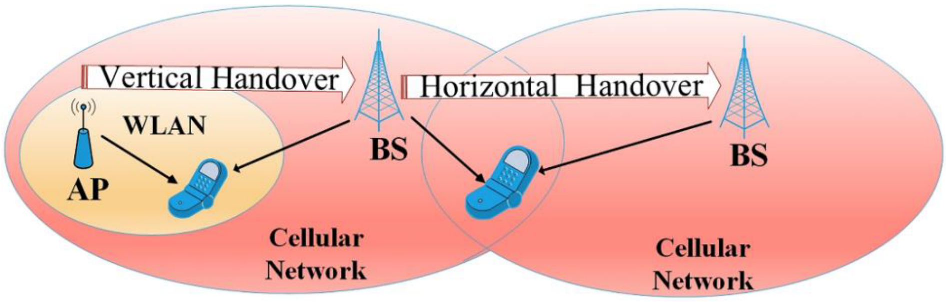

Figure 2.5: Vertical and horizontal handovers in heterogeneous network [4]

2.4.2 Frequency based Handover

This type of handover includes both intra-frequency and inter-frequency HO as depicted in the following [38]:

a) **Intra-frequency HO**: intra-frequency handover refers to the process of transferring communication connections between different BSs operating on the same frequency. For example, it is utilized in code division multiple access (CDMA) with frequency division duplex (FDD).

b) **Inter-frequency HO**: inter-frequency handover refers to the process of transferring communication connections between different BSs operating on different frequency bands. This type of HO is crucial for maintaining seamless connectivity and quality of service as the users move across areas covered by different frequencies in the wireless network.

2.4.3 Number of Connections

Categorized by the number of connections, HO can be classified as hard, soft and softer HO.

a) **Hard HO**: Hard handover, also known as "break before make" handover, is a process in cellular mobile communication systems where a mobile device switches its connection from one BS to another in such a way that the connection is broken before being established with the new BS as shown in Fig. 2.6. This short disconnection occurs on the radio channel to achieve the switch of carrier frequencies. This characteristic makes the hard HO approach more susceptible to call drops. Hard HO is used in GSM-R, unavoidably, this system experiences low data rates and unreliable link connection, which makes it inadequate to meet the broadband communication demands of HSR scenarios [8].

b) **Soft HO**: Soft handover, also known as "make before break" handover, refers to the process in which the user terminal maintains wireless connections with both the source BS and the target BS at the same time as depicted in Fig. 2.6. This allows for a seamless transition between BSs, enhancing the quality of communication connectivity and reducing the probability of call drops or service interruptions. In this way, the communication service quality is substantially improved [39, 40].

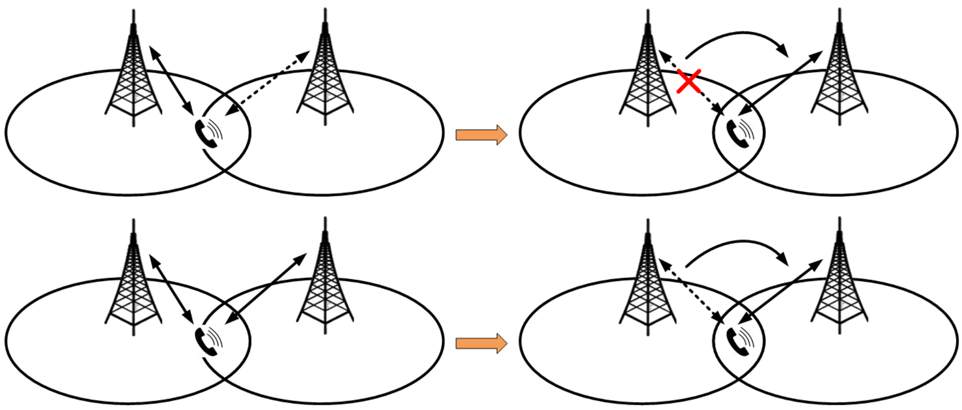

Figure 2.6: Hard and soft handovers

c) **Softer HO**: Another type of handover, known as softer handover, is similar to soft HO, except that the UE changes its radio channel in the same base station [8, 38].

2.4.4 User Control Allowance

According to the user Control Allowance, handover can be classified into two types: proactive and reactive HO:

a) **Proactive HO**: Proactive handover refers to a type of handover mechanism in wireless communication systems where the network initiates the HO process before the received power at the UE side drops below a certain threshold. This approach can decrease call drops, although it comes at the cost of increasing the average number of HOs and impacts the overall performance of the network [33]. In addition, it is especially useful in scenarios with high mobility, such as in vehicular communication, where a fast HO scheme is necessary to ensure continuous connectivity.

b) **Reactive HO**: In contrast to proactive handover, the reactive HO delays the initiation of the HO process as long as possible even after the received power at the UE side drops below a certain threshold. This approach aims to minimize the number of handovers while raising the likelihood of call drops or service interruptions [33].

To fully make use of the benefits offered by both proactive and reactive handovers, finding a balance between these two approaches is essential.

2.4.5 Necessity of Handover

Based on the necessity, HO can be divided into voluntary and obligatory types [38]

a) **Voluntary HO**: This type of HO is that the decision to trigger a handover process is made by the mobile device user itself, rather than being initiated solely by the network infrastructure. Moreover, users actively choose to switch their connection from one cell to another based on factors like personal preferences. It can also help to reduce unnecessary handovers initiated by the network, optimizing the overall user experience.

b) **Obligatory HO**: The UE has to trigger HO from one cell to another, in order to prevent service failure. In this type of HO, the network triggers the handover when specific predefined conditions are met such as signal degradation or coverage area changes.

2.4.6 Relay Handover

Relay handovers can utilize precise positioning technology, where the BS uses the user terminal's distance and direction parameters as supplementary data to decide whether a handover is necessary. This approach enhances the likelihood of successful HO, decreases the number of HOs and reduces overall system load. Moreover, relay HO can even occur between various cells operating on different frequencies [8].

2.5 Handover Performance Metrics

The performance of handover techniques is usually evaluated using the following metrics.

2.5.1 Handover Trigger Probability

Handover trigger probability (HTP) is defined as the probability that handover occurs at position x as a user moves between different cellular cells while traveling at high velocity [41].

2.5.2 Handover Failure Probability

Handover failure occurs when the handover trigger condition is met but the received power from the target BS is lower than a threshold β which is the minimum RSS to maintain a wireless communication link. So the handover failure probability at position x can be denoted as

$$\begin{aligned}
&P_{fail}(x) \\
&= \Pr\left\{R_t(x) < \beta \mid R_t(x) - R_s(x) > H\right\} \\
&= \frac{\Pr\left\{R_t(x) < \beta, R_t(x) - R_s(x) > H\right\}}{\Pr\{R_t(x) - R_s(x) > H\}} \\
&= \frac{\Pr\left\{R_s(x) < r - H \mid R_t(x) = r, R_t(x) < \beta\right\}}{P_{trigger}(x, H)} \\
&= \frac{1}{P_{trigger}(x, H)} \int_{-\infty}^{\beta} F^{(\mathrm{s})}(x, r - H) f^{(\mathrm{t})}(x, r) dr
\end{aligned} \tag{2.10}$$

where F is the Cumulative Distribution Functions (CDFs) and f is the Probability Density Functions (PDFs) [42, 43]. The handover trigger condition in equation (2.10) is based on A3 event. The system's performance decreases as the probability of HO failure increases. HO failures denote unsuccessful handover, which can lead to signal loss and interruptions in an active data session.

2.5.3 Ping-pong Handover

Handovers are adversely impacted by both HO failures and ping-pong handovers. Ping-pong handover occurs when UEs swiftly handover between two BSs within a short period of time which means that the UE is handed over from one BS to another but is quickly handed back to the original BS, generating unnecessary signaling overhead at BSs, degrading the QoS for other UEs and causing signal interruption [44]. The primary reasons contributing to ping-pong handovers include UEs' high mobility between BSs and high variations of signal quality between two different BSs that provide connectivity to the UEs. Therefore, in order to minimize the occurrence of ping-pong HOs, some research has been done to mitigate the negative effects of increased ping-pong HOs [45, 46]. In [45], the authors propose a self-organization algorithm to automatically adjust handover control parameters such as TTT and H in 5G networks in order to improve the HO performance by decreasing the ping-pong effects. A precoding based handover scheme is proposed in [46] to reduce the ping-pong HO trigger rate and improve the HSR system stability. More specifically, the precoding vectors of serving and target BS are jointly designed to maximize the difference in received signal power between the two BSs during the HO preparation stage. Nonetheless, there exists a trade-off between HO failure and ping-pong HO, where reducing the number of ping-pong HOs may potentially lead to an increase in the probability of HO failure, and vice versa. Hence, accurately assessing the occurrence of ping-pong HO in base stations within the network is essential to optimize system performance.

2.5.4 Number of Handovers

The total number of handovers is how many times the handover occurs during the journey. A higher value in this metric significantly degrades system performance and connectivity quality as the UE moves at high speed. For instance, frequent handovers can lead to signal interruptions in data transfers for train passengers which results in a poorer overall user experience. In addition, a large number of HOs generate additional signaling overheads in the wireless communication networks that can contribute to network congestion, affecting the performance of all connected devices.

2.5.5 Handover Delay

Delay is considered as one of the most significant QoS parameters, playing a vital role in enabling users to access real-time services within wireless cellular networks. This metric represents the time needed for transmitting packets from the source BS to the target BS. Higher delay time can impair QoS for these applications which are sensitive to real-time delay. If the delay time in a certain BS surpasses a predefined threshold value, then this particular BS will be avoided. Effective utilization of transmission delay D_i can effectively reduce the handover failure rate while simultaneously enhancing system throughput. The calculation of transmission delay D_i from i-th BS can be described as [47]

$$D_i = \frac{L}{R_i}, \tag{2.11}$$

where L represents the length of the data packet and a packet is the smallest data unit containing a specific number of data bits. R_i is the achievable bit rate from i-th BS. In addition, bit rate is the amount of bits that are transmitted per unit of time. Typically, it is measured in bits per second (bps). Higher bit rate can reduce the delay time to transmit a data packet.

Based on the literature, handover latency is defined as the time required for the entire handover process including HO preparation, execution and completion steps. Moreover, it is an important decisive performance metric when it comes to the provided QoS and user experience in wireless communication networks including cellular networks like

4G LTE and 5G. It measures the time it takes for a mobile device to handover from one BS to another while maintaining an ongoing communication link. The more detailed explanation of handover delay time can be found in the technical report [48]. The goal of network design and optimization is to minimize handover delay time to ensure a desired user experience. Since long handover delay time can result in dropped calls, interrupted data sessions and a poor user experience, particularly in high mobility systems such as HSR. Some research has been done to reduce delay time and improve the QoS for mobile users such as a predictive handover scheme to ensure smoother transitions between various BSs [49].

2.5.6 User Mean Throughput

Throughput is defined as the actual amount of data that can be moved from one place to another in unit period of time. In terms of user mean throughput, it refers to the average data transfer rate experienced by passengers or users. This metric is typically measured in Mbps and reflects the speed and quality of data connectivity available to users in the wireless networks. The throughput can be expressed as:

$$Throughput = B * \log_2(1 + SINR)(\text{ bits/s}) \tag{2.12}$$

where B is the bandwidth of the channel. A higher user mean throughput indicates better system performance and more reliable internet access for users. However, the high HO failure rate and frequent HOs can reduce the user mean throughput in the wireless network, especially when the speed of users increases.

2.6 Summary

This chapter comprehensively introduces the handover technology in wireless communications. First, the handover process is defined and then various handover trigger events from A1 to A6 are discussed. Second, the commonly used handover criteria in the wireless network are summarized in detail. Then, different handover types and different handover performance metrics are introduced. In summary, the content of this chapter gives a comprehensive introduction to the handover technique in wireless communications which lays a theoretical foundation for subsequent research on handover.

Chapter 3

Handover in High-Speed Railway Wireless Communication Systems

3.1 Introduction

This chapter introduces the HSR wireless communication systems and handover challenges in HSR and then gives a comprehensive literature review of the state of the art in the HSR scenario. The rest of this chapter is organized as follows. Section 3.2 gives an overall introduction to the HSR wireless communication system. In section 3.3, the existing handover challenges in the HSR are summarized. After, in section 3.4, the handover in the macro-only network of HSR is given. Then, section 3.5 provides a comprehensive literature review on the handover in the HSR network. At last, section 3.6 summarises this chapter.

3.2 HSR Wireless Communication System

The current railway system utilizes GSM-R for communications. However, LTE-R offers a more extensive range of services compared to GSM-R in terms of capabilities. These services include real-time surveillance and multimedia dispatching for trains. Nevertheless, there are still some limitations associated with LTE-R, such as providing massive internet access for passengers and supporting autonomous driving. As a potential future railway communication system, 5G-R is expected to offer superior performance compared to

GSM-R and LTE-R. It can provide higher data rate railway services that are not supported by LTE-R. The evolution of mobile communication systems from GSM-R to LTE-R and potentially to 5G-R is driven by commercial and marketing requirements, as well as the progressive maturity of 5G technologies. Thus, mobile wireless communication system plays a crucial role in establishing a vital connection between the ground and trains, making them one of the key technologies essential for the successful operation of HSR. They will be introduced in detail in the following sections respectively.

3.2.1 GSM-R

GSM-R was developed as a standardized digital communications protocol based on GSM technology. Its primary objective was to replace incompatible conventional analog systems along railway tracks with a cost-efficient solution. GSM-R implementation requires the installation of base stations along the railway track, accompanied by the deployment of directional antennas to ensure adequate coverage [50]. Prior to the introduction of GSM-R, there is no established standard for implementing both data and voice in a digital channel. The rapid adoption of GSM-R by the railway industry is to fill this gap. GSM-R operates at a data transmission rate of 9.6 kbit/s in a specific frequency band with duplex frequency separation [50]. The uplink frequency band of 876-880MHz and the downlink frequency band of 921-925MHz are used, supporting train velocities of up to 500km/h [5]. Moreover, GSM-R system offers various services that can be developed to meet railway communication requirements. Some of the typical services provided by GSM-R include voice group calls/broadcasts, functional addressing presentation, and location-dependent addressing. Figure 3.1 presents an overview of the GSM-R standard.

While GSM-R has successfully achieved its design goals, it does have some limitations that have come to light:

1) **Limited radio capacity**: The 4MHz bandwidth of GSM-R can only accommodate 19 channels with a bandwidth of 0.2MHz per channel. This capacity falls short of the increasing demands of HSR, which runs at significantly higher speed than traditional rail traffic. Expanding the capacity of the GSM-R network is challenging due to the occupied adjacent frequency bands.

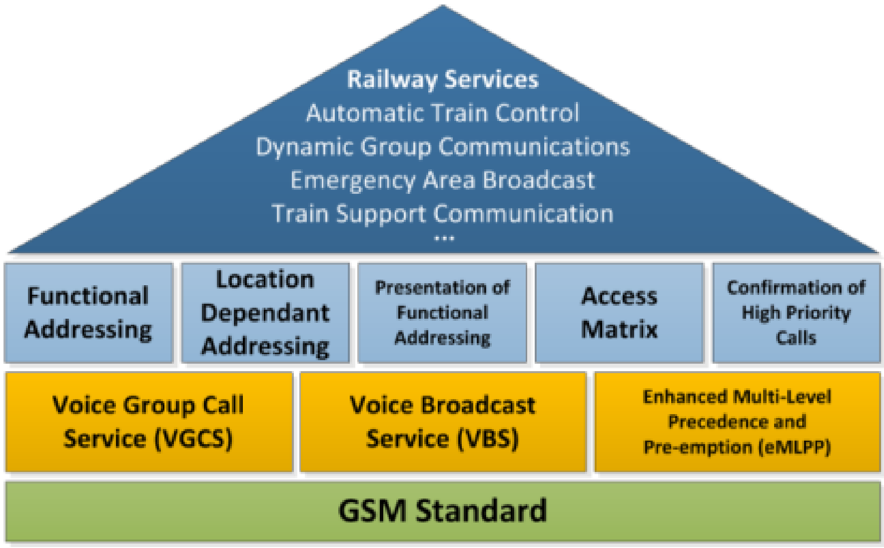

Figure 3.1: The GSM-R standard incorporates additional elements on top of the existing commercial GSM standard [5]

2) **Low link rate**: Each GSM-R connection has a maximum data transmission rate of 9.6 kbits/s, which is inadequate for video surveillance and image services. In addition, it has an approximate delay of 400ms, which is too large and not proper for any real-time interactive services.

3) **Insufficient passenger services**: GSM-R primarily supports train dispatching and control services, passengers can only access the internet through the public networks which leads to a poor user experience. With growing user demands and personal mobile devices, a high-quality network is necessary for good QoS.

4) **Interference**: Overlapping coverage areas between the GSM-R network and other public networks along railway tracks can result in interference, degrading the quality of data and voice communications.

In response to the evolving requirements of high-speed railways, GSM-R is expected to undergo further advancements. The International Union of Railways (UIC) has announced that the utilization of GSM-R will cease in 2025 due to its inability to meet the increasing bandwidth demands of modern internet applications and services, as well as the anticipated high passenger density at train stations.

3.2.2 LTE-R

LTE, also known as Long Term Evolution, is a 4G technology developed by the 3rd Generation Partnership Project (3GPP). It serves as a core network for various vehicular communication technologies. LTE offers several advantages over GSM, including the ability to operate over different bandwidths. Furthermore, the architecture supported by LTE enables a more efficient network, resulting in reduced data transmission delay. LTE-R, the railway-specific extension of LTE technology, has gained significant attention in international research, particularly for the HSR scenario. Some research papers [51] [52] have demonstrated that LTE-R provides improved passenger experience and network performance compared to GSM-R. This implies that LTE-R can offer a wider range of railway-related services, such as real-time video surveillance in critical scenarios like tunnels, bridges, and viaducts. In fact, LTE-R enhances efficiency, QoS, and passenger safety by offering services such as real-time monitoring, train multimedia dispatching, railway emergency communications, railway Internet of Things (IoT), and information transmission for control systems. Additionally, LTE-R can be utilized for data transmission in future automatic driving systems.

In order to meet the demanding QoS requirement, it is essential to explore the key techniques employed in LTE-R. These techniques include enhancing channel estimation, Doppler shift estimation, fast synchronization, wireless channel modeling, and employing massive Multiple-Input Multiple-Output (MIMO) techniques with strong correlation properties [21]. The frequency bands utilized by the upcoming generation of railway mobile communication systems are higher compared to those used in the GSM-R system. Consequently, the distance between two base stations becomes significantly shorter, necessitating more frequent handovers and dense deployment.

Some challenging issues of LTE-R implementation are further discussed as follows:

1) **Specific HSR scenario**: In paper [53], a channel model for high-speed railway is presented which considers the LTE standard. The model exclusively considers a hilly scenario and employs a Rayleigh fading channel model incorporating log-normal shadowing and path loss. Nevertheless, authors ignore the impact of resource alloca-

tion and assume that radio frequency resources are enough for all user equipments. In addition, paper [54] provides a comprehensive survey of channel fading behavior in various scenarios, and aims to help readers have a general understanding of channel modeling in different high-speed railway scenarios. A key feature of the channel is its non-stationary nature, thus the channel models must be dynamic, encompassing various scenarios such as cuttings, viaducts, tunnels, and bridges, to ensure a comprehensive representation of the real-world environment. Furthermore, the future channel model should possess self-learning ability by leveraging artificial intelligence (AI). Therefore, one of the biggest issues with current high-speed railway channel models is weak generality. While these models are only proper to be used in stable scenarios, their system performance deteriorates significantly when applied in more general environments.

2) **High velocity**: Typically, high-speed trains operate at a velocity of 350 km/h, while LTE-R is designed to provide support for speed up to 500 km/h. However, increased speed gives rise to various challenges. For instance, it will lead to a non-stationary channel model that degrades the performance of bit error rate (BER). Another challenge is the Doppler shift which is defined as the change in the wavelength or frequency of the waves with respect to the observer who is in motion relative to the wave source. Given that high-speed trains predominantly traverse established routes at predictable speeds, it becomes feasible to effectively trace and compensate the effects of Doppler shift. This is achievable through the utilization of real-time recorded data pertaining to the train's speed and position [51]. Some research [55] [56] has been done to reduce the impact of Doppler shift in the LTE-R system.

3) **Linear coverage**: In HSR, a strategy of linear coverage employing directional antennas positioned along the rail track is implemented. In this setup, directional BS antennas orientate their main lobe along the rail track so that it is power efficient. However, the influence of shadow fading results in a received signal power below a specific threshold in some positions within the coverage region. Therefore, determining the relationship between boundary coverage and the overall percentage of coverage area is very valuable for network deployment and link budget.

4) **Effect of train material**: Another major challenge is penetration loss in the LTE-R. This is mainly because most of trains are metal materials with large multi-layer glasses, which will lead to the loss of signals and reduction of signal-to-noise ratio (SNR) when going through the train body. For this limitation, a train relay station (TRS) can be utilized to improve the degraded signal strength.

LTE-R has a strong ability to offer competitive performance for the evolution of HSR. Based on the current evaluations of LTE-R, it may not meet the demanding requirements of high-speed railways in the future, which include increased system capacity, higher mobility and cell coverage at higher frequency bands. Therefore, more research and investigations about LTE-R are needed.

3.2.3 5G-R

5G-R is a promising solution for future high-speed railways, and this part presents and analyses the characteristics and key challenges of 5G-R.

In the current era, the prevailing GSM-R, functioning as a narrowband communication system with limited bandwidth and data transmission capabilities, falls short in catering to the demands of advanced railway services such as train multimedia dispatching communication, and intelligent operational and maintenance services. During early 2010s, LTE-R was contemplated as an evolutionary communication system for railways [51]. However, it was recognized that the LTE system was not able to satisfy the demands of highly reliable communication and extensive connectivity for future railway systems. Furthermore, it cannot provide the necessary support for the forthcoming intelligent services and sophisticated applications in the realm of railways such as predictive maintenance systems and autonomous train operations. In order to meet the demanding requirements for higher data rates and massive connectivity in the railways, there has been significant interest in the deployment of 5G-R technology nowadays. The functionalities offered by 5G-R encompass a wide range of services, such as maintaining seamless network coverage along railway tracks, providing coverage to high-traffic areas, monitoring critical infrastructure, and supporting high-speed intelligent applications. The current period is characterized by a continuous revolution in smart technologies, propelled by the swift development of artificial intelligence, cloud computing and 5G communication. Simultaneously, the railway industry is poised to transition from the information age to the era of intelligence. A survey [57] analyzes the distinctions between 5G-R, conventional cellular networks, and GSM-R networks based on the specific requirements of 5G-R and the unique physical environment and propagation traits. Their proposed standardized scenario categorization not only enriches the investigation into 5G-R but also fosters the development and implementation of existing advanced technologies in the HSR. The authors in [50] present a comprehensive survey of the current mobile communication systems utilized in railways and assess their future development trends and technical challenges. Obviously, the performance of railway mobile communication systems can be increased substantially by 5G-R [58].

Figure 3.2 presents a fully connected 5G-R system architecture and the corresponding services, where robust connectivity is established, encompassing train-to-ground link, train-to-train connection, and IoT, spanning all railway-associated devices.

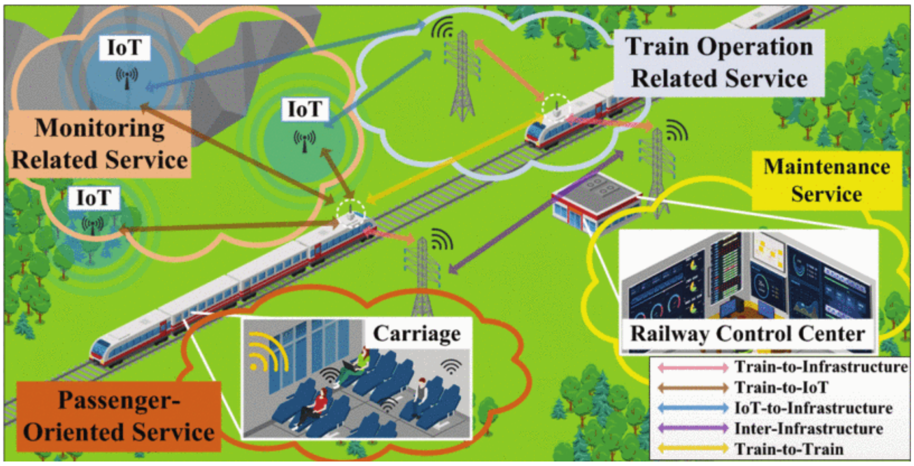

Figure 3.2: Fully connected 5G-R system architecture and services [6]

Nevertheless, 5G-R is regarded as a new research direction and it will encounter a range of challenges which can be summarized as follows:

1) **Limited network coverage**: Currently, the biggest problem for 5G-R is providing good network coverage along the railway road. Utilizing a higher frequency band has the potential to provide a high data rate, but serious path loss also limits network coverage.

2) **Dense small cell deployment and interference**: 5G-R requires the deployment of a dense network of small cells to ensure coverage along the rail route. Managing the deployment of small cells cost-effectively is a challenge. Furthermore, the densification of small cells can lead to more interference.

3) **Cost of deployment**: The deployment of 5G-R infrastructure, including small cells and advanced communication equipment, can be expensive. Balancing the costs with the benefits and ensuring a viable business model is a challenge for railway operators.

4) **Coexistence with other networks**: 5G-R necessitates harmonious coexistence with the general public mobile communication network. Therefore, how to avoid serious adjacent channel interference is a challenging issue in 5G-R. Additionally, HSR wireless communications need to support multiple access technologies in the future.

Besides, 5G technologies can be a potential solution to address the design problems on high throughput and high reliability for 5G HSR communications including rapid channel estimation, cell-free massive MIMO, millimeter-wave (mmWave) utilization, effective beamforming, and improved handover strategies [59]. Firstly, the demand for increased data rates in the context of 5G encourages more investigation of mmWave frequencies spanning from 30 to 300GHz [60]. Then, another significant research field for enhancing data rate is the implementation of massive MIMO, which has had a notable impact on initial 5G trials and has gained even greater prominence as 5G advances. While massive MIMO has been applied to 5G networks operating in traditional sub-6-GHz bands, it still has several research challenges [61]. Furthermore, the optimization of handover has attracted increasing attention in 5G and Beyond 5G systems. Paper [62] offers a comprehensive investigation of the challenges and remedies for HO optimization in the Beyond 5G system. As discussed above, 5G-R exhibits competitive performance which can support a wide range of railway services. Table 3.1 summarizes the major parameters used in GSM-R, LTE-R and 5G-R systems:

Table 3.1: System parameters of GSM-R, LTE-R and 5G-R [6]

Parameters	GSM-R	LTE-R	5G-R
Frequency	Uplink: 876-880MHz; Downlink: 921-925MHz	450MHz, 800MHz, 1.4GHz, 1.8GHz, 2.3GHz, 2.6GHz	900MHz, 1.9GHz, 2.1GHz, 5.9GHz
Bandwidth	4MHz	1.4-20MHz	10-20MHz, possible >100MHz at higher frequency bands
Peak data rate	downlink: 172Kbps uplink: 172Kbps	downlink: 50Mbps uplink: 10Mbps	downlink: 200Mbps uplink: 50Mbps
Peak spectral efficiency	0.33bps/Hz	2.55bps/Hz	downlink: 15bps/Hz uplink: 8bps/Hz
Delay	500ms	200-500ms	50-500ms
Mobility	Max. 500km/h	Max. 500km/h	Max. 500km/h
MIMO	No	2×2	4×4, 8×8, 16×16, 32×32
Handover type	Hard	Soft	Soft
Modulation	GMSK	QPSK/16-QAM	QPSK/16-QAM/64-QAM/256-QAM/OFDM
Cell range	8km	4-12km	1-6km
4k video	No	No	possible support with increased bandwidth

3.3 Handover Challenges in the High-Speed Railway

HSR expands globally as a result of increasing investment from different countries. Moreover, continuous efforts have been made to increase the operational speeds of high-speed trains, reducing travel times between major cities and making rail transportation more competitive than other modes of transportation. In addition, ongoing advancements in rail technologies include improvements in train design, energy efficiency and safety features, contributing to the overall development and efficiency of HSR systems. However, the demand for better communication in HSR is driven by the development of the HSR system. For example, passengers expect uninterrupted and seamless connectivity during their journeys. This requires effective handover mechanisms to ensure a smooth transition between different network cells as the train moves rapidly. Furthermore, some advanced technologies such as mmWave technology, massive MIMO and the addition of small cells are essential to meet the increasing demand for bandwidth to support data-intensive applications. Thus, as HSR networks continue to evolve, addressing these communication demands becomes a critical aspect of enhancing the overall quality of service.

The existing handover problems in the HSR wireless communication network are summarized as follows:

1) **Frequent handover**: The high speed of the train can cause frequent handover when the cell size of BS is small which means the train can pass through multiple BSs within a short time period, which can increase the probability of handover failures.

Moreover, when the minimum time for processing handover is larger than the time interval for the train to pass through an overlapping region between two neighboring BSs, then call drops will occur since the handover process is not completed and it can also cause packet loss. Therefore, a fast handover strategy needs to be proposed.

2) **Doppler shift**: The change in signal frequency at the receiver, resulting from the rapid movement of users, is known as the Doppler frequency shift. As shown in Fig. 3.3,

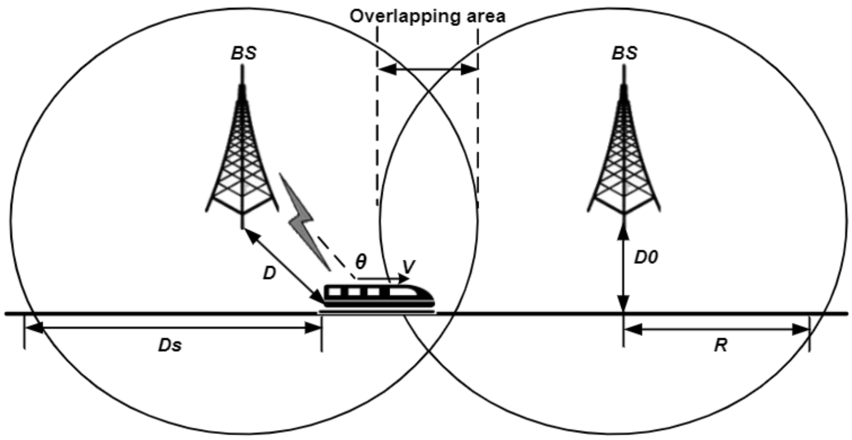

Figure 3.3: BS coverage model for HSR cellular communications

when the train moves along the rail track, Doppler shift f_ρ can be calculated by

$$f_\rho = f_d \times \cos\theta \tag{3.1}$$

where $f_d = \frac{v}{c} \times f$ is the maximum Doppler frequency, θ is the angle between the train's forward direction and the line of sight from BS to the train. As illustrated in Fig. 3.3, we can have $\cos\theta = \frac{R-D_s}{D}, 0 \le D_s \le 2R$. It can be seen that when the BS is placed far from the rail track, i.e., $D_0 \gg R$, f_ρ is relatively low since θ will be approximately $90°$. However, it will cause large path loss. Therefore, there exists a tradeoff between path loss and Doppler shift when optimizing BS deployment.

Additionally, to enable the practical application of wireless communications in the HSR system, we must address severe Doppler shift and rapid Doppler transition issues. On one hand, severe Doppler shift can lead to some problems such as the difficulty of synchronization and bit error rate. However, it's important to note that in most BS coverage areas, while the Doppler shift is large, its variation is relatively small. This allows for accurate estimation and easy compensation of the Doppler shift based on precise train speed and location information [55]. On the other hand, rapid Doppler transitions, especially in the cell center, significantly impact the accuracy of channel estimation and coding which can lead to irrecoverable errors in channel estimation and severe Inter-Cell Interference (ICI) [63] [64].

In the HSR system, as the train travels at high speed along the railway track, the relative motion between the users in the train and the BSs along the track will cause the signal frequency to shift and produce the Doppler effect, which causes serious problems in HSR communication systems. Doppler frequency shift will destroy the orthogonality between subcarriers in the Orthogonal Frequency Division Multiplexing (OFDM) system which can lead to ICI effect that can degrade the received power and also increase the system bit error rate. As a result, the attenuation of the received signal due to ICI effect can increase the probability of handover failure. Additionally, when the train passes the BS along the railway, the signal transmission direction is perpendicular to the train's running direction. At this time, the Doppler frequency shift is zero. But, when the train running direction is the same or opposite to the signal transmission direction, the Doppler frequency shift is maximum.

3) **Penetration loss**: Most of the high-speed trains are made of metal bodies with single or multi-layer glass windows which can cause penetration loss when signals go through the train carriages. In order to solve this problem, the two-hop model [65] can be used. In this model, wireless signals are fed to the antenna on the top of the train and then signals reach the UEs.

4) **High handover failure probability**: Handover failures occur when the system fails to establish a connection with the target BS during the handover process. Due to the high velocity of HST, the remote BS provides radio signal with poor strength leading to higher handover failure probability and communication disruptions. Especially, in the boundary region of two neighboring BSs.

5) **Group handover**: In the conventional mobile communication system, a handover request is generally initiated by a single user. Nevertheless, in the HSR system, there

will be a large number of users initiating handover requests simultaneously, due to the relatively concentrated passengers on the trains. This will trigger a "signaling storm", which will undoubtedly bring huge signaling overheads, and seriously impact system performance. For instance, even if only 30% of the 1000 UEs in the train are active, about 300 UEs will initiate handover requests at the same time.

6) **Handover delay**: Handovers between various BSs can introduce delays in data transmission. This delay becomes more critical in high-speed rail systems where trains are constantly moving with high velocity. Delays can lead to communication disruptions and affect passenger services.

7) **Load balancing**: High-speed railway networks must efficiently distribute the load among different BSs to ensure optimal network performance. Load balancing during handover is crucial to prevent network congestion and provide consistent service quality.

   Therefore, addressing these handover problems in HSR wireless communication networks requires a combination of technological advancements, network optimization and HO algorithm optimization. Efforts should focus on ensuring reliable and seamless high-speed connectivity to enhance safety, efficiency and passenger experience on high-speed trains.

3.4 Handover in Macro BS Only Network of HSR

A robust handover mechanism within macro BS only network is imperative, this network comprises a series of linearly deployed macro BSs which are providing wide-area coverage. These macro BSs are instrumental in delivering high-quality wireless communication services to passengers on board, catering to their diverse connectivity needs, such as voice calls, data streaming, and internet access.

The macro BS only deployment scenario focuses on continuous coverage along track in high speed trains. The key characteristics of this scenario are consistent passenger user experience and critical train communication reliability with very high mobility [7]. A dedicated linear deployment along the railway line is considered as in Fig. 3.4, and passenger

UEs are located in train carriages.

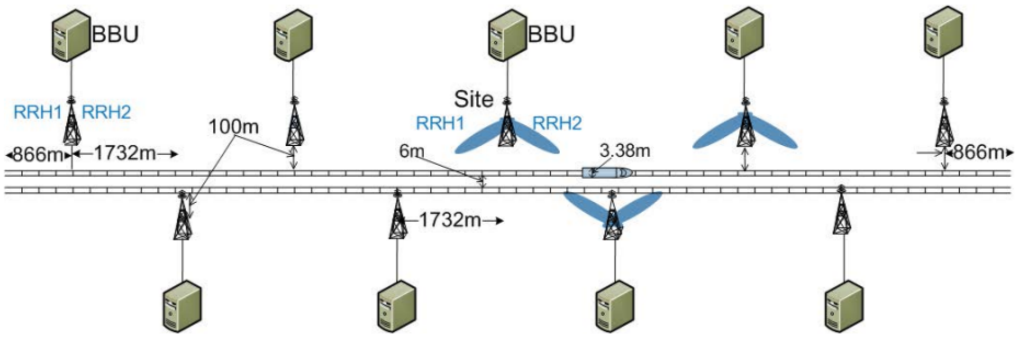

Figure 3.4: Macro-BS-only HSR network architecture [7]

where the baseband unit (BBU), known for its modular design, small size, high integration, low power consumption, and easy deployment, plays a vital role in dealing with baseband signal processing and managing the entire base station, including operational maintenance. Additionally, the remote radio head (RRH), also referred to as the remote radio unit (RRU), serves as a remote radio transceiver that establishes a connection with the operator's radio control panel through either electrical or wireless interfaces [66]. Its functions encompass several critical tasks such as converting radio frequency (RF) signals into data signals and vice versa, filtering and amplifying RF signals. Furthermore, the RRH is the distributed and integrated frequency unit that connects to the operator's network, facilitating communication with UEs like cell phones and mobile devices. Consequently, the macro BS is the integration of various radio unit like BBU and RRH. Notably, these radio units are currently installed in the tower beneath the antenna.

3.5 Literature Review

This section summarizes the technical developments of handover in the HSR wireless communication systems. It is divided into two main parts: handover schemes (focus on the handover problems based on different architectures) and handover algorithms (focus on the problem of handover decision-making during a handover process).

3.5.1 Handover Schemes

Handover schemes refer to systematically organized plans including network architecture deployment and handover procedure design based on different features of network configuration. These schemes play a crucial role in managing the handover process within mobile systems. The design of the handover procedure is vital, considering the nature of the network architecture. For instance, different network architectures such as distributed LTE and centralized GSM systems lead to various handover schemes. Thus, this section explores and classifies various network access and handover schemes, spanning from traditional to innovative, based on their various features as shown in Fig. 3.5.

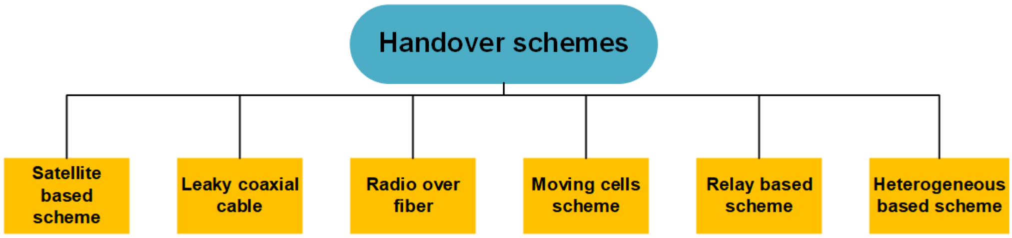

Figure 3.5: Different handover schemes [8]

1. Satellite to ground integrated network

Satellite communications offer extensive coverage and reduce the need for the installment of ground communication equipment along with tracks. Moreover, satellite communications are less affected by the Doppler frequency shift, ensuring more stable and reliable connections during high-speed rail movement. In addition, the broad coverage facilitated by satellites reduces the frequency of handovers, contributing to more continuous communication quality during the journey. Nevertheless, some special scenarios such as tall city buildings and tunnels can create blind spots in satellite coverage, leading to potential interruptions in communication. Also, satellite communication involves high costs, including satellite renting and user equipment expenses. The most obvious disadvantage of satellite communication is the large handover delay due to the vast distance involved, which makes it unsuitable for real-time applications that require rapid data transfer, such as multimedia streaming. On the other hand, satellite signals can experience high losses in adverse weather conditions which can deteriorate signal quality and reliability during periods of rain, snow, or storms.

Some research works about the handover schemes in the satellite-to-ground integrated network are summarized. Paper [67] presents a comprehensive overview of research on space-air-ground integrated network (SAGIN), including network design, resource allocation to performance analysis and optimization. This survey not only introduces the concept of SAGIN but also its significance in enhancing various services and applications. At last, it identifies technical challenges and suggests future research directions based on SAGIN. Thus, according to this survey, future work needs to consider system integration, resource allocation and mobility management in SAGIN due to its heterogeneity and time-variability. However, the existing surveys have not fully addressed the integration of all three network segments (space, air, ground). The authors in [9] discuss the development of SAGIN and its application in the HSR communication system. It acknowledges the potential of SAGIN to innovate communication network structures and integration technology, particularly in railway mobile communication systems. SAGIN provides a reliable and high-throughput communication service for smart HSRs by integrating space-air-ground network as shown in Fig. 3.6. This survey also points out that AI technologies are proposed

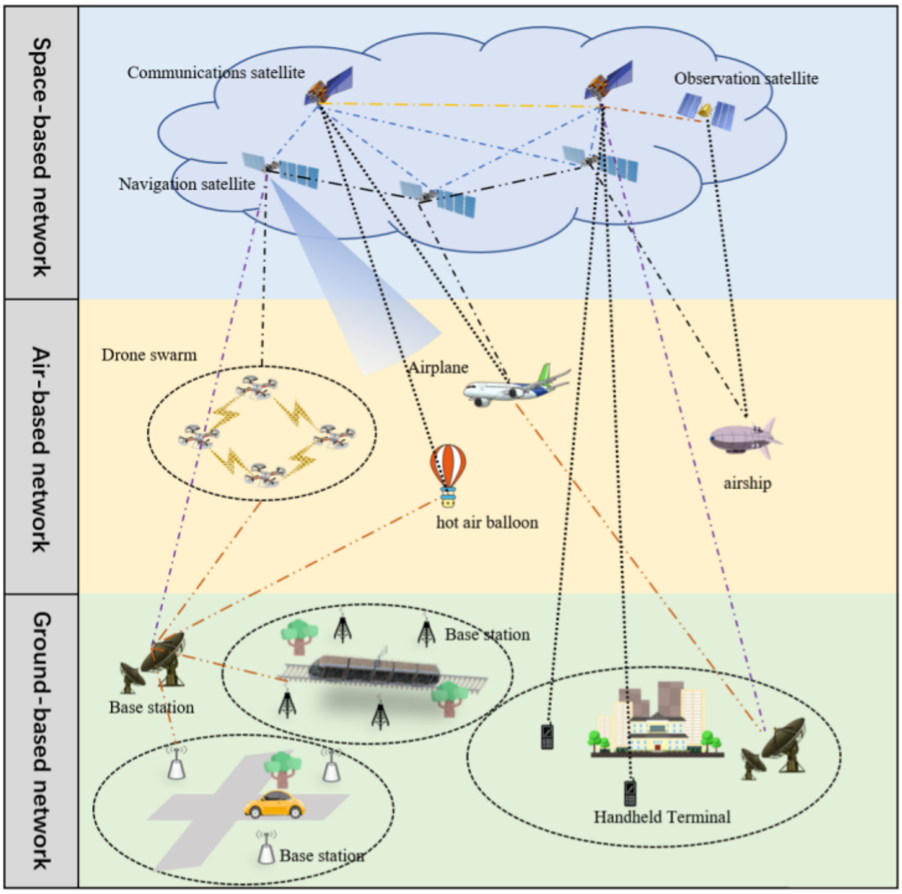

Figure 3.6: The SAGIN diagram [9]

as solutions for optimizing the SAGIN and enhancing efficient resource utilization in communication systems, but this paper implies that their application in SAGIN for HSRs is still a challenging task. Additionally, a lack of suitable time-varying channel models hinders the performance analysis of the SAGIN communication system for HSRs. The authors in [68] propose a fast and secure handover method using pre-authentication and security context transfer to reduce the handover delay in high-mobility SAGIN environments. Their results achieve a 16.92% reduction in handover delay compared to the baseline algorithm.

2. Leaky coaxial cable

The Leaky coaxial cable (LCX) is a type of coaxial cable that is designed to intentionally radiate signals along its length which have small slots or gaps along their outer conductor, allowing them to leak or radiate signals. This scheme is often used in tunnel scenarios in the HSR system. In addition, it is designed to offer average coverage of radio signals while avoiding mutual interference. It reuses several frequency points in the long segment of the HSR which means that the same frequency points are used multiple times to maximize the utilization of the available bandwidth [8]. In [69], the authors introduce an experiment of LCX in the 2.4GHz band for train-ground information transmission. In addition, they measure the transmission performance of LCX and analyze the availability of LCX in tunnels for Communication-Based Train Control (CBTC) train-ground communication. However, this paper does not compare with other frequency bands or technologies that could be used for CBTC train-ground communication. Also, the long-term durability and maintenance requirements of the LCX system are not addressed. On the other hand, because of the significant attenuation of radio signals in LCX, a large number of repeaters must be deployed which leads to a substantial increase in relay coverage expenses. The intricate nature of LCX's transceiver and repeater equipment has been the primary focus of recent research endeavors. In summary, the initial investment cost for establishing its network would be exceedingly high. Therefore, the LCX network coverage solution is not proper for the entire line of the HSR system. Instead, it serves as an alternative way to other wireless access methods specifically in scenarios such as tunnels.

The works in [70] develop a fading model for LCX in urban rail transit tunnels, using field measurements from Beijing Subway Line 7. It captures large-scale and small-scale fading characteristics and demonstrates that large-scale fading with LCX can be modeled linearly, while small-scale fading closely follows a normal distribution. Nonetheless, this model is based on measurements from a specific urban rail line, which may not be generalized to other environments. In [10], the authors address high-precision train positioning, a key aspect of intelligent transportation systems, particularly in tunnel scenario. They consider the LCX as an antenna in tunnels, highlighting its advantages for rail applications in the 5G and 6G eras. The structure of LCX is shown in Fig. 3.7.

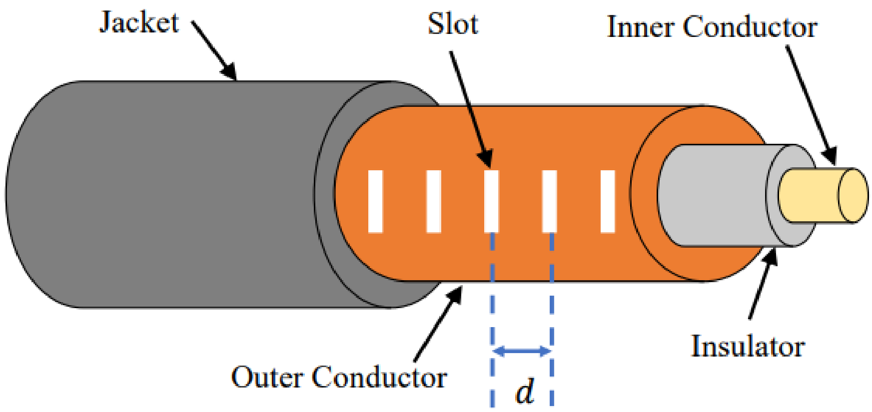

Figure 3.7: The structure of LCX [10]

where slots are cut on the outer conductor of an LCX and the outer conductor is typically made of metal. The slots are deployed periodically with a separation distance denoted as d that allows for the leakage of signal and radiated waves.

3. Radio over fibre

Radio over fiber (RoF) is a technology that combines radio frequency and optical communication systems. One of the primary motivations for implementing RoF in HSR communication systems is the potential for significant bandwidth enhancement. HSR systems demand high data rates and large system capacity for applications ranging from train control signaling to passenger information. Thus, the ability of RoF to support high-frequency signals over optical fibers addresses increased bandwidth requirements, contributing to improved communication performance. RoF technology [11, 71–73] has shown promising potential and is well-suited for HSR broadband wireless communication system.

The integration of DAS with RoF technologies has been implemented in HSR communications. Multiple Remote Antenna Units (RAUs) are positioned along the railway track and connected to the control center through a fiber ring. At the control center, RF signals are converted into optical signals for transmission over the fiber ring to the RAUs. Subsequently, they are reconverted into RF electrical signals, and then these signals are radiated by the antennas from the RAUs as shown in Fig. 3.8.

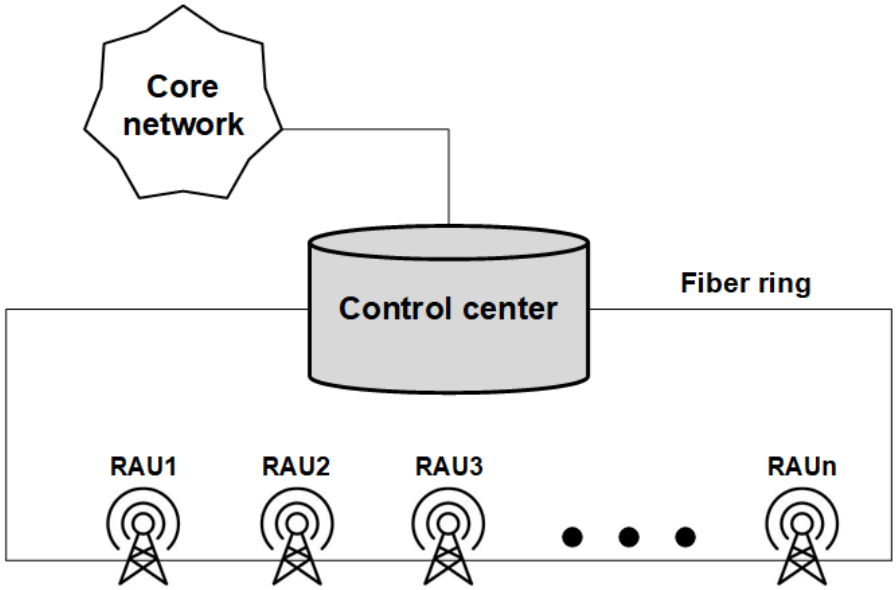

Figure 3.8: Linear cell planning based on the RoF network including control center, RAUs along the rail track, and an optical ring network connecting them all [11]

Furthermore, when the HST runs between RAUs within the same control center, there is no need for handover. Consequently, in comparison to the traditional HO method, the RoF scheme can significantly decrease the frequent handover. Additionally, the handover time can be minimized because of the swifter nature of optical switching compared to electrical switching.

4. Moving cells scheme

The concept of moving cells (MC) was first introduced in highway communications [74], where BSs are physically moving along the tracks. Consequently, the cell formed by a BS is also dynamically moving. Nevertheless, the practical implementation of numerous physically mobile BS is deemed impractical. As a result, it sparked some innovative ideas for realizing the function of moving cells.

Moving cells can be implemented by reconfiguring the optical network feeding the RAUs using current optical switching technology to allow trains to communicate on the same frequency throughout the journey, reducing handover issues [72]. Thus, there is no need for synchronization between the preparation of the new BS and the shift of the radio channel in the connection between the RAU and the train. The research work of moving cells is proposed in [12]. A Cell Array is defined as three cells (A, B, C) aligned along the path of HST. In terms of cell A, it is the current cell for the train connected. Cell B is the adjacent cell which is ahead of cell A along the railway track as shown in Fig. 3.9. As soon as the train is completely in cell B, cell A is no longer in the extended cell configuration. Then, The cell array reconfigures when the train completely leaves cells A and enters cell B.

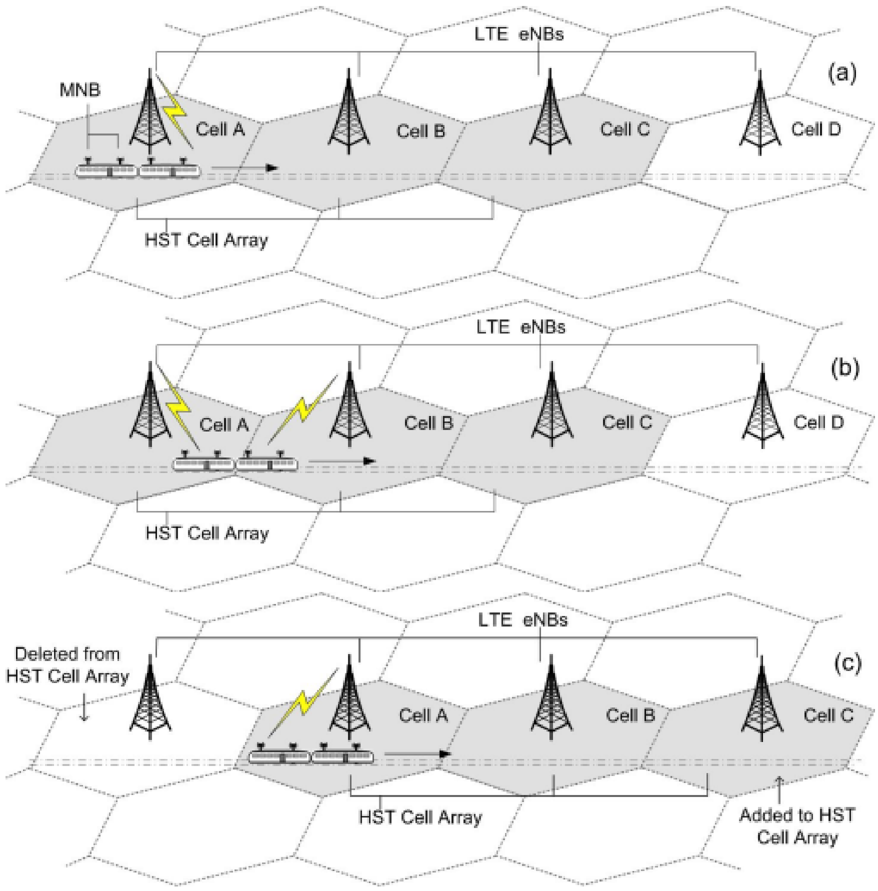

Figure 3.9: Moving cell array along the railway track [12]

During movement, cell D is introduced into the array, A is removed, and the array is renamed to ensure a seamless transition for the train. Utilizing information about the high-speed train's movement direction and speed, the cell array can predict the target LTE cell to achieve seamless handovers. The results of this paper demonstrate that handover latency time can be reduced and data throughput is increased compared to normal LTE access. However, we find that the proposed solution may encounter application challenges in real-world deployments with varying network and traffic conditions.

5. Relay based scheme

Relay station plays a crucial role in the HSR wireless communication system, facilitating the seamless transfer of signals and data along the high-speed railway network. These relay stations, often strategically positioned along the railway route, are instrumental in ensuring continuous and reliable communication. The following aspects highlight the key functionalities and characteristics of relay stations in HSR:

- **Signal coverage enhancement:** Relay stations are strategically located along the railway track to provide additional points of signal transmission and reception, relay stations help to mitigate signal attenuation and ensure a robust communication link between various BSs of the HSR system.
- **Handover management:** Relay stations aid in managing handovers efficiently, ensuring that the handover between different BSs is smooth and reduces signal loss or degradation.
- **Adaptation to HSR:** Relay stations are designed to operate effectively in high-speed scenarios. They can address challenges associated with rapid train movement, including frequent handovers and dynamic signal propagation.
- **Enhanced reliability and redundancy:** The deployment of relay stations adds redundancy to the communication system, ensuring that there are alternative paths for signal transmission in case of failures. However, installing a large number of relay stations will increase the cost and is very expensive. Thus, this is more suitable for some areas with weak signals.

The authors in [75] propose a mobile relay-based fast handover scheme to ensure timely handover execution by introducing two reference points. In addition, they implement pre-preparation and packet bi-casting to reduce communication interruption. A mobile relay node is adopted to mitigate penetration loss and support group mobility. Their simulation results demonstrate that this scheme can enhance performance in terms of handover delay time, communication interruption time, and bi-casting time.

Another handover scheme [76] utilizing dual antennas and a mobile relay station is proposed to enhance signal quality and reliability in train-to-ground communication. The proposed scheme is compared to standard LTE handovers, showing a significant reduction in handover frequency and improved seamless access. Their theoretical analysis and simulation results are provided to validate the effectiveness of the handover scheme. However, there is no comparison with other state-of-the-art handover schemes other than the conventional method. The authors in [77] introduce a network-assisted mobile relay-controlled fast handover method to improve HO performance. The mobile relay makes handover decisions based on advanced resource configuration from the serving BS. Specifically, the mobile relay continuously measures signal strength and compares it against handover conditions. Upon meeting the handover condition, the mobile relay will send a handover indication information to the serving BS. Then, the mobile relay synchronizes to the target BS using reserved resources after detaching from the serving BS. The simulation results indicate a reduction in handover failure and communication interruption probabilities. Note that the scheme's effectiveness in different network configurations is not discussed.

A positioning and relay-assisted handover scheme [13] is proposed that utilizes train position information and relay power control to ensure the signal coverage and strength to meet the minimum reliable communication requirements in the overlapping area. In this scheme (Fig. 3.10),

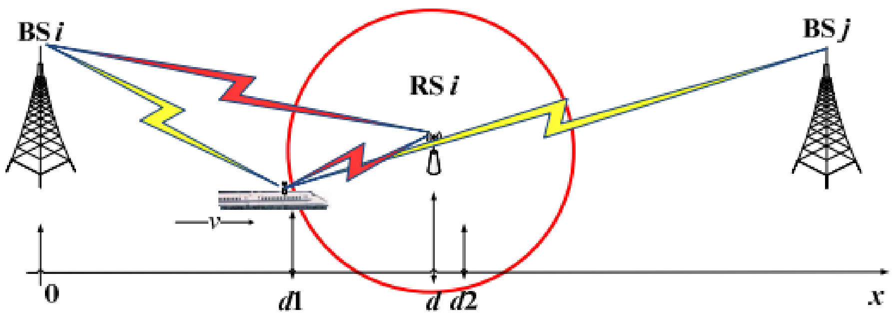

Figure 3.10: Relay based handover scheme [13]

the relay station is located at the middle point of the overlapping area and is only associated with the source BS. Additionally, the train positioning and speed information provided by the train control system is used to determine optimal handover points. For instance, the relay station first acts as a repeater (i.e. the train is at position $d1$). When the handover triggering condition is met, the power control of the relay station is initiated at specific locations (i.e. the train is at position $d2$) to maintain the minimal communication link between the train and the source BS, while the target BS is preparing for the handover. After the handover is completed, the relay station stops the transmission. In fact, the relay station power control leads to the postponement of the handover trigger, reducing premature triggering and extending the overlapping region's coverage. Besides, if more relay nodes are added to the communication system, it will result in more complicated interference analysis. Nevertheless, the Doppler Effect is not considered in this paper, which could impact the performance of relay based handover.

6. Heterogeneous based handover scheme

The deployment of wireless communication technologies in HSR environments has witnessed significant advancements, with the integration of heterogeneous ultra-dense networks playing a vital role. Seamless handover between different network technologies is crucial for maintaining continuous connectivity as trains traverse through diverse geographical areas. This section reviews the state-of-the-art research, challenges, and solutions about handover technology in heterogeneous HSR wireless communication.

Papers [78] [79] provide an overview of handover management in ultra-dense mobile heterogeneous networks. In [78], this paper reviews handover management in ultra-dense mobile HetNets, focusing on the challenges posed by 5G and beyond technologies. It discusses the introduction of small cells and mmWaves to enhance coverage and system throughput. But, small cells in ultra-dense networks cause frequent handovers and mmWave technology introduces high path loss and coverage issues. Additionally, various handover management algorithms and mobility management techniques for Dual Connectivity (DC) are investigated. The proposed solutions in the literature include self-organized networks (SON) and machine-learning approaches to minimize handover failures. Future research directions are identified, emphasizing the significance of machine learning in improving the handover process. The paper [79] gives a discussion about the evolution of mobile communication technologies and the necessity for high data speeds and continuous connectivity. It highlights the benefits of HetNets, such as increased network capacity and improved signal quality, while also acknowledging the complexity and challenges they face. Also, it identifies major challenges in handover management and categorizes current algorithms for performance evaluation. For instance, interference from adjacent cells can be more serious due to the ultra-dense small cells. Finally, the survey identifies open issues and future research directions for handovers in 5G and beyond. Thus, more research work is needed to develop advanced handover approaches to manage the increasing complexity in HetNets and enhance handover decision techniques with ML.

The works in [14] use a heterogeneous ultra-dense network (Fig. 3.11) with the gray model (GM) prediction based handover scheme to explore better mobility management in HSR. Specifically, the prediction algorithm uses historical data to forecast future values such as received power. Moreover, it also introduces a combination factor in the handover decision to select target BS.

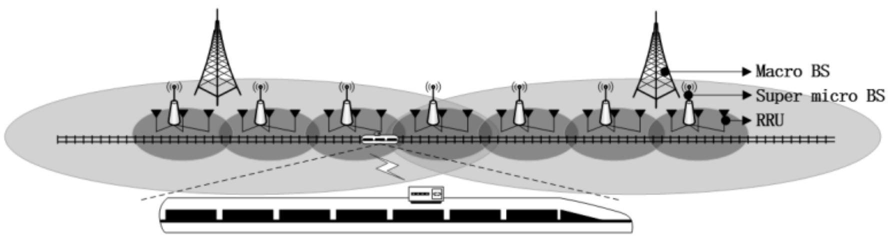

Figure 3.11: High-speed railway heterogeneous ultra-dense network [14]

In addition, this paper highlights the need for addressing handover lag and frequent handover issues to ensure QoS and service continuity in heterogeneous UDN of HSR. Their simulation results demonstrate that the proposed handover scheme significantly reduces handover failure probability compared to the traditional scheme and the gray model with multiple outputs helps prevent frequent ping-pong handovers and handover lag. Nonetheless, real-field measurements are not used to verify the proposed algorithm, indicating a need for practical validation.

A dual-link soft handover scheme to reduce handover outage probability and improve handover success probability is proposed in [80]. This work separates control and user data to improve communication reliability and capacity for high-speed trains. A TRS is also used to maintain a stable connection during the train's movement and two antennas on the train help in switching between BSs smoothly. Also, the bi-casting technique is adopted to decrease the communication interruption time. In addition, this paper identifies the need for improved network capacity and reliable frequency band usage in future HSR communication systems. Furthermore, it compares the proposed scheme with the existing LTE-R network and a gray prediction handover scheme, showing improved HO trigger probability, success probability, and reduced handover outage probability. The authors in paper [81] focus on a seamless handover scheme for 5G C/U plane split networks in HSR communication to improve QoS and reduce handover failure probability. Their work uses an optimized threshold that depends on the signal strength and hysteresis level. The front antenna on the train can execute handover to a new BS while the rear antenna maintains the connection with the current one. This method lowers the chance of losing service as the train moves at high speeds. Their simulation results support the effectiveness of the proposed handover procedure in reducing handover failure probability. Paper [82] proposes a dual-connectivity scheme for mobile UEs where they are connected simultaneously with two mmWave BSs to overcome frequent handover problems. This method allows a mobile UE to choose its serving link between the two mmWave connections based on signal strength measurements and then the corresponding BSs may forward duplicate packets to the UE. Nevertheless, this work might result in increased overhead due to maintaining two connections with a mobile UE.

3.5.2 Handover Algorithms

The handover algorithm is a systematic process designed to address the problem of handover decision-making during a handover process. This process, primarily executed by a computer processor, is assessed based on various metrics, including its complexity, accuracy, and reliability. The effectiveness of such an algorithm is crucial for the overall efficiency of a program and even the entire system.

A key concern in HSR is the handover failure which can lead to termination of an ongoing call. Handover failures occur due to many causes such as high mobility and signal degradation. These can have a direct impact on the service disruption and user experience leading to dropped calls and dissatisfaction among passengers. In order to satisfy the user QoS and to provide seamless coverage, handover procedures should be further discussed. Current research in the HSR networks primarily focuses on two aspects: the criteria for making handover decisions and prediction based handover algorithms. In this section, these algorithms are summarised and discussed, as illustrated in Fig. 3.12.

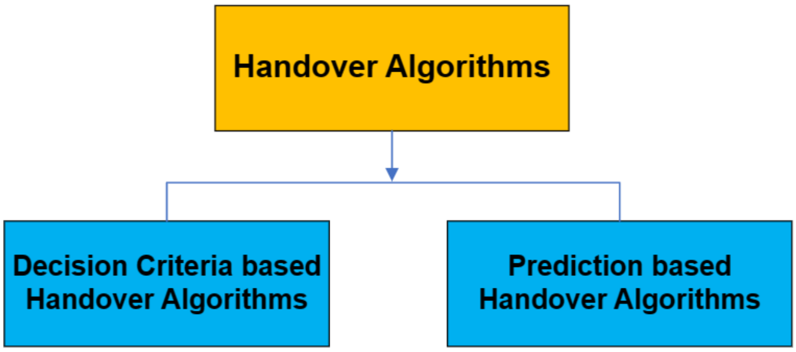

Figure 3.12: Classification of handover algorithms

1. Decision criteria based handover algorithms

Handover decision criteria are various based on different application scenarios. Handover decision making mainly relies on control parameters such as H, TTT and threshold level. However, small adjustments to these parameters are not able to improve the performance of the whole system. For decision making algorithms, the traditional A3 event-based HO algorithm is very prevalent. Event A3 is defined as a reporting triggering event that is fired when there exists a neighboring cell whose measured RSS is better than the measured RSS of the serving cell by a certain offset. However, it is mainly designed for low-speed scenarios such as speeds less than 120 km/h. Also, TTT and handover hysteresis parameters are fixed if they are chosen properly. The authors in [53] conclude that the traditional A3 event-based HO algorithm is not suitable for the HSR environment based on their simulation results. For instance, due to rapid changes in channel conditions in HSR, the A3 event-based algorithm may not be able to adapt quickly enough to these changes, leading to inefficient handover decisions or unnecessary handovers. In addition, A3 event-based algorithm typically rely on current measurements to trigger handover and may not have sufficient predictive capabilities to anticipate future changes in channel conditions, especially in dynamic environment. Moreover, even though the proposed optimization method can mitigate the influence of radio link failure in the HSR when compared with the conventional A3 event-based HO algorithm, the ping-pong effect cannot be eliminated. In [41], three conclusions are obtained as follows: firstly, handover trigger probability is affected by train velocity, measurement report period, handover decision number and overlap range. Secondly, the handover trigger probability will be more than 99.8% when the overlap range is more than 30% of the radius of a cell. At last, the report period can be calculated according to the train speed and overlap range to meet the requirement of handover trigger probability. Moreover, the report period such as 200ms is very proper to be used in high speed systems with less than 540km/h speed. In [83], the authors adopt two kinds of handover algorithms based on the optimization of handover hysteresis and TTT in different velocity scenarios and different SINR when compared with the conventional A3 event algorithm. The measurement of UE velocity is used to avoid the increase of handover failure when UE velocity increases. Furthermore, simulation results show that optimized handover algorithms have higher handover success rate than traditional A3 event algorithm based on different velocity scenarios. The authors in [84], propose an adaptive handover algorithm based on distributed antenna. In the process of switching, the hysteresis threshold value of handover is dynamically varied depending on different velocities at different positions. Simulation results demonstrate that the handover success rate is improved substantially, and the interruption probability of handover is decreased. However, the proposed method in [84] has increased the average number of handovers. In summary, adjusting the handover control parameters to find the most proper values is very helpful in triggering handover timely for high-speed scenarios.

Besides, previous handover works use single parameter for HO decision in HSR. Various HO criteria can be combined such as RSS, SINR, capacity, train speed, distance, etc. Therefore, the HO algorithms have various types such as RSS based HO algorithm [85–87] and user speed based HO algorithm [88]. Generally, the RSS based algorithm is the simplest system, so it is the least accurate. So far, there have been a few works done on multiple criteria handover decision algorithms [89–92]. In [89], the authors propose a parameter-adaptive handover mechanism suitable for LTE-R communication systems.

It uses Temporal-Difference (TD) learning-based reinforced agents to assess handover performance and network performance under various train speed and handover parameter combinations to enhance handover accuracy and adaptability. Work in [90] adopts a novel handover control scheme named H2, designed to monitor the received power variation level of the serving BS and target BS to prevent late handover and stably execute handover in rapidly changing conditions. Their simulation results demonstrate that the H2 scheme can reduce handover failure probability compared to traditional LTE handover mechanisms, offering a promising approach for maintaining reliable and high-quality communication in the HSR system. In addition, an optimized relay handover algorithm in HSR is proposed in [91]. The authors address the issue of handover decision-making, focusing on improving the handover success rate and reducing outages. They introduce a dynamic adjustment of TTT based on the train speed and other factors aiming to optimize handover performance in high-speed environments. Their simulations demonstrate its effectiveness in reducing handover failures compared to traditional methods. However, the limitation of this proposed scheme is the higher signaling load and less system throughput. In [92], the authors propose a novel approach to improve the handover decision-making in 3GPP LTE systems for HSR environments. This handover algorithm is based on the Grey Model, which considers both RSRP and RSRQ for handover decision making. The aim is to ensure QoS by making more accurate handover decisions, using the Grey Model to predict signal strengths and manage handover processes efficiently. Moreover, the authors in [93] propose an improved handover scheme for LTE-R systems in HSR scenarios. It focuses on optimizing handover decisions by dynamically adjusting the handover hysteresis margins of the RSRP and RSRQ based on the train speed and position. The proposed scheme also uses beamforming technology to improve received signal strength which can enhance handover performance such as handover trigger probability and success probability effectively.

2. Prediction based handover algorithms

According to the space-time features of the train operation, the available information that they may have in the HSR network can be utilized to assist handovers. Prediction techniques help to make handover decisions based on the expected future values of the received power or other network information. Therefore, the handover processing time can be decreased and the handover failure probability will be reduced. Various handover prediction algorithms have been proposed to enhance the handover performance. Here is a summary of some prediction based handover algorithms in high-speed railways:

A work in [94] is proposed to predict the future movement and handover reference point. The authors use GPS information including train speed, location and direction, the system first finishes the channel assignation and activation, and then makes handover decisions based on the measurement reports before the train arrives at the overlapping area between serving BS and target BS. Their simulation results demonstrate that the proposed algorithm can reduce the handover delay. Nevertheless, a large amount of statistical location information is needed to obtain a good prediction of the handover reference point. Moreover, the handover reference point should be updated periodically. The authors in [95] present a Bayesian regression based algorithm for improving handover decisions in HSR scenarios which uses this model to predict the cell boundary crossing time to initiate the handover process earlier than conventional algorithm. Simulation results show that the proposed algorithm performs better than the conventional and H2-based schemes in terms of communication reliability at the cell boundary regardless of the train speed. Additionally, this paper also discusses the computational complexity and the necessity of using Bayesian regression for better handover decisions. In [14], the authors propose a gray model prediction based handover algorithm in the heterogeneous ultra-dense HSR network to reduce the handover failure rate. The results show that the proposed handover algorithm reduces the handover failure probability, and ping-pong effect and improves mobility management for HSR.

In addition, with the development of AI, utilizing ML for handovers in 4G/5G HSR networks holds the promise of enhancing overall wireless network performance and efficiency through making more precise handover decisions. Nonetheless, several key research challenges need to be solved to fully realize the potential of machine learning in this area. These challenges encompass issues related to the availability and quality of datasets, concerns about privacy and security, online and offline learning, frequent handovers, signaling overhead, energy efficiency, and even the necessity of load balancing. Addressing these challenges is paramount to the development of accurate and efficient ML algorithms for handover decisions in the HSR network. Thus, a few works [32, 96–98] have utilized machine learning algorithms to analyze patterns in handover data and predictive models trained on historical handover events can provide insights for more efficient handover decisions in the future. The authors in [32] provide a comprehensive survey of recent handover decision algorithms in mobile HetNets, categorizing them based on decision techniques. The survey concentrates on current and future requirements related to mobility and handover management in mobile communication networks. In addition, this survey discusses self-optimization, soft computing, and machine learning/deep learning approaches for handover and mobility management. Finally, it highlights technical challenges and opportunities related to handovers in mobile communication networks. In [96], the authors propose an intelligent handover algorithm based on the Elman network for HSR scenarios which introduces a neural network prediction system to continuously observe and predict handover decision parameters like RSRP and RSRQ. The proposed method aims to reduce link failure and enhance user experience under the LTE-R system. Simulation results illustrate that the Elman-based handover algorithm performs better than the gray prediction model-based handover algorithm. Nevertheless, the Elman network's performance in this study is not compared against other types of neural networks or machine learning models. In [97] and [98], various machine learning approaches are applied for future mobility prediction. An adaptive optimization decision mechanism for handover parameters in LTE-R communication systems using Generative Adversarial Networks (GAN) theory is presented in [97]. The authors use GAN to learn from discrete samples and train the network to generate continuous handover situations for the selection of optimal handover parameters. At last, feedback from successful handover situations is used to improve the GAN's learning and optimization. The results demonstrate that the GAN based handover parameter prediction algorithm can help in making more accurate handover decisions and enhance the LTE-R system performance. However, the adaptability of the proposed HO algorithm to other types of communication networks beyond LTE-R is not discussed and the robustness of the system against adversarial attacks or network disruptions is not addressed. In [98], a speed model based on the gradient descent algorithm to locate the high-speed train is proposed which helps to identify the location of abnormal remote radio units and BSs. The results demonstrate that the probability of accurately locating the HST can be up to 87.17%. Nevertheless, the accuracy of the speed model is dependent on the amount of user information available including start time, user name, BS HO type, BS HO status, cell number and BS label. In other words, with less data, the model may be less accurate. Thus, expanding user information could reduce errors and improve the model's accuracy.

3.6 Summary

The studies reviewed reveal the complexities associated with seamless communication during the handover process between cells or BSs. To overcome these challenges including high handover failure probability, frequent handovers caused by high mobility, high penetration loss, Doppler shift, signaling overhead and group handovers, etc, advanced system architecture designs, handover schemes choice and handover algorithms designs are needed to achieve better handover performance in HSR wireless communication mobile network. According to the reviewed research works in this chapter, we can conclude that RSS based handover is the most common handover decision metric used in the literature. Other criteria such as speed, position and interference, etc also can have a big impact on handover performance for high-speed railways. Additionally, incorporating HO hysteresis margin and TTT plays a significant role in RSS comparison which could reduce the ping-pong HOs to some extent. Furthermore, the incorporation of machine learning-based approaches is considered as a promising solution for addressing the handover issues in the HSR network.

Based on the literature, a lot of work has been done to address the HO issues in HSR wireless networks, but important work still needs to be done in this area. Handover management is a critical aspect of wireless networks, particularly in 5G-R where there is a high density of small cells and a large number of devices connecting to the network with dynamic requirements like high-speed mobility, low latency, high reliability and high data rates, etc. Similarly, a lot of emerging technologies such as intelligent HSR, millimeter technology and massive MIMO, etc. will introduce new challenges in HSR dense network. Previous research works in the HSR scenario are already focusing on various techniques such as different handover system architecture, and decision criteria based HO to improve handover management and to make it more efficient and flexible. There are still many open areas to be addressed such as the need for considering multiple criteria to enhance the HO decision and better prediction algorithms to reduce the handover challenges such as frequent HOs and handover failure.

To this extent, this thesis focuses on the handover problem in HSR wireless communication mobile networks aiming to address the following aspects:

1) To improve the handover triggering process and handover success probability by considering multiple significant criteria in the HSR system.

2) To reduce the signaling overhead, penetration loss and group handover problem by adopting dual-layer architecture (two-hop model).

3) Designing a handover technique to minimize the number of handovers, handover latency and handover failure probability, in addition to improving the user experience in terms of enhancing the throughput for 5G HSR system. This is done by a reinforcement learning algorithm based handover strategy.

4) Designing an intelligent handover method based on machine learning algorithm that selects the target BS by jointly considering multiple HO criteria that account for reducing the number of handovers, handover failure probability and handover latency, in addition to improving the user experience in terms of enhanced user mean throughput.

Chapter 4

Multi-Criteria Handover Scheme for Distributed Antenna System based High-Speed Railway Wireless Communication

4.1 Introduction

In recent years, the HSR has been developed rapidly and considered as the most flourishing transportation mode. The maximum train moving speed has reached nearly 575 km/h. This brings multiple challenges to HSR wireless communications, such as the handover technology. In wireless communication of HSR, BSs are deployed along the rail track and handover happens in the overlapping areas. The increased handover failure rate and frequent handover caused by high mobility results in serious communication interruption.

The DAS based system architecture [99] [15] provides a dual purpose, aiming to improve both spectrum efficiency and handover performance, as shown in Fig. 4.1.

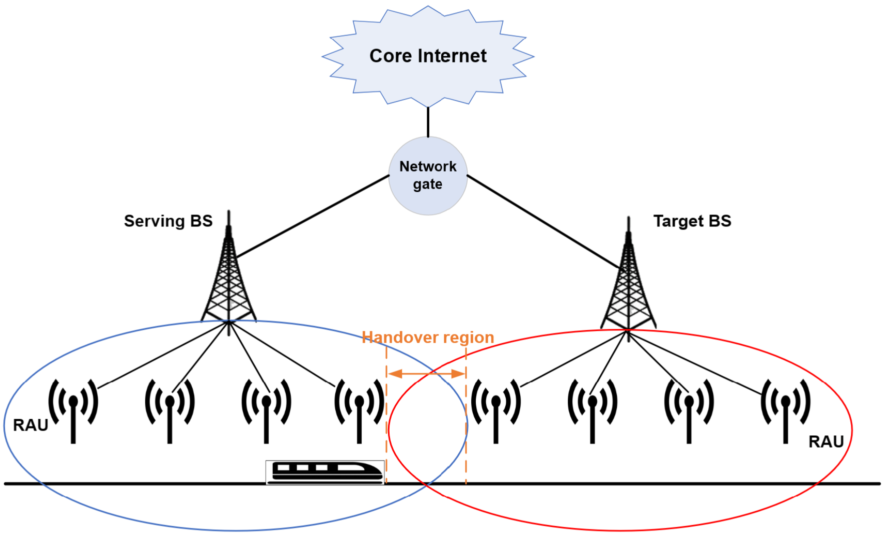

Figure 4.1: Distributed antenna system [15]

Specifically designed for the HST scenario, this DAS network architecture comprises numerous RAUs arranged linearly along the train track. These RAUs are controlled by the same CU which is used for signal processing operations [100, 101] and linked to a CU via optical fiber or wireless links. Notably, the CU's range of control extends up to 20 km, and the need for handovers between RAUs is eliminated due to the utilization of a frequency switch (FSW) scheme as proposed in [102]. The FSW scheme offers a robust mobility signaling process, ensuring a successful transition between frequencies instead of HO. This approach effectively decreases the probability of RLF when compared with the conventional HO process in HSR system. Consequently, DAS is considered as the most effective solution to deal with the HO issue in HSR mobile wireless communications. However, a HO is still needed when the train moves between RAUs controlled by different CUs.

Additionally, there are many problems existing in high-speed rail handover process. For instance, the train passes through the overlapping area so fast that the handover procedure can not be accomplished timely. Additionally, most of the high-speed trains have metal bodies with single or multi-layer glass windows which can cause penetration loss when signals go through the train carriages. This will degrade the received signal strength and lead to handover failure. Meanwhile, a handover will be triggered every 20s assuming a speed of 360km/h along an area with a coverage of 2km. In this case, it is hard to maintain the QoS and system performance is deteriorated. Moreover, when the minimum time required to process handover is larger than the time interval for the HST takes to pass through an overlapping region, then call drops will occur and it causes packet loss. In addition, in HSR, there are a great number of handover requests from different users at the same time. However, because passengers congregate in the small areas on the train, simultaneous requests of handover from a large number of users result in "signaling storm", which will undoubtedly cause huge signaling overhead, and seriously deteriorate system performance [103]. For example, even only 30% of the 1000 UEs in the train are active, there will be 300 simultaneous handover requests.

Although the handover problems for high-speed railway have been widely investigated, most of them consider only the A3 event method to trigger handover in the DAS architecture. A3 is triggered when the received power of the neighboring cell becomes better than that of the serving cell by an offset. In fact, other criteria can also affect the handover performance such as speed, RSRQ, and position, etc. Thus, this chapter aims to propose an improved handover scheme based on dual-layer DAS architecture for HST. In the proposed method, we jointly consider multi-criteria including speed, location, RSRP, and RSRQ, in order to have more precise handover decisions. In addition, the received signal strength and SINR obtained from the measurement report are used as the RSRP and RSRQ respectively.

The main contribution in this chapter can be briefly given as follows:

- We develop an improved handover decision scheme by jointly considering multiple HO criteria including train speed, location, RSRP and RSRQ to achieve better HO decision making.
- Implement, evaluate and compare the proposed method with the conventional HO method in terms of HO trigger probability, HO success probability and outage probability.

The rest of this chapter is organized as follows: In section 4.2, some of the related works are given. In section 4.3, the network architecture model and usage scenarios are introduced. Section 4.4 is devoted to presenting the proposed handover scheme. Section 4.5 analyzes the performance of the proposed scheme. Simulation results are presented and discussed in section 4.6. Finally, Section 4.7 concludes the chapter.

4.2 Related Works

The above challenges have attracted intensive research interest. A communication system using dual-antenna and mobile relay station proposed in [104] can improve the quality of the received signal and provide reliable communication for train-to-ground network to enhance the handover performance. LTE-based HST mobile communication system is proposed in [105], which adopts a faster handover algorithm to reduce the failure probability by decreasing the handover latency with the target of providing reliable broadband services. Based on special features of railway architecture such as straight rail line and motion direction of train, a number of handover schemes based on DAS are proposed [106] [107] [108]. The SINR can be increased especially for users near cell boundaries by DAS in a multicell environment [106]. Additionally, dual-antenna for HSR DAS is proposed in [107] to guarantee higher handover success rate and improve system throughput. A novel handover scheme based on DAS and mobile-relay-aided two-link network is introduced in [108] taking into account the effect of inter-carrier interference to reduce the handover failure rate and link outage probabilities. Furthermore, an expedited predictive handover scheme based on DAS is presented to lower the handover failure rate by starting handover preparation phase in advance via inferring the train position [49]. A handover optimization algorithm based on multi criteria is proposed in [109], considering joint RSRP and RSRQ to trigger handover in order to reduce the handover failure rate. Table 4.1 summarizes the criteria used by existing works and our proposed scheme that are based on DAS architecture.

Table 4.1: Handover trigger conditions in DAS

References	Criteria
[104][107][108]	RSS
[105][106]	SINR
[109]	RSRP, RSRQ
Proposed scheme	RSRP, RSRQ, Speed, Location

4.3 System Model

Considering that the train is running at relatively constant speed and direction on a straight railway line, we take advantage of the specialized DAS network architecture to meet the requirement of high quality wireless communication service. In addition, dual-layer configuration which is similar to the two-hop model [103] is also used to reduce penetration loss and avoid the group handover problem.

4.3.1 HSR Network Architecture

A) **Dual-layer configuration**

The dual-layer configuration has been well recognized in the broadband wireless communication for HST [65, 104], as shown in Fig. 4.2.

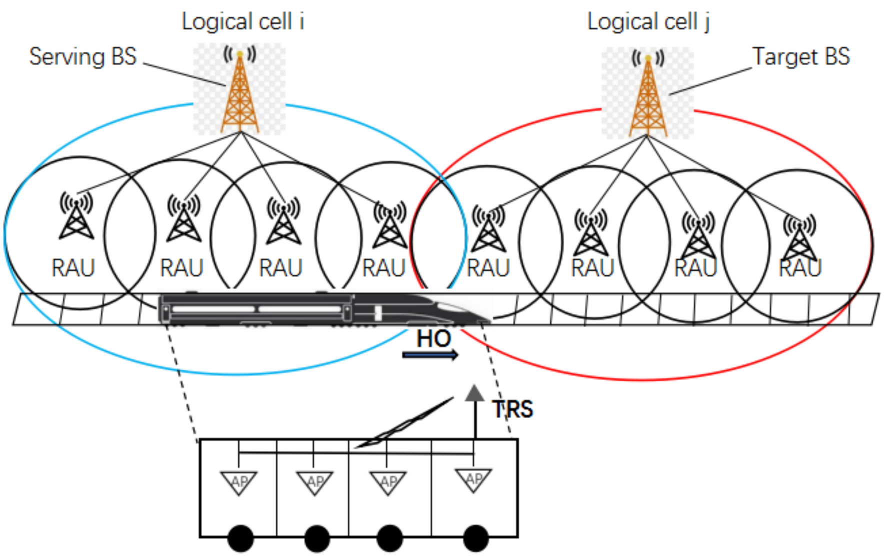

Figure 4.2: System architecture.

The users communicate directly with the access points (AP) which are located inside each carriage of the train. All access points are able to support various types of wireless access technologies such as 3G,4G and WLAN, and are controlled by train relay station. All information collected by the TRS will be sent to BS. Actually, in dual-layer configuration, the TRS acts as a single big user to execute handover from the serving BS to the target BS, which substantially reduces the handover signaling overheads and avoids the group handover problem. However, for a conventional handover scheme, all the active mobile terminals in a train will request a handover at the same time which can result in “signaling storm”.

B) **DAS configuration**

DAS system has been certified as a promising system for HST communication by shortening the distance between transmitter and receiver to provide better communication services [71]. Unlike the conventional collocated antenna system, DAS is divided into several sectors including BSs and RAUs. Several RAUs in the DAS system are geographically distributed in the cell and all RAUs are connected to a central unit or base station via cable or optical fiber in a cell (see Fig.4.2). The left four RAUs constitute a logical cell i which are centrally controlled by the serving base station i. Similarly, the right four RAUs constitute a logical cell j which are centrally controlled by the target base station j. RAUs are only considered as the antenna units, and their aim is to transmit and receive radio signals for their BSs. Because all RAUs are connected to one BS in a cell, spatial diversity technology can be utilized to improve spectral efficiency. Moreover, several RAUs are controlled by the same BS, which can be considered as repeaters for the same signal to form spatial diversity and there is no need to handover when train moves from one RAU to the next one in the same cell. Handover between RAUs in the coverage of the same BS can be avoided via moving frequency concept (MFC) [72] [110]. Using frequency switching algorithm to initiate MFC and then the frequency of RAU can be changed through the implementation of fiber optic radio technology within MFC. The fundamental concept of the “moving frequency” is to alter the radio frequencies of RAUs alternatively in a predetermined sequence based on the position of the HST. Thus, the radio frequency used by the HST remains constant, while the frequencies of the RAUs dynamically adjust to synchronize with the train’s movement. In other words, the frequencies of RAUs are “moving” with the HST. Nonetheless, when the train enters the coverage area of another BS, the handover is still necessary. It’s worth noting that this approach significantly reduces the number of handovers. More specifically, as shown in Fig. 4.3.

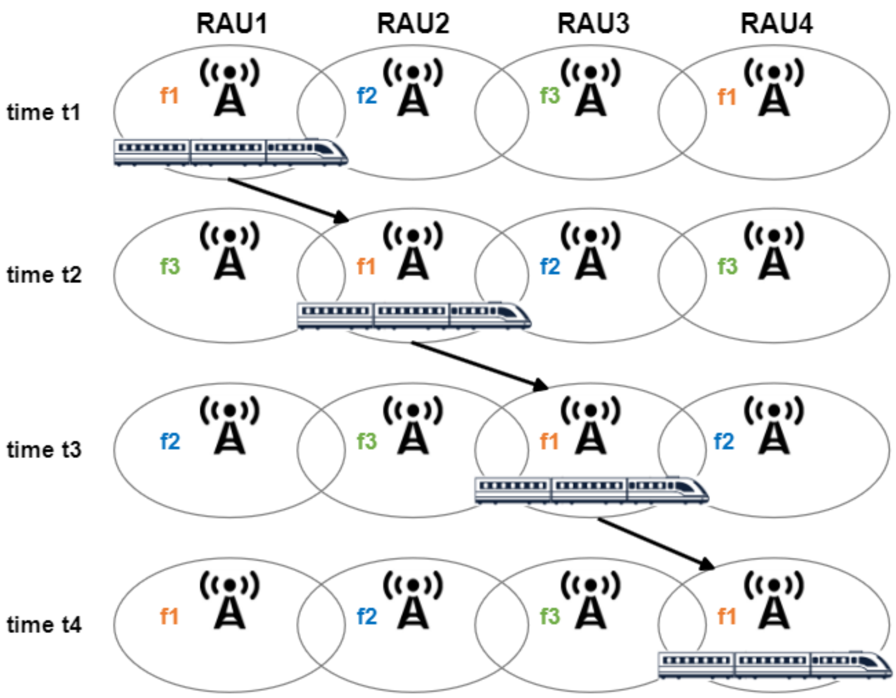

Figure 4.3: Moving frequency concept.

When the HST is in the position of time $t1$, the radio frequencies of RAU1, RAU2, RAU3 and RAU4 are $f1$, $f2$, $f3$, and $f1$. After, when the train is running in the position of time $t2$, the RAU1, RAU2, RAU3 and RAU4 use frequencies as follow: $f3$, $f1$, $f2$ and $f3$, respectively. The operation frequency of the HST during the movement between the different RAUs is unchanged, and the frequency of the RAU moves with the HST. As a result, if the HST is within the coverage of a BS and the handover is not needed. The concept of moving frequency is introduced briefly above, but this is not my focus on this chapter.

Since there is usually more data transmission in downlink than uplink, this work will mainly concentrate on the downlink scenario.

4.3.2 Channel Model

In this section, only the main path signal will be considered since high-speed train runs through rural or viaduct areas most times, so multi-path effect can be ignored. At time t, let a high-speed train be located at x in the i-th cell (current serving logic cell). The received signal power from the cell i is:

$$R(i,x)[\mathrm{dBm}] = Pt[\mathrm{dBm}] - PL(i,d_i)[\mathrm{dB}] - \varepsilon(0,\sigma_i)[\mathrm{dB}] \tag{4.1}$$

where Pt is the transmission power of the i-th cell. $PL(d) = A \cdot d^{-\gamma}$ is the path loss, where A is a constant, d is the distance, γ is the path loss exponent. $\varepsilon(0,\sigma_i)$ is a Gaussian distributed random variable with a zero mean and a standard deviation to characterise the shadow fading.

Similarly, the received signal from target logical cell j is:

$$R(j,x)[\mathrm{dBm}] = Pt[\mathrm{dBm}] - PL\left(j,d_j\right)[\mathrm{dB}] - \varepsilon\left(0,\sigma_j\right)[\mathrm{dB}] \tag{4.2}$$

The co-frequency interference for the serving cell can be given by:

$$I(i,x)[\mathrm{dBm}] = 10\log_{10}\left(\sum_{n_i=1}^{N_{cell}} 10^{R(n_i,x)/10}\right) \tag{4.3}$$

Similarly, the co-frequency interference for the target cell can be given by:

$$I(j,x)[\mathrm{dBm}] = 10\log_{10}\left(\sum_{n_j=1}^{N_{cell}} 10^{R(n_j,x)/10}\right) \tag{4.4}$$

Where N_{cell} is the number of co-frequency cells. $R(n_i,x)$ means the received interference signal power of the serving cell from the co-frequency cells. $R(n_j,x)$ is the received interference signal power of the target cell from the co-frequency cells.

Thus, at location x, we can express the SINR from source BS and target BS respectively as follows:

$$SINR(i,x)[\mathrm{dB}] = R(i,x)[\mathrm{dBm}] - (I(i,x) + B \cdot N_0)[\mathrm{dBm}] \tag{4.5}$$

$$SINR(j,x)[\mathrm{dB}] = R(j,x)[\mathrm{dBm}] - (I(j,x) + B \cdot N_0)[\mathrm{dBm}] \tag{4.6}$$

Where N_0 is the noise power spectral density (dB/Hz). B is signal bandwidth (Hz).

4.4 Proposed HO Decision Scheme

According to [106], several transmission schemes such as RAU selection transmission scheme and blanket transmission scheme can be adopted for data transmission in a DAS cell. In the blanket transmission scheme, each BS allocates its available power to its RAUs equally, even during the handover process. For the RAU selection transmission scheme, it is a strategy used in wireless and cellular communication systems to optimize the selection of antennas within a DAS network, which selects one RAU with minimum propagation path loss for transmission. [65] shows that the RAU selection transmission is more appropriate for wireless HSR. In this chapter, we are concern about the handover between the last RAU of the source cell and the first RAU of the target cell, namely the handover between two adjacent logical cells.

In this section, a handover trigger scheme is proposed considering four parameters, i.e., speed, location, RSRP, and RSRQ. The handover decision algorithm determines whether to trigger a handover according to the measurement reports sent by the UE. A good handover decision algorithm can guarantee the success probability, avoid the occurrence of the ping-pong effect, and ensure good quality of service and experience of passengers.

RSRP and RSRQ are the key parameters of physical layer measurement. RSRP is a key measurement of signal level for modern LTE networks. When a mobile device moves from one cell to another, it must measure the signal strength received from the neighboring cells. In a low-speed environment, the RSRP received by the UE has small fluctuations and the received signal strength is more accurate, so the RSRP based handover decision algorithm is widely used. However, the RSRP measurement values for the two adjacent cells are close and if channel conditions are not good, RSRP based scheme may not trigger a handover in time, causing link failure. In addition, in a high-speed environment, because the channel environment changes very quickly, considering RSRP only to trigger a handover easily results in ping-pong effect.

The RSSI is the average power of the received signals of the UE in the measurement bandwidth and it is used to indicate the interference condition of the channel. It includes the impact of interference and noise in wireless communication. Specifically, interference and noise can both impact the strength and quality of the received signal, which in turn can affect the RSSI value. Interference and noise can cause the received signal to be weakened, leading to fluctuations in the RSSI value. RSRQ is defined as the ratio of RSRP to RSSI, expressed as

$$RSRQ = N \times \frac{RSRP}{RSSI} \tag{4.7}$$

where N is the number of resource blocks within measurement bandwidth. RSRQ is determined by reference signals, interference singals and noise signals, which is similar to the SINR of the link. Thus, RSRQ can effectively reflect the received signal quality. During the handover, RSRQ is seriously affected by the system load and measurement strategy. Considering RSRQ only may lead to increased drop rate. Therefore, a handover decision algorithm that considers RSRP and RSRQ jointly can enhance the overall performance of handover. The proposed handover decision algorithm, first checks the RSRP. When the RSRP of the target BS is higher than that of the source BS by a certain threshold, handover can be triggered. Otherwise, the algorithm then checks RSRQ of the target BS. If it is higher than that of the source BS by a certrain threshold, handover can be triggered.

Additionally, speed can also have a big impact on handover performance [41]. In a given measurement period T, with higher train speed, the train moves further during TTT, then the handover trigger position can be further away from the right position which will cause higher handover failure probability. For this reason, speed is also used as one of the parameters to make handover decision. Conventional A3 event-based handover algorithm is very prevalent. However, it is mainly designed for the low speed scenarios such as speed less than 120 km/h [41]. Thus, when the velocity is higher than 120km/h, the effect of speed on handover performance should be considered. In the proposed handover scheme, if the speed is higher than 120km/h, then the algorithm checks the location of users. Otherwise, traditional A3 trigger event will be adopted to determine a handover. Another criterion, location is also used to check if the train reaches the midpoint of the overlapping area. Specifically, when the train does not reach the midpoint of the overlapping area, A3 trigger event with considering TTT is used to avoid false triggers and ping-pong effect caused by signal fluctuations in high-speed scenario and to ensure the stability of handover triggers. When the train reaches the midpoint of the overlapping area, checking the RSRP without considering TTT parameter to decrease the handover delay. According to the analysis of [111], A3 event has the best performance in terms of HO probability or HO success rate when compared with A4 and A5 events for high-speed railway scenario.

Generally, a handover will take place between two neighboring BSs when the criteria of handover is met. We assume that the handover criteria is met when the RSRP of the target station is larger than that of the serving station by a certain threshold for a certain period of time called TTT. If the TTT is too long, the probability of radio link failure might increase before the handover is finished. This is mainly because the time interval available to execute the handover decreases as the train velocity increases based on the same overlapping area. Similarly, in case of 5G system, the time available to perform handover successfully between base stations also decreases substantially as the cell size smaller. Thus, in the proposed scheme, when the train enters the overlapping area between two neighboring BSs, the handover should be triggered at the predefined location where the handover criteria is met without considering the waiting time TTT. So, the time for performing HO is decreased by omitting the time delay caused by TTT, especially in the high-speed railway scenario. Specifically, the handover will be triggered at one location where the RSRP or RSRQ of the target BS is larger than that of the serving BS by a certain threshold. The flow chart of the proposed handover decision algorithm is shown in Fig. 4.4.

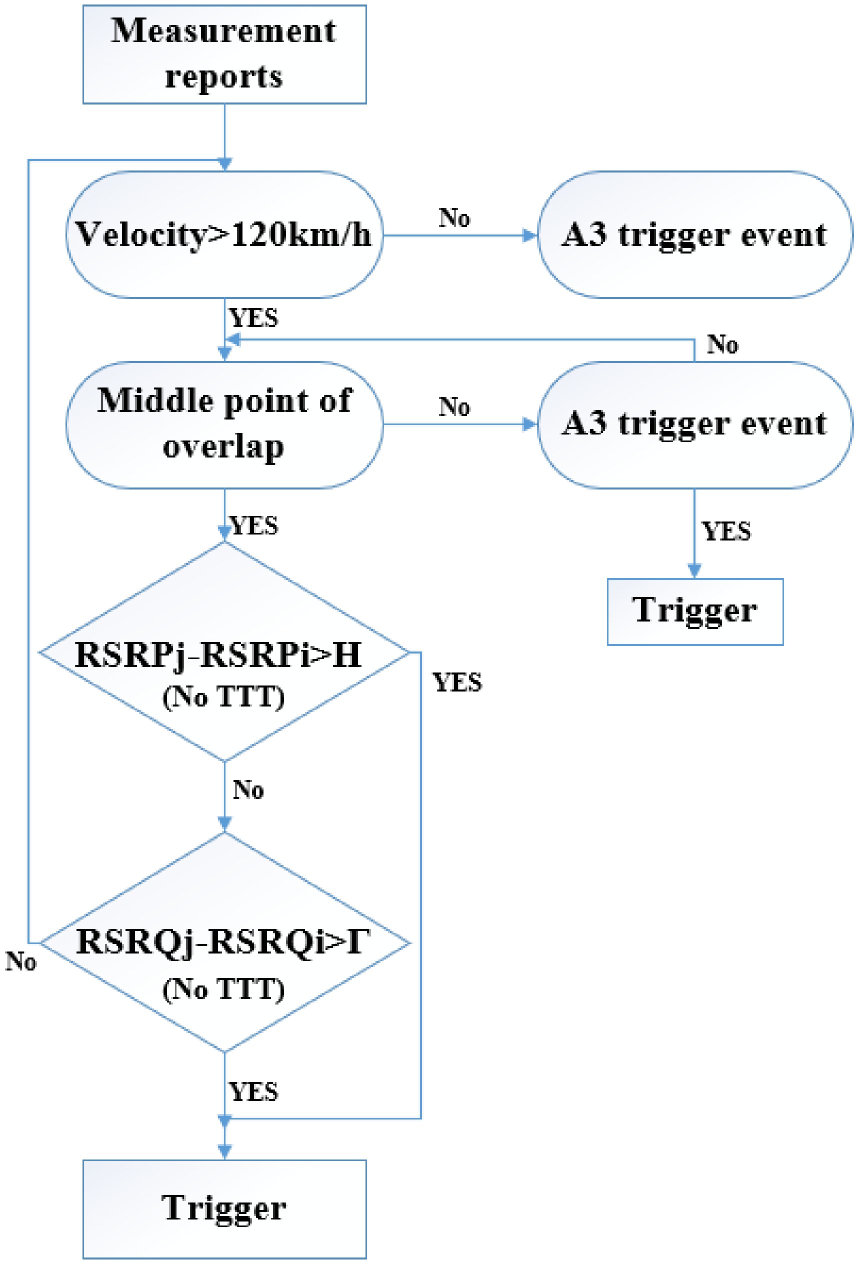

Figure 4.4: The flow chart of the proposed handover decision algorithm.

4.5 Performance Analysis

The performance of the proposed handover trigger scheme is evaluated in terms of handover trigger probability, handover success probability and the outage probability.

4.5.1 Handover Trigger Probability

Handover trigger probability is defined in this paper as the probability that handover occurs at position x. According to the proposed handover decision algorithm in DAS cell, the handover trigger probability of HST can be written as:

$$\begin{aligned} P(x)_{\text{trigger}} =& P[R(j,x)-R(i,x) \geq H] \\ &+ P[R(j,x)-R(i,x) < H] \cdot P[SINR(j,x)-SINR(i,x) \geq \Gamma] \end{aligned} \tag{4.8}$$

Based on equations from (4.1) to (4.6), we can obtain the following equations:

$$\begin{aligned} &P[R(j,x)-R(i,x) \geq H] \\ &= P\left[-PL\left(j,d_j\right)+PL(i,d_i)-\varepsilon\left(0,\sigma_j\right)+\varepsilon\left(0,\sigma_i\right) \geq H\right] \\ &= P\left[\varepsilon\left(0,\sigma_i\right) \geq H+PL\left(j,d_j\right)-PL(i,d_i)+\varepsilon\left(0,\sigma_j\right)\right] \\ &= P\left[\varepsilon\left(0,\sigma_i\right) \geq H+PL\left(j,d_j\right)-PL(i,d_i)+\varepsilon_0 \mid \varepsilon\left(0,\sigma_j\right)=\varepsilon_0\right] \\ &\cdot P\left[\varepsilon\left(0,\sigma_j\right)=\varepsilon_0\right] \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma_j^2}} Q\left(\frac{H+PL\left(j,d_j\right)-PL(i,d_i)+\varepsilon_0}{\sigma_i}\right) \\ &\exp\left(-\frac{\varepsilon_0^2}{2\sigma_j^2}\right) d\varepsilon_0 \end{aligned} \tag{4.9}$$

$$\begin{aligned} &P[R(j,x)-R(i,x) < H] \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma_j^2}} \left[1-Q\left(\frac{H+PL\left(j,d_j\right)-PL(i,d_i)+\varepsilon_0}{\sigma_i}\right)\right] \\ &\exp\left(-\frac{\varepsilon_0^2}{2\sigma_j^2}\right) d\varepsilon_0 \end{aligned} \tag{4.10}$$

$$
\begin{aligned}
&P[SINR(j,x) - SINR(i,x) \geq \Gamma] \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma_j^2}} \\
&Q\left(\frac{\Gamma + PL\left(j,d_j\right) - PL(i,d_i) + I(j,x) - I(i,x) + \varepsilon_0}{\sigma_i}\right) \\
&\exp\left(-\frac{\varepsilon_0^2}{2\sigma_j^2}\right) d\varepsilon_0
\end{aligned}
\tag{4.11}
$$

Where, *Q(x)* is defined as follows:

$$
Q(x) = \frac{1}{\sqrt{2\pi}} \int_{x}^{\infty} \exp\left(-\frac{t^2}{2}\right) dt \tag{4.12}
$$

Based on the above analysis, the handover trigger probability of proposed scheme can be expressed as follows,

$$
\begin{aligned}
P(x)_{\text{trigger}} = &\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma_j^2}} Q\left(\frac{H + PL\left(j,d_j\right) - PL(i,d_i) + \varepsilon_0}{\sigma_i}\right) \exp\left(-\frac{\varepsilon_0^2}{2\sigma_j^2}\right) d\varepsilon_0 \\
+ &\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma_j^2}} \left[1 - Q\left(\frac{H + PL\left(j,d_j\right) - PL(i,d_i) + \varepsilon_0}{\sigma_i}\right)\right] \exp\left(-\frac{\varepsilon_0^2}{2\sigma_j^2}\right) d\varepsilon_0 \\
\cdot &\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma_j^2}} Q\left(\frac{\Gamma + PL\left(j,d_j\right) - PL(i,d_i) + I(j,x) - I(i,x) + \varepsilon_0}{\sigma_i}\right) \exp\left(-\frac{\varepsilon_0^2}{2\sigma_j^2}\right) d\varepsilon_0
\end{aligned}
\tag{4.13}
$$

4.5.2 Outage Probability

Another important metric to evaluate the handover performance is outage probability. Communication outage happens if the received SINR of the TRS is less than a threshold. The outage probability is:

$$
P(i,x)_{\text{outage}} = P[SINR(i,x) < \Upsilon] \tag{4.14}
$$

$$
P(j,x)_{\text{outage}} = P[SINR(j,x) < \Upsilon] \tag{4.15}
$$

where Υ is the minimum SINR required for normal communication.

Substituting equations (4.5)(4.6) into equations (4.14)(4.15), then:

$$
\begin{aligned}
&P(i,x)_{\text{outage}} \\
&= P[SINR(i,x) < \Upsilon] = P[R(i,x) - (I(i,x) + B \cdot N_0) < \Upsilon] \\
&= P\left[Pt - PL(i,d_i) - \varepsilon(0,\sigma_i) - (I(i,x) + B \cdot N_0) < \Upsilon\right] \\
&= P\left[\varepsilon(0,\sigma_i) > Pt - PL(i,d_i) - (I(i,x) + B \cdot N_0) - \Upsilon\right] \\
&= Q\left(\frac{Pt - PL(i,d_i) - (I(i,x) + B \cdot N_0) - \Upsilon}{\sigma_i}\right)
\end{aligned}
\tag{4.16}
$$

$$
\begin{aligned}
&P(j,x)_{\text{outage}} \\
&= P[SINR(j,x) < \Upsilon] = P[R(j,x) - (I(j,x) + B \cdot N_0) < \Upsilon] \\
&= P\left[Pt - PL\left(j,d_j\right) - \varepsilon\left(0,\sigma_j\right) - (I(j,x) + B \cdot N_0) < \Upsilon\right] \\
&= P\left[\varepsilon\left(0,\sigma_j\right) > Pt - PL\left(j,d_j\right) - (I(j,x) + B \cdot N_0) - \Upsilon\right] \\
&= Q\left(\frac{Pt - PL\left(j,d_j\right) - (I(j,x) + B \cdot N_0) - \Upsilon}{\sigma_j}\right)
\end{aligned}
\tag{4.17}
$$

4.5.3 Handover Success Probability

In our proposed scheme, the handover success probability can be defined as the probability when the following three conditions are met simultaneously. At first, there is no communication interruption before the handover is triggered. Then, the handover is triggered at appropriate position and successfully executed. Finally, there is no communication interruption after the handover is completed. Thus, when the train is at location x, the probability of successful handover can be expressed as:

$$
\begin{aligned}
&P(x)_{\text{success}} \\
&= [1 - P(i,x)_{\text{outage}}] \cdot P(x)_{\text{trigger}} \cdot [1 - P(j,x)_{\text{outage}}]
\end{aligned}
\tag{4.18}
$$

4.6 Simulation Results and Discussion

The performance of our proposed handover scheme is also evaluated by simulation based on dual-layer DAS cell architecture. The speed of the train is assumed to be 360km/h. Based on the HST channel model recommended by 3GPP [112], a single-path fading channel is used in the simulation. In the obtained simulation results, each point of the curves was obtained as the average of over 1000 channel realizations. The parameters used in the simulation are listed in Table 4.2.

Table 4.2: Simulation parameters

Parameters	**Value**
Bandwidth of System	10 MHz
Carrier Frequency	2.5 GHz
Total Transmit Power Pt	86 dBm
Path Loss Model	31.5+35log10(d)
Shadow Fading Deviation	4 dB
Noise Density	-145 dBm/Hz
Signal Threshold	-30 dB
Distance from BS to Rail	Dbs,rail=100 m
Distance from RAU to Rail	Drau,rail=60 m
Distance between two cells	ds=3000 m
Number of RAU in one cell	N=4
Hysteresis Threshold H	1 dB

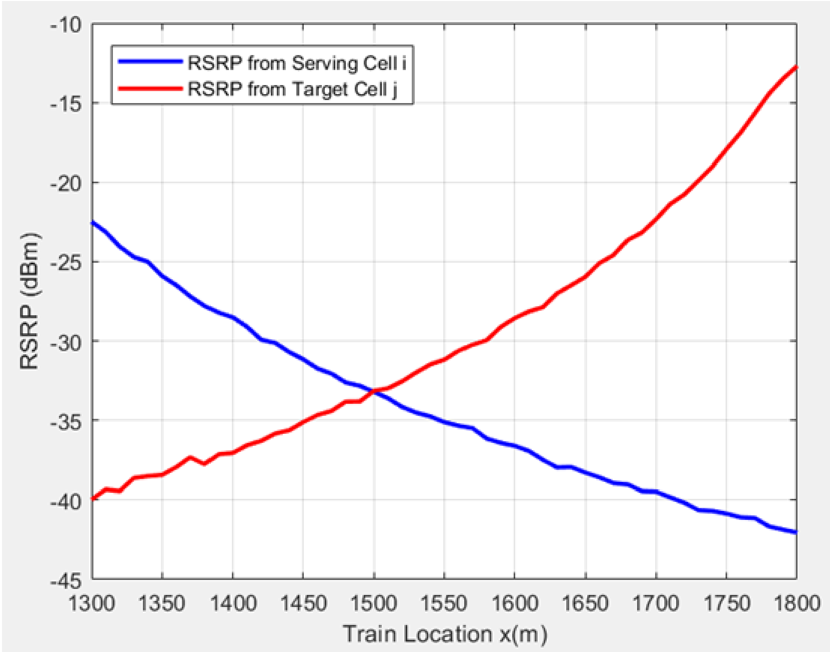

Figure 4.5: Relationship between RSRP and train location

Fig. 4.5 shows the RSRP simulation results received by the mobile train at different train positions. The blue line represents the RSRP received from the serving cell i, and the red line represents the RSRP received from the target j. From the results, we can conclude that the received signal strength from the serving and target cells is identical at location 1500m which indicates that the train arrives at the middle point of overlapping area between two cells.

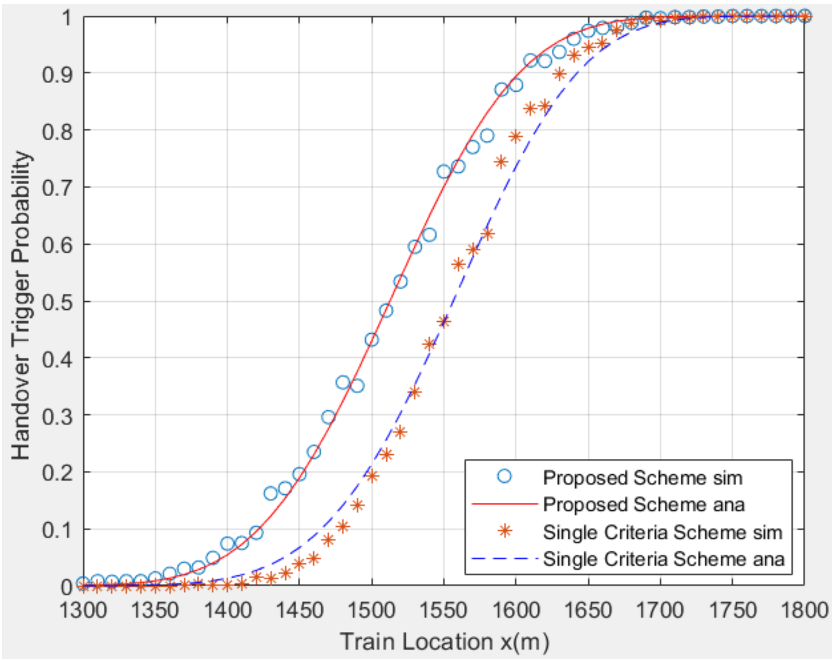

Figure 4.6: Handover trigger probability

Fig. 4.6 presents the handover trigger probability versus the train location *x*. To verify the analysis in section 4.5, both analytical results and simulation results are presented, and it is clear that the simulation results closely align with the analytical results. The curve with red star indicates the cumulative distribution probability of conventional handover which only considers RSS as the triggering metric. For example, when the train location is at 1500m, the TRS has 20% handover triggering probability. The curve with blue circle is the cumulative distribution probability of the proposed handover triggering method. As shown in the figure, the handover triggering probability of the proposed scheme is 22% higher than that of the traditional scheme when the train is located at 1500m.

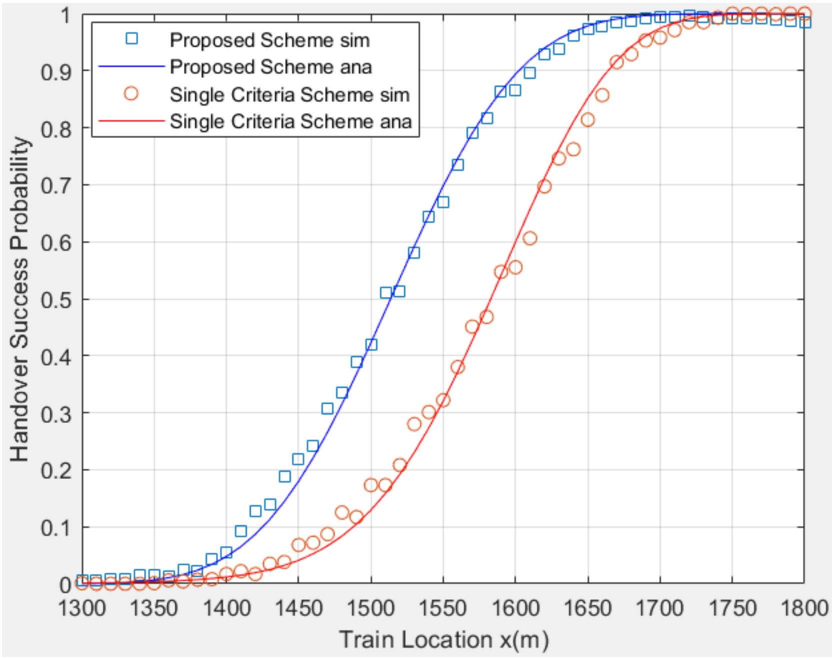

Figure 4.7: Handover success probability

Fig. 4.7 is the probability of handover success. In this figure, both simulation results and analytical results are presented, and we can see that the simulation results closely align with the analytical results. The curve with orange circle describes the A3 event method. The curve with blue square is the result of the proposed scheme. In addition, it is obvious that the proposed scheme can substantially enhance the handover performance when compared with the A3 event method. At distance 1600m, the successful handover probability rises from 60% to 90%.

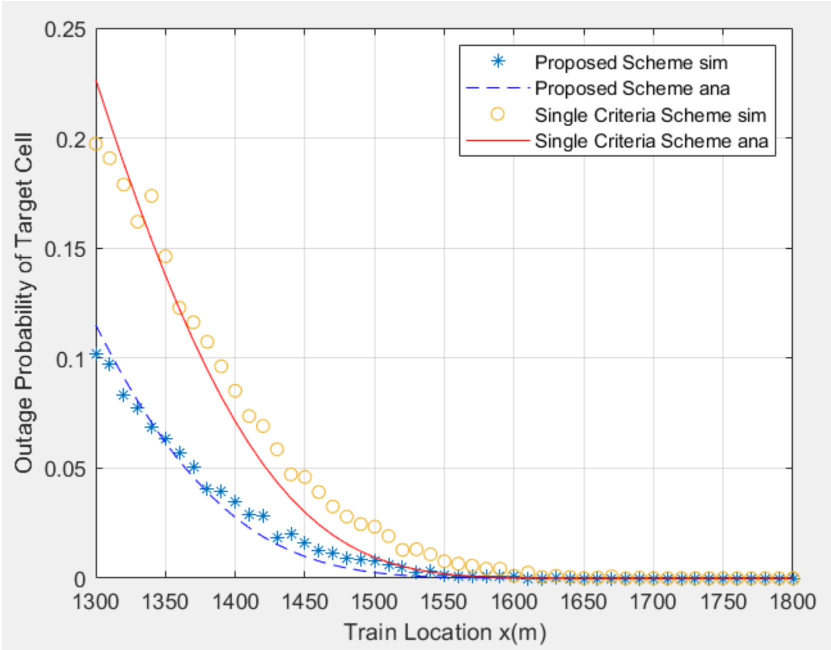

Figure 4.8: The outage probability

Fig.4.8 illustrates the outage probability of the target cell vs. train location. Again, the simulation results closely align with the analytical results. The curve with orange circles is the conventional method. The curve with blue stars is the result of the proposed scheme. In addition, the proposed scheme can substantially reduce the outage probability when compared with the conventional one. With the high-speed train moving towards the target cell, its outage probability is gradually reduced. It is close to 0 when the handover location exceeds 1550 m. This clearly indicates that the proposed scheme can achieve a seamless handover when the handover location is properly selected.

4.7 Summary

This chapter has investigated the handover problems of high speed rail wireless communication systems. Based on dual-layer and DAS architecture, an improved handover scheme employing multi-criteria is proposed to achieve better handover performance, which is analyzed in terms of handover trigger probability, handover success probability and outage probability. The simulation results coincide with the theoretical analysis perfectly and they both show that the proposed scheme can improve the handover success probability in comparison with the conventional RSRP based handover decision algorithm.

Chapter 5

Q-learning based Handover Algorithm for High-Speed Railway Wireless Communication

5.1 Introduction

In recent years, HSRs have offered transport that runs considerably fast in excess of 300km/h for the new lines and 200km/h for the existing ones. HSRs can immensely increase mobility and improve customer satisfaction [113]. With the booming development of HSRs, the evolution of high speed rail wireless communications systems from LTE-R to 5G is essential for reliable safety-related communications and seamless mobile connectivity of passengers, but also faces new challenges, one of which is the handover problem. In the next generation wireless cellular networks, the HO process has even more significant impact on the network performance. For instance, the number of HOs between different cells will increase in the 5G network because of the smaller coverage area of the cells. Frequent HOs cause disruption of data service, higher signaling overhead and decreased network throughput. Moreover, most of the high-speed trains have metal bodies with large windows of a single layer or multi-layer glasses, therefore, at higher frequency bands, such as mmWave that is used in 5G system, signals passing through the compartments will experience more significant attenuation. Additionally, handover failure occurs if the

UEs fail to connect to the target cell. The probability of handover failures increases as the speed of train increases. In cellular networks, the mobility management entity is tasked with facilitating a seamless handover operation without interruptions, as emphasized in [114]. With users on trains moving at significantly high speed and the utilization of higher frequency bands in the 5G HSR system, there arises a question about the capability of existing 5G mobility management designs to support uninterrupted high-speed mobility. Furthermore, as 5G is expected to connect a larger number of users, resulting in a denser network, it becomes imperative to develop a handover management scheme that ensures optimal performance for users. This poses an open research challenge within 5G technology. Thus, the objective of this chapter is to propose a novel HO scheme that can reduce the number of HOs and enhance the HO success probability without compromising the performance of the communications in HSR.

In this chapter, an efficient HO scheme is developed based on RL for 5G high-speed rail wireless communication system. In a dense 5G high-speed rail deployment, HO decisions can be optimized by utilizing Q-learning algorithm in order to reduce the unnecessary HOs and improve the network performance. In addition, the criteria are selected considering the characteristics of the high-speed rail movement. The proposed scheme aims to reduce the HO rate and achieve better throughput by jointly considering four parameters including HO cost, SINR, RSRP and ToS as rewards. By using the Q-learning algorithm, the current HO decision of users is associated with the long-term benefit so that the overall HO performance can be improved. The effectiveness of this method is demonstrated using simulation results.

The rest of this chapter is organized as follows. In section 5.2, the related works are given. Then, the network architecture model and usage scenarios are introduced in section 5.3. Section 5.4 illustrates a brief introduction of the RL and the proposed Q-learning based HO scheme. Section 5.5 evaluates, compares and analyses the performance and simulation results of the proposed method with other methods in the literature. Finally, section 5.6 concludes this chapter.

5.2 Related Works

There already exist several different schemes which solve the HO problems mentioned above. Work in [115] adopts mobile relays (MRs) to decrease the signaling load caused by the large number of HO procedures triggered by the UEs in the same carriage at the same time. In addition, authors of [49] adopt a DAS that combines with the two-hop architecture and proposed a fast predictive HO algorithm to reduce the HO latency and handover command failure probability. In [116], type-2 fuzzy-based HO optimization for advanced long term evolution (LTE-A) network was proposed, with the aim to reduce call drops and the number of handovers. In the future, machine learning will play a crucial role in many innovative 5G technologies and architectures [117]. It is an algorithm which has a self-learning ability to the systems and improve the performance from its training experience. Machine learning based algorithms are required to accurately apply HO decisions which enable the system to automatically adjust parameters according to the users' demand and requirements. As a result, frequent HO can be reduced efficiently [78]. The authors in [118] introduce a HO optimization algorithm based on RL for 5G system. They model the HO problem as a contextual multi-armed bandit problem and then solve it using Q-learning algorithm to improve the performance of the link beam. In [119], the authors present a solution based on RL to select the optimal BS for proactive decision HO in mmWave wireless communication. Their results illustrate that intelligent self-learning agent can decrease the number of HOs. A self-learning framework is proposed in [120], which jointly optimizes the performance of HO and beamforming for mmWave network. RL algorithm is exploited to determine the optimal backup BSs along user trajectory which can help reduce the overhead signalling during channel estimation and minimize the number of HOs. In [121], the authors introduce a RL based mobility model for drone UEs. This model provides flexible HO decision making for a specific flight trajectory. It achieves a trade-off between the number of HOs and the received signal power. Optimization of HO by jointly considering AHP-TOPSIS and RL algorithm is discussed in [122], where optimal BS is selected by using AHP-TOPSIS, then RL approach is utilized to set the parameters for the triggering point such as TTT and hysteresis according to the different speed of users for LTE-Advanced network. Their results show that the proposed scheme can minimize the handover failure rate and ping-pong effect. The algorithms in [96] and [123] present methods to decrease the link failure during HO process. An intelligent handover scheme based on the Elman neural network in the HSR scenario is proposed in [96], where the authors associate past measurement parameters such as RSRP and RSRQ with future handover decisions, which can accelerate the HO execution and optimize the HO process. In [123], fuzzy Q-learning is used to optimize two contradictory handover problems, which are RLF and the ping-pong effect for LTE network. The aim is to minimize RLF for real time users and reduce ping pong effect for non-real time users while keeping RLF within acceptable limits. Although the machine learning based HO methods have been widely investigated in the literature, the problem of HO optimization using machine learning for HSR still remains an open problem. As aforementioned, there are a few that consider RSRP measurements, HO success probability, etc. as rewards. However, different reward configurations such as SINR, train speed and downlink-throughput, etc. can also be taken into account to achieve the optimized HO decision in 5G HSR system. Thus, we use a reinforcement learning approach (Q-learning) to select the target BS from available possible neighbor BSs. For this selection, four criteria are considered, and consequently, the intelligent selection of the target BS is impacted by the weighted combination of four criteria as reward.

5.3 System Model

5.3.1 HSR Network Architecture

In [124], the challenges of high speed railway communications are to maintain consistent user experience and reliable and seamless communication. In high-speed rail system, dedicated BS deployment along railway track is considered, as illustrated in Fig. 5.1. The network is covered by macro BSs with carrier frequency 4 GHz. The distance between the BS and railway track is 100 m. Passenger UEs are located in train carriages. For the passenger UEs, all information is collected by the TRS which is placed on the top of the train. The TRS acts as a single big user to execute HO from the serving BS to the target BS to reduce the HO signaling overheads and avoid the penetration loss caused by the train. The passengers can get access to the internet via WiFi or access points that are connected to the train relay by wired link.

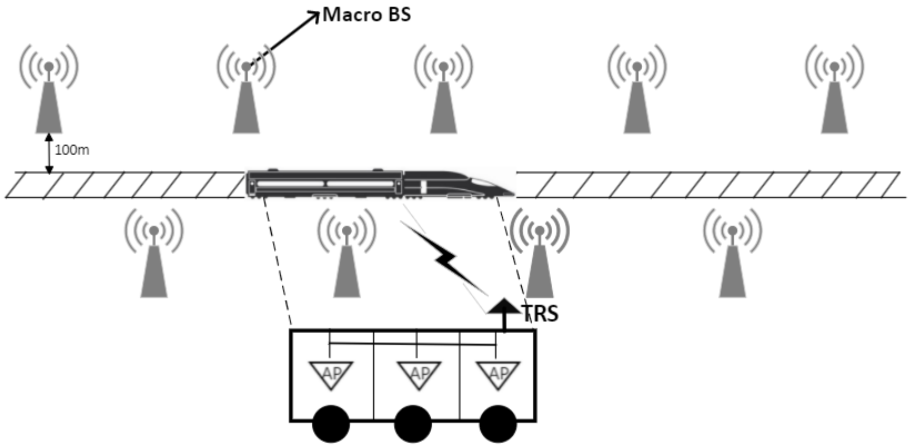

Figure 5.1: System architecture

5.3.2 Channel Model

Since high-speed train runs through rural or viaduct areas most of the time, so multi-path effect is disregarded in this chapter. At time t, let a high-speed train be located at position x in the i-th cell. The received signal power as well as RSRP from the cell i is:

$$RSRP(i,x)[\text{dBm}] = Pt[\text{dBm}] - PL(i,d_i)\,[\text{dB}] - \zeta\,(0,\sigma_i)\,[\text{dB}] \tag{5.1}$$

where Pt is the transmission power of the i-th cell. $PL(i,d_i)$ is the path loss. $\zeta\,(0,\sigma_i)$ is a Gaussian distributed random variable with a zero mean and a standard deviation to describe the shadow fading.

The pathloss model is proposed in [89, 125], which is applicable to HSR scenario. Thus, the path loss (in decibels) of the i-th cell can be expressed as

$$PL(i,d_i)\,[\text{dB}] = 10\theta\log(d_i) + PL(d_0)\,, \tag{5.2}$$

where θ is the path loss exponent, d_i is the distance between the BS and the moving train, $d_0 = 1$ m is close-in free space reference distance thus $PL(d_0) = 10\theta\log(4\pi d_0/\lambda)^2$, λ is the wavelength.

We consider inter-cell interference caused by the nearest neighboring cells and express it as

$$I_i[mW] = \sum_{n=1, n \neq i}^{N_{nei.i}} 10^{RSRP(n,x)/10}, \tag{5.3}$$

where $N_{nei.i}$ is the number of the co-channel cells. $RSRP(n,x)$ obtained from (5.1) is the received interference signal power from the co-channel cells.
Then, we can obtain the SINR of the i-th cell as

$$SINR_i[dB] = 10\log\left[\frac{10^{RSRP(i,x)/10}}{I_i + N_0}\right], \tag{5.4}$$

where N_0 is the noise power.

5.4 Proposed Handover Scheme

In this chapter, we first introduce the Q-learning algorithm in the RL framework, and then a Q-learning based HO algorithm is designed to solve the frequent HO problem while maintaining the reliable connection.

5.4.1 Background of Reinforcement Learning

RL is a type of machine learning which can make an agent learn in an interactive environment by a great number of trials and errors utilizing feedback from its own actions and experiences. It is about taking suitable action to maximize the reward in a particular situation [126] [127].
RL is different from the supervised learning. Specifically, in the supervised learning, the training data has the answer key with it so the model is trained with the correct answer whereas in RL, there is no answer but the reinforcement agent decides what to do to perform the given task. In the absence of a training dataset, it is bound to learn from its experience. As compared to unsupervised learning, RL differs from it in terms of their respective goals. The goal in the unsupervised learning is to find similarities and differences between data points. While in the RL, the goal is to find a suitable action model which can maximize the total cumulative reward of the agent. The relationship between agent,

action and environment is shown in Fig. 5.2. The agent learns the best policy by multiple interactions with the environment.

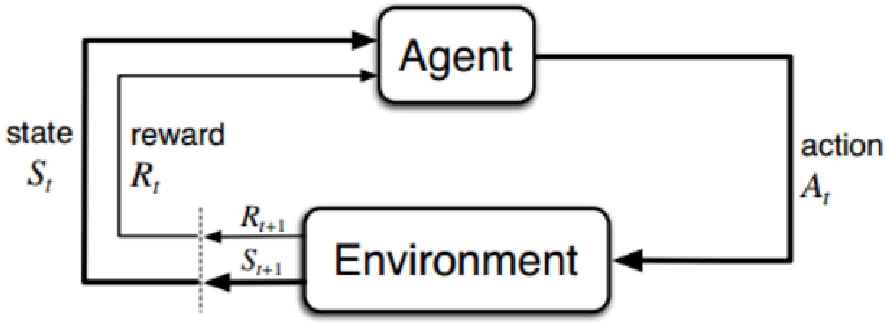

Figure 5.2: Framework of RL [16]

The model comprises the following components:

- Set of states: A state is utilized to characterize the present condition of an agent under specific environmental circumstances. Let the set of states be S.
- Set of actions: The set of all possible actions that an agent can take in a given state is known as the action space. The action space defines the range of choices available to the agent in each state. The new state (next state) can be obtained by taking an action on the current state. Let A represent the set of available actions.
- The probability of state transition from state s_t at time t to a new state s_{t+1} when taking action a_t can be denoted as $P(s_{t+1} \mid s_t, a_t)$.
- Policy: A policy refers to the strategy or a set of rules that guides the agent's decision-making process in choosing actions in different states of an environment. The policy determines how the agent selects actions to maximize its cumulative rewards over time.
- Reward: A reward R, generally refers to a scalar quantity provided by the environment to the learning agent to indicate the immediate feedback. Reward is used to represent whether the action taken is good, bad or neutral in the current state.

Markov Decision Processes (MDPs) are mathematical frameworks which are often used to describe an environment in RL and almost all RL problems can be formulated using MDPs. An MDP usually consists of a set of finite environment states S, a set of possible actions

A, a real valued reward function R and a transition model P. The learning procedure is detailed as follows: At time t, the agent observes the state of the environment, $s_t \in S$, where S is the set of possible states. After observing state s_t, agent will take an action, $a_t \in A(s_t)$ where $A(s_t)$ is the set of possible actions at state s_t. After selecting the action a_t from state s_t, agent will receive the immediate reward R_{t+1} from the state-action pair (s_t, a_t). The selected action in state s_t moves the agent to state s_{t+1} at time t+1. It is crucial for the environment to have state dynamics such that $P(s_{t+1} \mid s_t, a_t)$ exists. This is the probability that the action a_t in state s_t at time t will lead to new state s_{t+1} at time t+1. Specifically, the probability that the process moves into its new state s_{t+1} is impacted by the chosen action. However, real world environment is more likely to lack any prior knowledge of environmental dynamics. Model-free RL method such as Q-learning, is more suitable. Since Q-learning is a model-free RL algorithm to learn the value of an action in a particular state. It does not require a model of the environment. For any finite MDP, Q-learning can find an optimal policy to maximize the expected value of the total reward over all successive steps, starting from the current state, and it can identify an optimal action in a given situation based on trial and error. The model is composed of a set of states, agent and set of actions per state. An action is performed by an agent and it will move from one state to another state. The reward is provided to the agent after executing the action in the specific state. The maximum total reward is the main goal of the agent by learning the optimal action in each state.

In a given state, the action-value-function is utilized to calculate the anticipated utility for an action chosen by the agent. More specifically, this function quantifies the effectiveness of the state-action combination. For state s and action a, the action-value function, often referred to as the Q-function or Q-value, is represented by $Q(s, a)$. Thus, in general, the Q-value can be described as follows:

$$Q : S \times A \mapsto \mathbb{R}. \tag{5.5}$$

At the beginning, the Q-value is chosen by the designer and can be an arbitrary value. After taking an action in the current state, the agent transits to a new state (next state), and the Q-value is computed for the current state. Therefore, the core component of the Q-learning algorithm is based on value iteration, where the old Q-value is updated by considering the current information that is related to the chosen action. The update of the Q-value is given as

$$Q(s,a) \leftarrow Q(s,a) + \alpha \left[R + \gamma \max_{a'} Q\left(s', a'\right) - Q(s,a) \right] \tag{5.6}$$

where $\max_{a'} Q\left(s', a'\right)$ is the maximum Q-value for the next state s' over all possible actions a', indicating the highest expected cumulative reward that the agent could achieve from the next state onwards, considering all available actions. It is used in the Q-learning update rule to adjust the current Q-value for the action a in state s. In addition, the parameter α is the learning rate, γ is the discount factor, respectively. R is known as immediate reward. These parameters are defined as follows:

- Learning rate (α): This parameter determines the extent to which new information gained from experiences in the environment is used to update the agent's Q-value. Additionally, it impacts how quickly or slowly the Q-value is updated, balancing the importance of new information with the agent's prior knowledge. For instance, when $\alpha = 0$, the agent does not learn anything from the environment; Whereas, when $\alpha = 1$, the Q-value is updated entirely based on the most recent information only. It completely overrides the old Q-value with the new estimate. Most often, the learning rate α is set to a value between 0 and 1. In this case, the Q-value update is a weighted average of the old Q-value and the newly acquired estimate. In fact, a constant value, such as 0.1, is used as the learning rate since the learner can be assigned a significant time to obtain information about the environment.

- Discount factor (γ): This factor represents the importance of future rewards. It must be a value between 0 and 1, where a smaller γ makes the agent focus more on short-term rewards, and a larger γ increases its emphasis on long-term rewards. If γ exceeds 1, the Q-value might diverge.

5.4.2 Q-learning Based Handover Algorithm

In this work, we regard the high-speed train as a reinforcement agent and consider the physical network environment of the HO area as the external environment, so the interaction process between the RL agent and the external environment can be modelled as a MDP. The MDP can be represented by the following four tuples: $\langle S, A, R, \pi \rangle$, which are defined as follows:

1) **State**

   The state is represented by S consisting of the train's location and current serving base station.

2) **Action**

   In the conventional HO events such as A3, the users will always select the target BS with the highest RSRP or SINR according to the HO trigger condition resulting in the sub-optimal decision and frequent HO problem. Hence, we define the action $a \in A(s)$ as the scalar representation of the serving BS at current state s. The action space $A(s)$ includes all BSs along with the train railway track. For instance, choosing an action is analogous to select an appropriate BS to handover or stay with the current BS. The aim is to select an action in a given state that maximizes the expected reward.

3) **Reward Design**

   The reward R is a measurement of the rationality of the agent's behavior in a specific environment state. The selection of the reward will impact the learning efficiency of the agent and reflect the environmental feedback in RL. The significance of reward is to motivate the agent to learn to reach the target through reward maximization, and our objective is to reduce the HO number while maximizing the throughput. In a handover process, the reward is closely related to handover and network performance. Thus, we design the immediate reward so that multiple parameters that affect the HO performance can be considered and it can be expressed as follows:

$$R = -\omega 1 \times HOcost + \omega 2 \times SINR + \omega 3 \times RSRP + \omega 4 \times ToS \tag{5.7}$$

where $\omega1$, $\omega2$, $\omega3$ and $\omega4$ represent the weights for the HO cost, SINR, RSRP and ToS respectively. In addition, their range is from [0,1] and $\omega1+\omega2+\omega3+\omega4 = 1$. Based on RL, we can find out the best state-action pair with the maximum R.

The reward function is a weighted combination of four parameters: HO cost, RSRP, ToS and SINR. The definition of RSRP and SINR can be found in section 5.3.2. HO cost is used to describe whether a HO occurs between the current and the next state. For example, HO cost=0 if there is no HO occurring from the current state to the next state, and vice versa. In terms of ToS, it is the expected time that the users will stay in the coverage area of a BS. When combined with the train speed, this parameter can be used to decrease the unnecessary HO number. If the train is always connected to the BS with the best received signal without considering how long the users stay in that cell, it will result in frequent HOs, and consequently, also increase overheads. Ideally, if the time of stay in one cell is too short, then the handover to this cell can be avoided.

4) **Policy**

In a given HO route, policy π refers to the set of rules that the agent uses to determine which actions to take in different states. In the proposed scheme, we select the ε-greedy algorithm [128] [129] as the policy π, which is a common exploration strategy used in Q-learning algorithm and other RL algorithms to balance exploration (trying new actions) and exploitation (choosing the best-known actions). This algorithm helps the agent to gradually learn the optimal policy while exploring the environment. In other words, the ε-greedy policy deals with the trade-off between exploration and exploitation. Moreover, ε means the probability that the agent chooses the next state s'. Its range is from 0 to 1. In the ε-greedy policy, the concepts of exploration and exploitation can be elucidated as follows:

- Exploration: An action is chosen randomly from the set of available actions with the probability of ε. This encourages the agent to explore new actions and states, even if they have lower estimated Q-values.
- Exploitation: An action is selected to achieve the highest Q-value for the current state with the probability of 1-ε.

The ε-greedy algorithm for Q-learning based HO is described in Algorithm 5.1. A Q-table

Algorithm 5.1 Q-learning with ε-greedy algorithm for HO situation

1: Initialization: input parameters
2: Train trajectory $position = \{\mathscr{P}_i \mid i = 0, 1, \ldots, l-1\}$
3: Set $\mathbf{Q} \leftarrow \mathbf{0}_{l \times m \times m}$; $\mathbf{HO}_{s,s'} \leftarrow \mathbf{0}_{m \times m}$
4: Set $w_{HO_{cost}}, w_{RSRP}, w_{SINR}, w_{ToS}, \epsilon, \alpha, \gamma$
5: $C_{m_s} \leftarrow m$ BSs at starting waypoint $\mathscr{P}_0$
6: **for** i in length $(position) - 1$ **do**
7: $\quad C_{m_{s'}} \leftarrow m$ BSs at waypoint $\mathscr{P}_{i+1}$
8: $\quad RSRP_{m_{s'}} \leftarrow$ RSRP values for BSs in $C_{m_{s'}}$
9: $\quad SINR_{m_{s'}} \leftarrow$ SINR values for BSs in $C_{m_{s'}}$
10: $\quad ToS_{m_{s'}} \leftarrow$ ToS values for BSs in $C_{m_{s'}}$
11: $\quad \forall h \in \mathscr{I}_m$, set $\mathbf{Q}[i,:,h] \leftarrow w_{RSRP} \cdot RSRP_{m_{s'}} + w_{SINR} \cdot SINR_{m_{s'}} + w_{ToS} \cdot ToS_{m_{s'}}$
12: $\quad$ Binary matrix $\mathbf{HO}_{s,s'} \leftarrow C_{m_s} \neq C_{m_{s'}}$
13: $\quad \mathbf{Q}[i,:,:] \leftarrow \mathbf{Q}[i,:,:] - w_{HO} \times \mathbf{HO}_{s,s'}$
14: $\quad C_{m_s} \leftarrow C_{m_{s'}}$
15: **end for**
16: Reward matrix $\mathbf{R} \leftarrow \mathbf{Q}$;
17: **for** $n \leftarrow 0$ to number of training **do**
18: $\quad j = 0$
19: $\quad$ **for** i in length $(position) - 1$ **do**
20: $\qquad$ Take a random variable g uniformly from [0,1]
21: $\qquad$ **if** $g < \varepsilon$ **then**
22: $\qquad\quad j_{new} \leftarrow$ select a random number from $\mathscr{I}_m$
23: $\qquad$ **else**
24: $\qquad\quad j_{new} \leftarrow argmax_{y \in \mathscr{I}_m} \mathbf{Q}[i, j, y]$
25: $\qquad$ **end if**
26: $\qquad$ Observe next state and reward
27: $\qquad$ Update the Q values as:
28: $\qquad Q = \mathbf{Q}[i, j, j_{new}]$
29: $\qquad Q = (1-\alpha)Q + \alpha \mathbf{R}[i, j, j_{new}] + \alpha\gamma \arg\max_{k \in \mathscr{I}_m} \mathbf{Q}[i+1, j_{new}, k]$
30: $\qquad \mathbf{Q}[i, j, j_{new}] = Q$
31: $\quad$ **end for**
32: $\quad j = j_{new}$
33: **end for**
34: Return $\mathbf{Q}$

is created by trying different actions for the received states and recording the observed reward. To simplify the model, we limit the action space A to the neighboring m candidate BSs for each state in order to reduce complexity. Let us define a set $\mathscr{I}_m = \{0, 1, \ldots, m-1\}$ which is the m neighboring candidate BSs for every state. For a train railway trajectory with l waypoints, the resulting Q-table is stored in memory and updated based on equation (5.6). The initial Q-table for the given train trajectory is generated in steps 6-15. In step 12, a binary square matrix of size m is generated such that its entry is 0 if the BS at state s is the same as the BS at state s', and it is 1 otherwise. The Q-value iterations for each training episode are performed in steps 17-34, where steps 21-25 perform the ε-greedy exploration while step 29 implements the update equation (5.6). After a great number of iterations, the RL converges and the state with the largest R value selected by the agent will remain stable. In other words, the largest Q value stored in the Q-table represents the optimal selection. Therefore, we can find the optimal HO decision in a given railway trajectory by selecting the largest Q value in each state.

5.5 Performance Evaluation and Simulation Results

In this section, we compare the performance of the proposed Q-learning based algorithm with that of the conventional method and the methods presented in [49] and [130] in terms of the throughput, HO number, HO success probability and HO latency. Firstly, we describe the experiment setup and then present the simulation results with discussions.

5.5.1 Experiment Setup

In order to evaluate the handover performance of the proposed algorithm in the HSR system, we set up a HSR communication network model in MATLAB. In this model, the overall length of the railway track is 10 km and ten BSs are distributed alternately along both sides of the track. Based on the analysis in section 5.4.2, we build a Q table during the training phase of the Q-learning algorithm (Algorithm 5.1). In this phase, the agent will take random actions (HO to certain BS) based on ε-greedy exploration and then update the Q table. After the training phase, the table we obtained can be used in HO decision making.

In addition, the whole track is divided into 100 waypoints evenly and the train will take steps in a straight direction along the railway track. In each step, the measurement reports will be sent to the agent which exploits the Q table to take an action that maximizes the Q values for each state.

5.5.2 Simulation Results

The performance of the proposed HO scheme is evaluated in terms of the average number of HOs, HO success probability, throughput and HO latency, and the performance is the average of 1000 runs. The parameters used in the simulation are listed in Table 5.1.

Table 5.1: Simulation Parameters

Parameters	**Value**
Bandwidth of System	20 MHz
Carrier Frequency	4 GHz
Transmit Power of BS Pt	46 dBm
Shadow Fading Deviation	4 dB
Noise Density	-174 dBm/Hz
Distance from BS to Rail	100 m
Number of BSs	10
Learning Rate α	0.1
Discount Factor γ	0.9
Exploration Coefficient ε	0.1

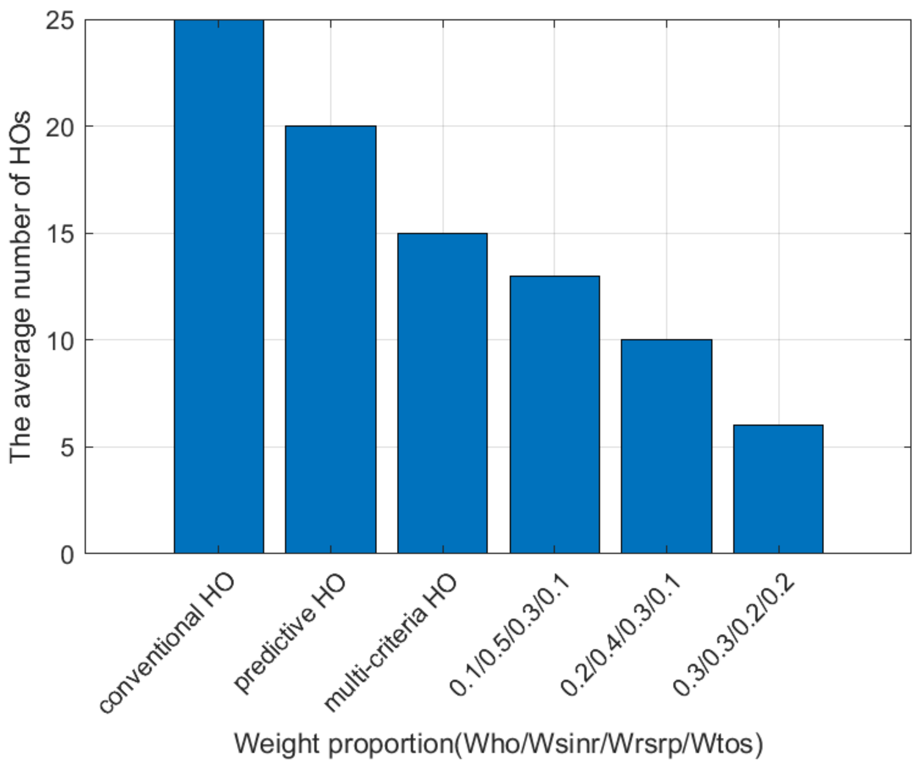

Figure 5.3: Average number of HOs for different weight proportion

In Fig. 5.3, we compare the average number of HOs of the proposed scheme with different weight combinations considering four criteria including HO cost, SINR, RSRP and ToS with the conventional HO method and the existing algorithms [49] [130]. In the conventional HO scheme, the train users always connect to the BS with the strongest RSRP. From the figure, it can be clearly observed that the proposed scheme outperforms the existing algorithms as well as the conventional method and achieves the lowest HO rate. Since the UEs do not have to perform HO frequently due to the incorporation of UE's ToS and HO cost. The proposed scheme can be decreased by 48% when weight proportion $\omega1/\omega2/\omega3/\omega4$=0.1/0.5/0.3/0.1, compared to the conventional scheme. Overall, the simulation result clearly indicates that the proposed scheme is effective in reducing the number of unnecessary handovers.

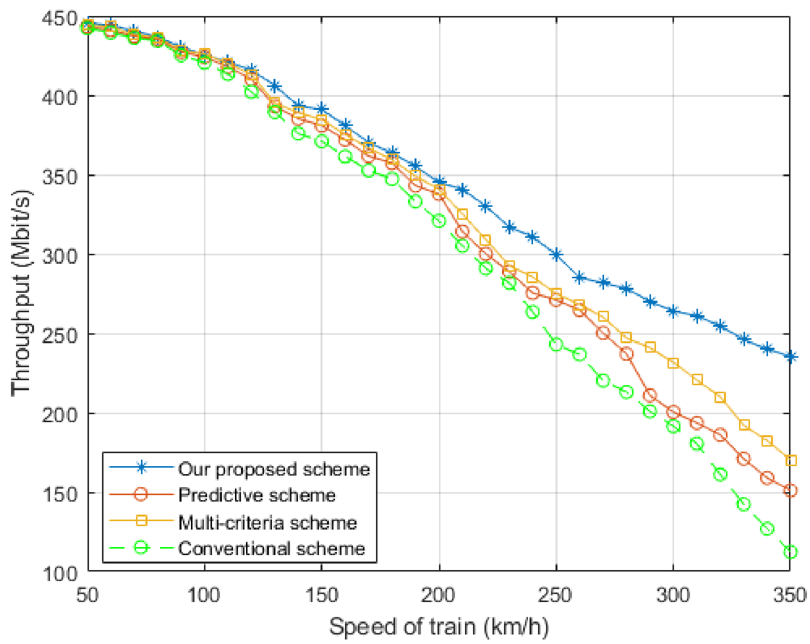

Figure 5.4: Throughput for HSR system

The throughput vs speed of train is shown in Fig. 5.4. The proposed method clearly outperforms the conventional method and other algorithms [49] [130] in the literature. Specifically, when the speed of train is 300 km/h, the throughput of the proposed scheme is increased by about 12.2%, 23.9% and 27.4% respectively in comparison to multi-criteria, predictive HO scheme and conventional HO scheme. The reason is that the proposed method can effectively reduce the HO number, especially when the train moves faster.

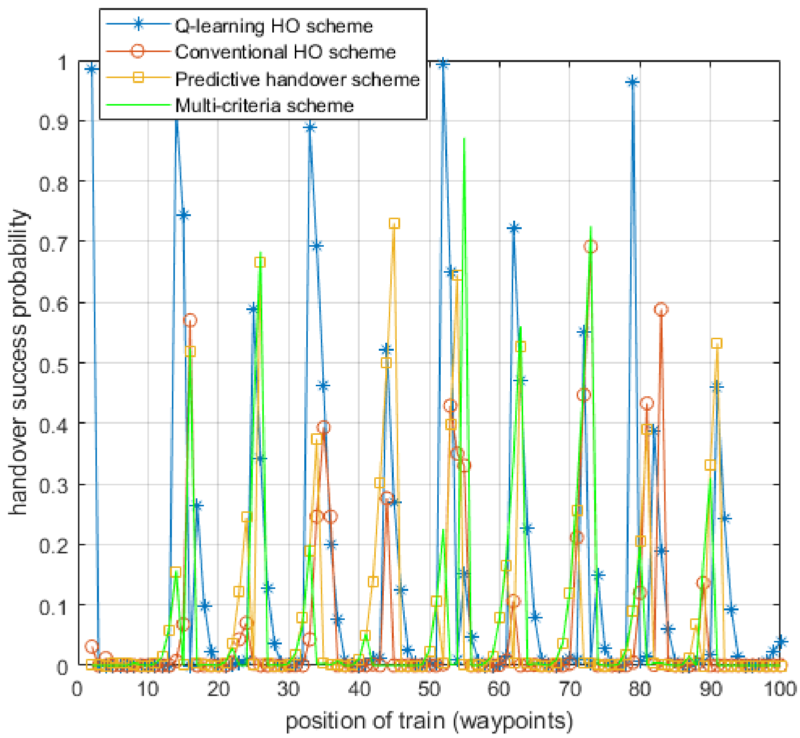

Figure 5.5: Handover success probability

Fig. 5.5 illustrates the handover success probability at different train location (waypoints) for different HO methods. The proposed scheme achieves the best performance at most of the train locations when compared with the conventional HO method and the existing HO algorithms [49] [130]. For those locations with a success probability of 0, that means no HO occurring.

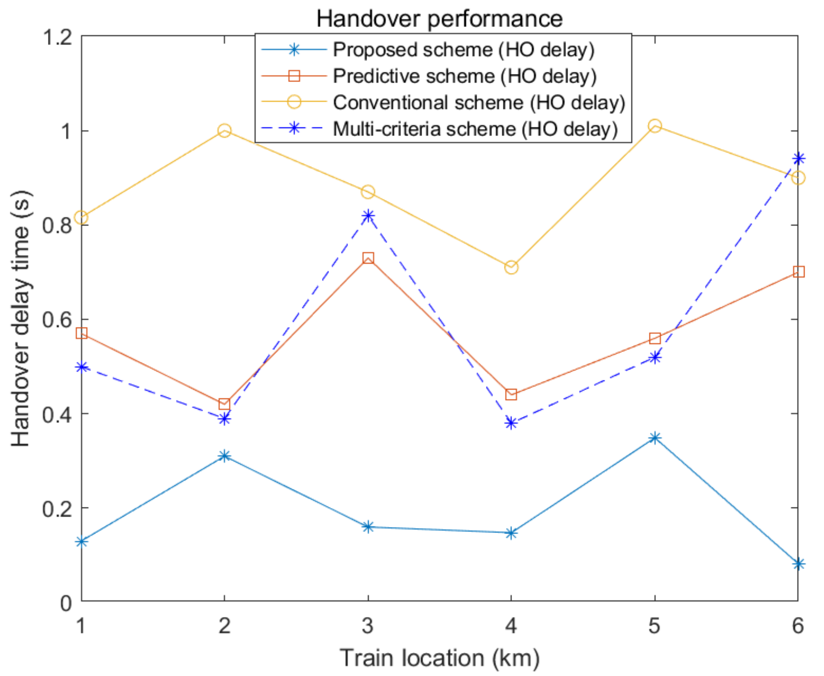

Figure 5.6: Handover latency

Handover latency is another decisive performance metric when it comes to QoS. The handover latency time can be defined as

$$T = Psuccess \times T_s + (1 - Psuccess) \times T_{rec} \tag{5.8}$$

where T_s and T_{rec} are the handover latency in case of successful and failure situations respectively. Note that the calculation of T_s and T_{rec} is based on [27] [48] [49]. More specifically, T_s=1/14 ms is the symbol duration, and $T_{rec} = T304 + T311 + T301$ where $T304$, $T311$ and $T301$ are 50 ms, 100 ms and 1000 ms respectively.

In Fig. 5.6, the abscissa is the train locations that we pick along the whole train railway track and we use the numerical values to represent for convenience. It is clear that the proposed scheme has the lowest HO latency time and also achieves the best user QoS when compared with the conventional HO scheme and other existing schemes [49] [130].

5.6 Summary

This chapter has investigated the HO problems in the HSR wireless communication system. A novel HO scheme based on Q-learning algorithm is proposed to achieve better HO performance. This chapter establishes a Q-learning framework for the HSR HO, and provides an effective method for optimal HO decision making which considers multiple criteria as weighted rewards along the railway track. As demonstrated by the simulation results, the proposed scheme outperforms the conventional approach and existing algorithms in terms of throughput, HO number, HO success probability and HO latency. In summary, the Q-learning based handover scheme offers an intelligent solution for effectively managing high mobility scenarios by intelligently selecting the target cell. Through simulations, it indicates that the proposed scheme has the potential to greatly enhance the HO process in the context of 5G, enabling more efficient management of high-speed mobility in 5G network.

Chapter 6

Machine Learning based Handover Algorithms for High-Speed Railway Heterogeneous Ultra-Dense Network

6.1 Introduction

HSRs have emerged as a transformative mode of transportation, revolutionizing the way people and goods move across long distances. With their remarkable speed, efficiency, and environmental advantages, HSRs have gained significant popularity and are becoming a vital form of transportation. As the demand for fast, reliable, and sustainable transportation grows, effective management of HSR mobility becomes crucial to ensure smooth operations and enhance passenger experience. Thus, how to provide users with the best quality of high-speed mobile network services attracts increasing attention in the research of 5G wireless communications. One of the big challenges is the handover problem, such as the frequent handover, unnecessary handover, etc due to the increased density of small cells and increased speed of trains. In addition, seamless handover procedures are vital to ensure safety, and uninterrupted services in the high-speed rail network. UDN is deployed in 5G-Railway to meet the increasing demand for mobile internet services for large number of train passengers. In addition, mmWave is considered as a promising technique used in the 5G communication for UDN. However, both the small coverage of mmWave and the ultra-dense small cells can result in more severe frequent HO problem. Frequent HO causes disruption of data service, higher signalling overheads and decreased network throughput. Moreover, most of the high-speed trains have metal bodies with large windows of a single layer or multi-layer glasses, which cause great attenuation for high frequency signals. Additionally, handover failure occurs if the UEs fail to connect to the target cell, especially in the ultra-dense network using the mmWave which offers high data rates but experiences higher propagation loss. This means that the signal strength degrades rapidly over distance. As a user moves away from a BS, the signal strength drops quickly, increasing the probability of handover failure. In ultra-dense networks, a large number of small cells are deployed in a limited geographical area, leading to increased interference levels. Co-channel interference can hinder seamless HO and decrease the received signal strength leading to handover failure. Thus, handover problem is becoming more serious in the UDN of HSR that must be addressed.

In this chapter, we develop an efficient intelligent HO scheme based on Kalman filter and machine learning algorithms for 5G heterogeneous ultra-dense network of high-speed railways. This HO scheme follows a two-step methodology. In the first step, it employs a Kalman filter to estimate future (a posteriori) RSRP values of the neighboring BSs. These estimations rely on measurement reports obtained from neighboring BSs. Subsequently, considering the predicted RSRP, in the second step, HO decision can be optimized by utilizing support vector machine algorithm in order to select the target cell and reduce the number of HOs and improve the network performance. The proposed SVM based handover scheme determines the optimal hyper-plane using labeled training data to classify the available target cells by jointly considering multiple HO criteria such as RSRP, SINR and time. In addition, we conduct extensive simulations to demonstrate the effectiveness of our proposed HO algorithm. The simulation results show that the proposed method can improve the HO performance in terms of the number of HOs, handover failure probability, user mean throughput and handover latency in comparison with other machine learning approaches including Decision Tree (DT), K-Nearest Neighbors algorithm (KNN) and Artificial Neural Network algorithm (ANN) and existing methods [131] in the literature.

The contributions of this chapter are summarised below:

- We design a Kalman filter to calculate the posteriori of the RSRP of the neighboring BSs for the handover. Consequently, according to the prediction, the estimation of the future signal quality can help the handover decision-making in advance, ensuring that the chosen target BS can sustain the necessary network performance in such dynamic, high-mobility environments.
- We develop an intelligent handover scheme based on SVM algorithm which selects the target cell by jointly considering multiple HO criteria including RSRP, SINR and time. For the selection of target BS, the RSRP estimation by the Kalman filter is considered, and thus the intelligent selection of the target BS is impacted by the predicted future signal quality of neighboring BSs.
- We introduce the standard deviation (SD) technique to assess the priority weight and effect of different attributes in the BS selection algorithm.

The rest of this chapter is organized as follows. In section 6.2, the related works are reviewed. Then, the network architecture model and usage scenarios are introduced in section 6.3. Section 6.4 introduces the Kalman filter and different machine learning algorithms and then presents the proposed HO scheme. Section 6.5 evaluates the performance of the proposed method in comparison with other methods in the literature. Finally, section 6.6 summarises this chapter.

6.2 Related Works

There already exist several different schemes which solve the HO problems mentioned above. Work in [132] adopts a gray model prediction based handover scheme to prevent the handover delay problem, and they used a combination factor in the handover decision-making in order to select the best target cell in heterogeneous UDN for HSR. Paper [133] provides a comprehensive overview of the current state-of-the-art and future trends for wireless technologies aiming to realize the concept of HSR communication services. Furthermore, authors of [134] adopt a dual-link soft handover algorithm by deploying a TRS and two antennas on top of train to decrease the outage probability of handover for control/user plane split network. In [135], the authors introduce a handover decision algorithm called the relative velocity-aided handover algorithm, which uses the speed of the user's device relative to the current and potential new access points to decide when to switch connections. This work considers the user's relative speed to all nearby access points and uses a speed threshold to determine if a handover should occur, aiming to reduce the chance of a connection drop. If the relative speed is lower than the threshold, the traditional handover method is used; if it's higher, the UE device avoids connecting to the AP with high relative velocity (higher than the threshold). Works in [136] [137] propose MRs as a solution to improve communication quality by reducing handover delays. Paper [138] provides a comprehensive overview of the current state-of-the-art and future trends for exploring the utilization of mmWave technology for high-speed railway communication systems. In [139] and [140], different handover schemes based on 5G mmWave mobile communication systems in the HSR scenario are proposed. Work in [139], the authors present a novel approach to improve inter-beam handover (IBH) in HSR scenarios for 5G mmWave communication systems. The authors propose a Jointly Optimized Dynamic Inter-Beam Handover (JOD-IBH) scheme, which is based on multiple beams cooperation. Simulation results show that the proposed method can decrease handover failure rates and achieve a balance between handover failure and resource occupation rates in the HSR scenario. In [140], the authors propose a Long Short-Term Memory (LSTM) based method to improve handover in 5G mmWave networks for high-speed railways. This method aims to address the high handover delay and link interruption issues of the traditional A3 handover approach. In [141], authors proposed a handover optimization method that utilizes fuzzy logic to decrease the number of HOs. Their simulation results demonstrate that the ping-pong effect has been significantly reduced. With the development of the 5G and beyond networks, AI will play a crucial role in many innovative 5G technologies and architectures in the future. The main reason is that AI techniques are increasingly being utilized in wireless communication due to their powerful learning and optimizing ability to analyze complex data patterns, make intelligent decisions, and adapt to dynamic environments [142]. Additionally, ML techniques can be used to refine the models by analyzing the behavioral traits of moving objects from historical records and predicting future movements using trained models. For handover management, it can be used to achieve intelligent mobility prediction and optimal handover solutions to guarantee seamless communication connectivity. An intelligent mobility management system based on improved multi-objective optimization method by ratio analysis and reinforcement learning approach has been discussed in [143]. In this scheme, authors first adopt the multi-objective optimization method to decrease the ranking abnormality when it selects a target BS. Then, the reinforcement learning method is used to select the optimal triggering points, aiming to reduce the number of frequent HOs and handover failure rate for better QoS. However, the authors only consider handover occurring in an intra-macro cell scenario. In [144], authors propose and evaluate a ML based handover algorithm to improve users' QoS in the presence of signal-blocking obstacles. Their simulation results demonstrate that the proposed scheme is able to choose the BS that will provide the highest QoS, even in severe channel propagation situations. Nevertheless, they do not consider the effect of velocity on handover performance. Work on [145] uses the received power of the serving and the target BS as input features for handover failure prediction based on a machine learning model. However, they do not consider the effect of interference which is severe in UDN. In [146] and [147], different machine learning approaches are applied for future mobility prediction. Machine learning based user-level handover decision making scheme in 3GPP networks is presented in [146], where the authors employ various classification algorithms, including neural networks, to address the problem of identifying the optimal HO target BS. They also utilize regression approaches to estimate the duration and percentage of download progress of a mobile user, specifically when the user is in motion and has initiated a HO request. In [147], a novel framework is introduced to predict the future network connection of a mobile node in WLAN environments. The proposed framework incorporates a decision-making process that relies on classification algorithm, as well as the user's position and acceleration data, to anticipate network change. The proposed method achieves an impressive prediction accuracy of up to 96.75% when forecasting the connection of the future network. A Bayesian Regression Based LTE-R Handover Decision Algorithm [95] for HSR is proposed to predict the cell boundary crossing time. Their proposed method maintains stronger signal strength between mobile relays and the serving BS both before and after the handover. This improvement results in a reduced likelihood of RLF. However,

they do not consider the Doppler shift effect in HSR scenario.
Despite the improvements achieved by the existing works, they may not perform efficiently in 5G HSR mmWave UDNs. This is primarily due to the challenges posed by the small coverage area and high propagation loss of the mmWave technique. Therefore, the HO challenges in 5G HSR mmWave UDNs cannot be solved by the existing methods. In order to fill in this gap, this chapter focuses on investigating handover performance enhancement using various machine learning algorithms in 5G mmWave ultra-dense network of HSR. Leveraging the characteristics of HSR scenario, the study develops a Kalman combined with SVM algorithm prediction method to find the optimal target cell. This approach takes into account a range of appropriate handover criteria and the Doppler effect to enhance handover performance and deliver satisfying QoS to train users of HSR system.

6.3 System Model

6.3.1 HSR Network Architecture

In this chapter, we consider a heterogeneous UDN deployed for high-speed rail, as depicted in Fig. 6.1. The network consists of two layers. The first layer utilizes macro BS operating at a carrier frequency of 2 GHz to ensure uninterrupted coverage. The second layer employs small BSs operating at mmWave frequency, which offers high-speed data services to passengers thanks to their abundant bandwidth resources. Thus, the mmWave technology can support a massive number of connected devices simultaneously, especially in HSR environments. However, the mmWave signal suffers from high path loss and significant propagation attenuation, resulting in limited coverage [148]. Consequently, a dense deployment of small BS is required in the HSR scenario.
To mitigate penetration loss caused by trains, a relay station is deployed on the train's roof. This relay assists passengers in connecting to the small BS or macro BS located along the trackside. Passengers can access the Internet through AP connected to the TRS via wired link. The TRS acts as a single entity to execute HO from the serving BS to the target BS, aiming to minimize handover signaling overhead and avoid the penetration loss caused by the train.

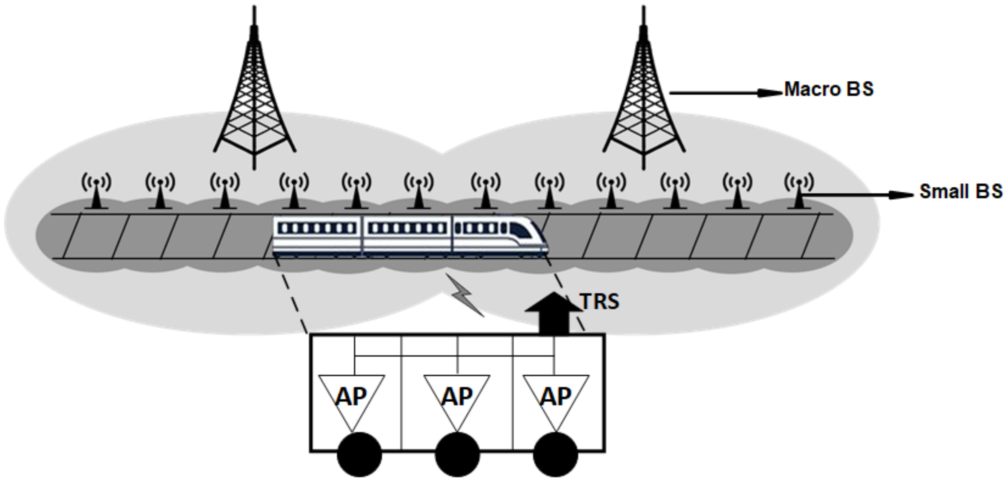

Figure 6.1: System architecture

6.3.2 Channel Model

In this chapter, the linear radio coverage topology is adopted in HSR and only the main path signal is considered since high-speed train runs through rural or viaduct areas most of the time, so multi-path effect can be ignored. At time t, let a high-speed train be located at position x in the coverage area of i-th BS. The received signal power as well as RSRP from the BS i is considered in dBm as follows:

$$RSRP(i,x)[\text{dBm}] = Pt[\text{dBm}] - PL(i,d_i)\,[\text{dB}] - \zeta\,(0,\sigma_i)\,[\text{dB}] \tag{6.1}$$

where Pt is the transmission power of the i-th BS. $PL(i,d_i)$ is the path loss. $\zeta\,(0,\sigma_i)$ is a Gaussian distributed random variable with a zero mean and a standard deviation to describe the shadow fading. Let

$$R(i,x)[\text{mw}] = 10^{RSRP(i,x)/10} \tag{6.2}$$

where R is the RSRP in mw units.

Doppler shift is another important feature we should consider for high speed rail [149]. Doppler Shift refers to the phenomenon observed in mobile communication, where the fast movement of the mobile station and the varying distance between the receiving terminal and the base station cause the synthesized frequency to fluctuate both above and below the center frequency [150]. The Doppler shift will result in ICI effect which will decrease the received power. Thus, the RSRP with the normalized noise power at the time t and position x from the i-th BS can be modified as

$$RSRP(i,x)[\text{dBm}] = 10\log_{10}\left[\frac{R(i,x)}{R(i,x)\Delta+1}\right], \tag{6.3}$$

where $\Delta = \int_{-1}^{1}(1-|x|)\left(1-J_0\left(2\pi f_d T_s x\right)\right)dx$ is the ICI power and details can be found in [108] [49] [151]. $J_0(\cdot)$ is the zeroth-order Bessel function of the first kind, f_d represents the maximum Doppler frequency shift and T_s is the symbol duration.

Pass loss models are different for macro BSs and small BSs. The COST 231 Hata model for macro BS is proposed in [93], which is applicable to HSR scenario. The simplicity and availability of accurate factors make this model widely used for radio propagation in the urban, suburban and rural areas for HSR system. Since the HSR is generally built in rural areas, we use the open rural formula to calculate the path loss which can be written as follows,

$$\begin{aligned}
\text{PL} &= A + B\log(d) + C,\\
A &= 46.3 + 33.9\log\left(f_c\right) - 13.82\log\left(h_b\right) - a\left(h_m\right),\\
B &= 44.9 - 6.55\log\left(h_b\right),\\
C &= 0,\\
a\left(h_m\right) &= \left[1.1\log\left(f_c\right) - 0.7\right]h_m - \left[1.56\log\left(f_c\right) - 0.8\right],
\end{aligned} \tag{6.4}$$

where A, B and C are decided by frequency and antenna height, and f_c is given in MHz, d is given in kilometer, h_b and h_m are given in meter.

The path loss model in the i-th small BS can be represented as

$$PL\left(i,d_i\right)[\text{dB}] = PL\left(d_0\right) + 10\theta\log\left(d_i\right), \tag{6.5}$$

where d_i is the distance between the i-th small BS and the mobile train, θ is the path loss exponent for the mmWave frequency band, and $PL(d_0) = 10\theta\log\left(4\pi d_0/\lambda\right)^2$, λ is the wave length. Path loss depends on the distance between the transmitter and receiver.

For macro BS, we consider the inter-cell interference caused by the nearest neighboring base stations with same frequency and can be denoted as

$$I_m[mW] = \sum_{n_m=1, n_m \neq m}^{N_{nei.m}} 10^{RSRP(n_m,x)/10}, m \in M \tag{6.6}$$

where $M = \{m_1, m_2, m_3 ... m_h\}$, h is the number of macro BS and $N_{nei.m}$ is the number of the nearest neighboring macro BS. *RSRP*(n_m, x) obtained from (6.3) is the received interference signal power from the co-channel macro BS.

Similarly, the interference power received from the nearest neighboring small base stations can be expressed as

$$I_s[mW] = \sum_{n_s=1, n_s \neq s}^{N_{nei.s}} 10^{RSRP(n_s,x)/10}, s \in S \tag{6.7}$$

where $S = \{s_1, s_2, s_3 ... s_p\}$, p is the number of small BS and $N_{nei.s}$ is the number of the nearest neighboring small BS. *RSRP*(n_s, x) obtained from (6.3) is the received interference signal power from the co-channel small BS.

Therefore, at location x, we can express SINR from the i_m-th macro BS and the i_s-th small BS respectively as follows:

$$SINR_{i_m}[dB] = 10\log\left[\frac{10^{RSRP(i_m,x)/10}}{I_{i_m} + N_0}\right], \tag{6.8}$$

$$SINR_{i_s}[dB] = 10\log\left[\frac{10^{RSRP(i_s,x)/10}}{I_{i_s} + N_0}\right], \tag{6.9}$$

where N_0 is the noise power, *RSRP*(i_m, x) and *RSRP*(i_s, x) obtained from (6.3) is the received signal power from the macro and the small BS respectively.

6.4 Proposed HO Scheme

In this section, we present the proposed handover scheme. This scheme comprises two modules, as illustrated in Fig. 6.2.

1) The Kalman filter based module is proposed to estimate the posteriori of RSRP for the neighboring BSs, relying on the measurement report as its primary input.

2) Following the first module, the SVM module considers the output of the Kalman filter and SINR, time criteria of neighboring BSs as its input to determine the most suitable target BS for a handover.

For the purpose of comparison, we will consider other supervised learning algorithms and the existing works [131] in the literature.

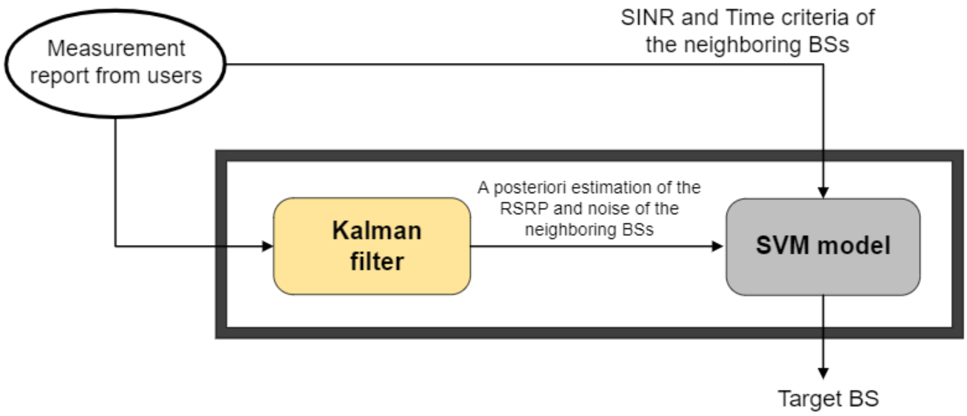

Figure 6.2: System model overview

6.4.1 Kalman Filter based RSRP Estimation

The Kalman filter [152] is a recursive mathematical algorithm used for estimating a system's next state based on only the previous state and the current measurement. It is particularly useful in some situations where noise or uncertainty exists in the measurements. The reasons for using the Kalman filter are described as follows:

- As a recursive model, the Kalman filter does not need the knowledge of the entire measurement history. It only relies on the information from the last measurement to estimate the desired output [152]. Consequently, this approach reduces response time and saves memory space.

- The Kalman filter has a strong ability to estimate system states from noisy measurements [153]. Thus, it can be employed to estimate the received signal strength of the channel and effectively monitor fluctuations in channel quality while taking noise factor into account.

- When estimating unknown variables, the utilization of the Kalman filter generally tends to be more accurate compared to a single measurement [154].

In addition, classical machine learning algorithms need a lot of historical data for training, where the estimation is based on the size and quality of the dataset. Nonetheless, the RSRP based on a long time sequence of previous measurements will not be estimated fast enough to adapt to rapid variation of the wireless environment in HSR communications systems.

Kalman filters are employed for state estimation based on linear dynamical systems in state space format. The process model defines the state's progression from time $k-1$ to time k as follows:

$$\boldsymbol{x}_k = F\boldsymbol{x}_{k-1} + \boldsymbol{w}_k \tag{6.10}$$

where F is the state transition matrix which is applied to the previous state vector $\boldsymbol{x}_{k-1}$ and $\boldsymbol{w}_k$ is the process noise vector that is assumed to be zero-mean Gaussian with the covariance Q, i.e., $\boldsymbol{w}_k \sim \mathcal{N}(0, Q_k)$.

The process model of the Kalman filter is combined with the measurement model which describes the relationship between the state and the measurement at the current time step k denoted as:

$$\boldsymbol{z}_k = H\boldsymbol{x}_k + \boldsymbol{\nu}_k \tag{6.11}$$

where $\boldsymbol{z}_k$ is the measurement or the observation vector, H is the measurement matrix, and $\boldsymbol{\nu}_k$ represents the measurement noise vector that is assumed to be zero-mean Gaussian with the covariance R, i.e., $\boldsymbol{\nu}_k \sim \mathcal{N}(0, R_k)$.

Kalman filter algorithm consists of two phases: prediction and update [152] [17] [155]. During the prediction phase, the previous state estimate is utilized to predict the state at the current time step. It's important to note that the predicted state is an estimate of the current time step, does not account for the current observed information. Consequently, the predicted state is commonly referred to as the "$a\ prior$" state. In the subsequent update phase, the current observed information is integrated with the $a\ priori$ prediction to enhance the state estimate. This refined state estimate is known as the "a posteriori" state estimate.

The Kalman filter algorithm is summarized as follows:

1) **Predict phase**

- Predicted (Prior) state estimate: $\hat{\boldsymbol{x}}_k = F\boldsymbol{x}_{k-1} + \boldsymbol{w}_k$
- Predicted (Prior) error covariance estimate: $\widehat{P}_k = FP_{k-1}F^T + Q_k$

2) **Update phase**

- Kalman gain: $K_k = \widehat{P}_k H^T \left(R_k + H\widehat{P}_k H^T\right)^{-1} = \frac{\widehat{P}_k H^T}{H\widehat{P}_k H^T + R_k}$
- Updated (Posteriori) state estimate: $\boldsymbol{x}_k = \hat{\boldsymbol{x}}_k + K_k \mathbf{z}_k - H\hat{\boldsymbol{x}}_k$
- Updated (Posteriori) error covariance: $P_k = (I - K_k H)\widehat{P}_k$

In the above equations, let $\hat{\boldsymbol{x}}_k$ and $\boldsymbol{x}_k$ be the a prior and a posteriori estimation of the state, respectively. $\widehat{P}_k$ means the a prior estimate error covariance matrix and P_k denotes the a posteriori estimate error covariance matrix. Let K represent the Kalman gain, which is the relative weight assigned to the current state estimate and measurements. K can be tuned to obtain a particular performance. For instance, a higher Kalman gain gives more weight to the measurements, while a lower Kalman gain relies more on the prediction. Thus, a higher gain leads to more frequent fluctuations in the estimation. In contrast, a lower gain tends to reduce the impact of noise, resulting in a smoother estimate but with reduced responsiveness.

The Kalman filter information flow is described in Fig. 6.3.

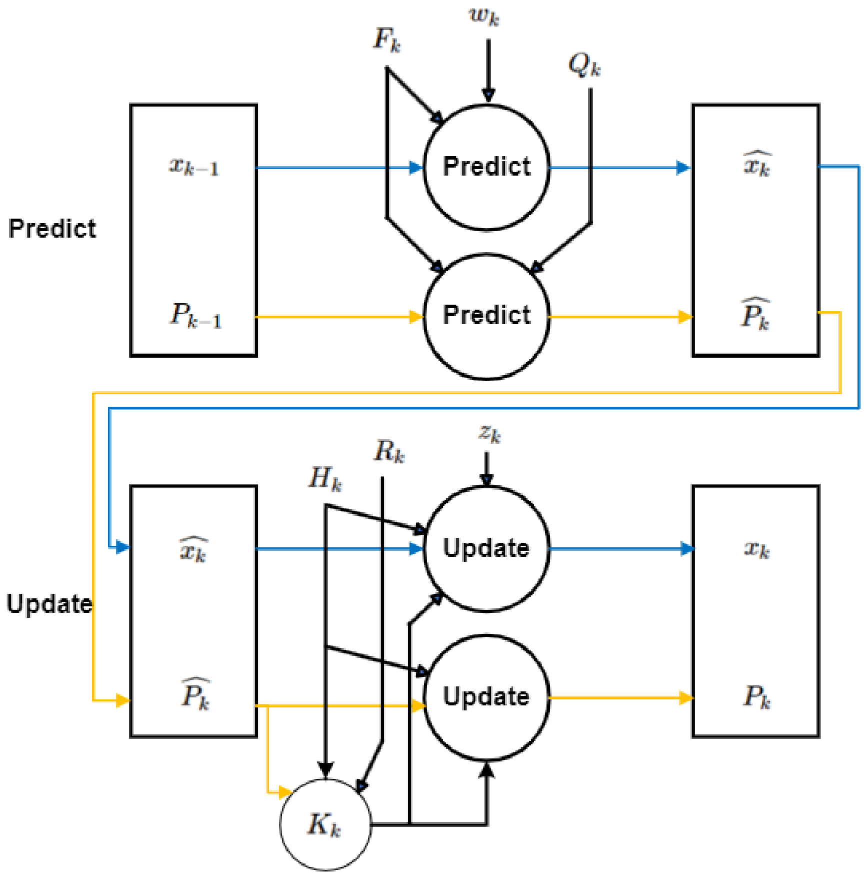

Figure 6.3: Kalman filter information flow [17]

Since the initial value of P will not have an impact on the converged value of K, we initialized P with a non-zero matrix.

Based on the introduction above, we consider $S_{RSRP,k}$ and $S_{RSRP,k-1}$ as the RSRP values measured at time k and $k-1$, respectively. We assume that $N_{e,k}$ is the environment noise that is generated from various RF devices, including power generators, WiFi and motors, etc. Additionally, the HSR wireless network is a fast time-varying system in which measurement and environment noises have a significant impact on the received signal strength of a BS. Accurate signal quality estimation is vital for handover decision making. Hence, it is necessary to enhance signal quality estimation by eliminating measurement noise. The Kalman filter's promising ability is to estimate and predict the states of a dynamic system based on noisy measurements. It is particularly effective in filtering out noise and providing accurate estimates of the system states. Therefore, we can design a Kalman filter to estimate the signal strength of the channel and accordingly track the changes in channel quality, accounting for the influence of the noise. In our model, the Kalman filter estimates the variation of RSRP (S_{RSRP}) and environment noise (N_e) by filtering the value of the measurement noise. Therefore, the system can be denoted as follows,

$$\begin{aligned} S_{RSRP,k} &= S_{RSRP,k-1} + w_{s,k} \\ N_{e,k} &= N_{e,k-1} + w_{e,k} \end{aligned} \tag{6.12}$$

where $w_{s,k}$ is the effect of path loss and shadow fading on the signal strength RSRP and $w_{e,k}$ represents the environment noise. Moreover, we assume that $w_{s,k}$ and $w_{e,k}$ are independent of each other and they both follow Gaussian distributions with zero mean. The values of environment noise and RSRP are correlated because the RSRP values can be reduced as the effect of environment noise increases and vice versa. As a result, the likelihood of the measured RSRP values is related to the environment noise. At time k, the correlation between the RSRP and environment noise is described by a covariance matrix denoted by P_k. Equation (6.12) can be expressed in a vector representation, which results in a state evolution formula as follows,

$$\boldsymbol{x}_k = F\boldsymbol{x}_{k-1} + \boldsymbol{w}_k \tag{6.13}$$

Where, $\boldsymbol{x}_k = \begin{bmatrix} S_{RSRP,k} \\ N_{e,k} \end{bmatrix}$ is the estimated values of the system states. Our aim is to predict the next state at time k according to the current state at time $k-1$. F is the prediction matrix and $F = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. It is used to estimate the next state of the dynamic system. In a wireless signal model, we can use a radio frequency integrated circuit (RFIC) to measure the state variable, and this variable is relevant to the measurement noise. Furthermore, $\boldsymbol{w}_k = \begin{bmatrix} w_{s,k} \\ w_{e,k} \end{bmatrix}$. Based on the introduction of Kalman filter above, let Q_k be the covariance of the zero mean Gaussian distribution followed by $\boldsymbol{w}_k$ and consequently, $\boldsymbol{w}_k \sim \mathcal{N}(0, Q_k)$. Besides, in terms of the observation (measurement) equation, it can be represented as

$$\boldsymbol{z}_k = H\boldsymbol{x}_k + \boldsymbol{\nu}_k \tag{6.14}$$

where $\boldsymbol{z}_k$ is the measured or observed value of the system state. H is the measurement matrix and $H = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. Further, $\boldsymbol{\nu}_k$ represents the observed noise or measured noise, assumed to follow a zero mean Gaussian distribution with covariance R_k, and thus $\boldsymbol{\nu}_k \sim \mathcal{N}(0, R_k)$.

In our proposed scheme, the state value $\boldsymbol{x}_k$ is used in the next module as shown in Fig. 6.2, where the handover decision will be made utilising the SVM model. This learning algorithm and details of the handover mechanism will be discussed in the following section.

6.4.2 SVM based HO

This section first provides a brief introduction to the SVM algorithm and then the proposed SVM based HO scheme will be introduced.

SVM is a popular algorithm in supervised learning, commonly employed for both classification and regression tasks. However, its predominant usage lies in classification problems. In addition, SVM is not limited to linear classification tasks, it can also effectively handle non-linear classification through the use of the kernel trick. By doing so, SVM can implicitly map the inputs into high-dimensional feature spaces, thereby enabling efficient non-linear classification. The main objective of the SVM algorithm is to construct an optimal line or decision boundary, known as a hyperplane, capable of effectively separating different classes within an n-dimensional space. This facilitates accurate categorization of future data points into their appropriate categories.

SVM identifies crucial points or vectors, known as support vectors, which play a vital role in defining the hyperplane. Support vectors are data points that are closer to the hyperplane and significantly impact its position and orientation. By leveraging these support vectors, we aim to maximize the margin of the classifier. A margin is a separation gap between the two lines on the closest data points. It is calculated as the distance from the hyperplane to support vectors. The relationship between hyperplane, support vectors and margin is shown in Fig. 6.4. The basic principle of SVM is to find a decision hyperplane among linearly separable sets. The expression of the hyperplane can be shown as follows:

$$f(\mathbf{x}) = \langle \boldsymbol{\omega}, \mathbf{x} \rangle + b \tag{6.15}$$

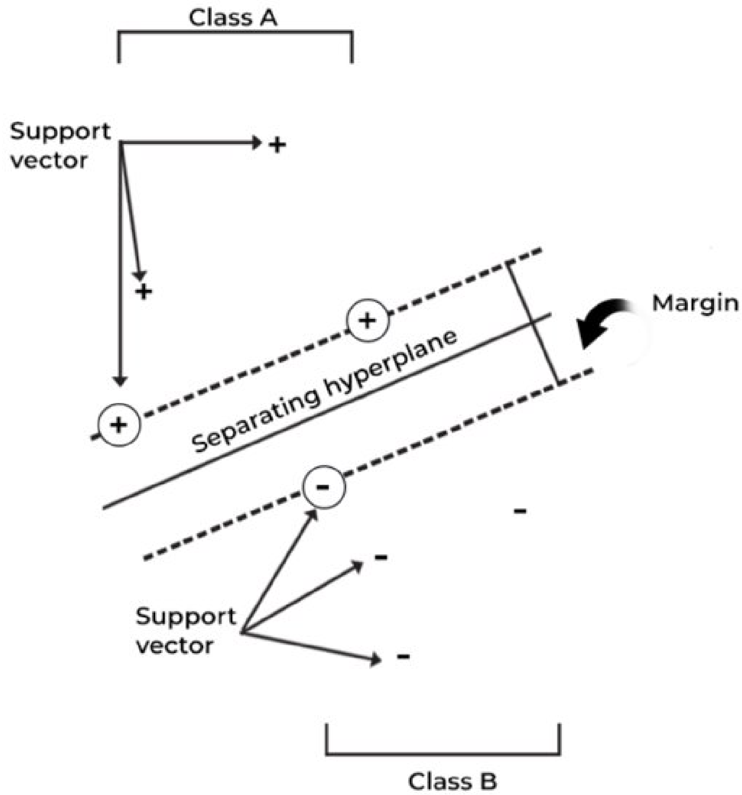

Figure 6.4: SVM optimizes margin between classes or support vector [18]

where $\mathbf{x}$ is the feature input, $\boldsymbol{\omega}$ is the weight matrix and b is a bias term.

The cost function is given by the following formula:

$$J = \min \frac{1}{2}\|\boldsymbol{\omega}\|^2 + C\sum_{i=1}^{N}\mu_i \tag{6.16}$$

where $\mu_i \geq 0$ and $\|\boldsymbol{\omega}\|^2$ is the Euclidean distance from the original to the hyperplane. Moreover, N is the number of training set, C is a parameter for regularization which can be used to control the number of errors that is permitted on the training dataset.

The prediction model in equation (6.15) can be expressed as a linear combination of the training data $\mathbf{x}_i$ [156]. It can be represented as follows:

$$f(\mathbf{x}) = \sum_{i=1}^{N}\left(\alpha_i - \alpha_i^*\right) K\left(\mathbf{x}_i, \mathbf{x}_j\right) + b \tag{6.17}$$

where $\alpha_i, \alpha_i^* \in [0, C]$ and $K(\mathbf{x}_i, \mathbf{x}_j)$ is the kernel function. In our proposed scheme, we utilize the radial basis function (RBF) as the kernel function:

$$K(\mathbf{x}_i, \mathbf{x}_j) = exp\left(-\frac{\left\|\mathbf{x}_i - \mathbf{x}_j\right\|^2}{2\sigma^2}\right) \tag{6.18}$$

where σ is the Gaussian kernel width. It implicitly determines the data's distribution after being mapped to the new feature space. $\|\mathbf{x}_i - \mathbf{x}_j\|^2$ represents the squared Euclidean distance between vectors $\mathbf{x}_i$ and $\mathbf{x}_j$ in the original feature space. An equivalent definition involves a parameter $\gamma = \frac{1}{2\sigma^2}$:

$$K(\mathbf{x}_i, \mathbf{x}_j) = exp\left(-\gamma \left\|\mathbf{x}_i - \mathbf{x}_j\right\|^2\right) \tag{6.19}$$

where γ is the kernel coefficient for kernel function. Without loss of generality, let $\gamma = \frac{1}{2}$. We can obtain the following equations using the multinomial theorem [157]:

$$\begin{aligned}
& exp\left(-\frac{1}{2}\left\|\mathbf{x}_i - \mathbf{x}_j\right\|^2\right) \\
& = exp\left(-\frac{1}{2}\|\mathbf{x}_i\|^2 + \mathbf{x}_i^\top \mathbf{x}_j - \frac{1}{2}\left\|\mathbf{x}_j\right\|^2\right) \\
& = exp\left(-\frac{1}{2}\|\mathbf{x}_i\|^2\right) exp\left(\mathbf{x}_i^\top \mathbf{x}_j\right) exp\left(-\frac{1}{2}\left\|\mathbf{x}_j\right\|^2\right) \\
& = \sum_{m=0}^{\infty} \frac{\left(\mathbf{x}_i^\top \mathbf{x}_j\right)^m}{m!} exp\left(-\frac{1}{2}\|\mathbf{x}_i\|^2\right) exp\left(-\frac{1}{2}\left\|\mathbf{x}_j\right\|^2\right) \\
& = \sum_{m=0}^{\infty} \sum_{n_1+n_2+\cdots+n_k=m} exp\left(-\frac{1}{2}\|\mathbf{x}_i\|^2\right) \frac{x_{i_1}^{n_1} \cdots x_{i_k}^{n_k}}{\sqrt{n_1! \cdots n_k!}} \\
& exp\left(-\frac{1}{2}\left\|\mathbf{x}_j\right\|^2\right) \frac{x_{j_1}^{n_1} \cdots x_{j_k}^{n_k}}{\sqrt{n_1! \cdots n_k!}} \\
& = \left\langle \varphi(\mathbf{x}_i), \varphi\left(\mathbf{x}_j\right) \right\rangle
\end{aligned} \tag{6.20}$$

where is a mathematical function that takes two input data points in the original feature space and computes their similarity or inner product in the transformed space. It is particularly useful when dealing with non-linearly separable data in high-dimensional feature spaces. Thus, the kernel function enables the transformation of input data into a higher-dimensional space where it might become linearly separable. Mathematically, given two data points represented as vectors $\mathbf{x}_i$ and $\mathbf{x}_j$, the kernel function is denoted as

$K(\mathbf{x}_i, \mathbf{x}_j)$. The kernel function can be thought of as a measure of how similar or related the data points are. In other words, the kernel function can implicitly represent the high-dimensional feature space without explicitly computing the transformed data points.

After the brief introduction of the SVM algorithm, the introduction to the proposed SVM based HO scheme will be divided into two parts: (A) Data Preparation And Process; (B) The SVM-based Predictive Handover Scheme.

A. Data Preparation And Process

We jointly consider three major criteria for the HO decision making process, i.e., RSRP, SINR and travel time (T).

1) **RSRP**: RSRP refers to reference signal receive power and measures the power of the reference signal received at the antenna of the mobile device. It indicates the radio link quality between the mobile device and the serving cell. This criterion is from the output of the Kalman filter module.

2) **SINR**: SINR often serves as an indicator of the interference level on the radio link and evaluates the connection's overall quality. In the proposed scheme, this parameter helps to select a BS that provides higher SINR value, and in turn, decreases the HO failure rate.

3) **Travel time**: The travel time (T) is defined as how long the train takes to reach the far edge of a BS coverage area. In other words, it is the time that the train takes to arrive at the intersection point of the railway track and BS coverage boundary as shown in Fig. 6.5. This metric is utilized to reduce the HO number. In Fig. 6.5, we assume that the circle drawn is the communication range that a BS can cover.

Then, a multiple criteria matrix is formed and can be used as model input to predict an optimal BS for HO. The dataset is expressed in the equation (6.21), where m is the number of samples (waypoints), X_m denotes the inputs of the sample m and y_m stands for the outputs of the sample m as well as the different classes in the SVM

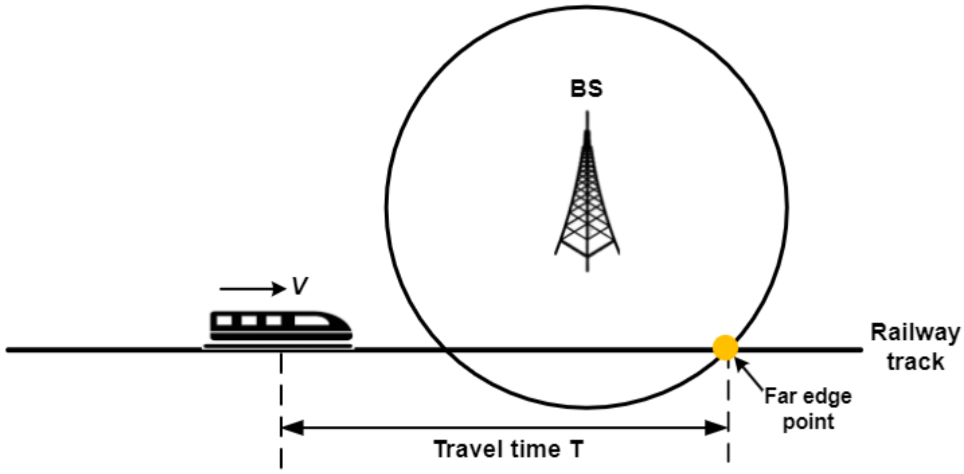

Figure 6.5: Travel time diagram

algorithm.

$$D = \{(X_1, y_1), (X_2, y_2), ..., (X_m, y_m)\} \tag{6.21}$$

The inputs contain features including the RSRP, SINR, and T in the network. Furthermore, in addition to an input matrix X, the corresponding output matrix y is also generated and the elements in matrix y are the chosen BS index for each sample (each waypoint). Here, we use index p_i={1,2,3,4,5} to describe the 5 different classes in the output matrix which are the next 5 neighboring BSs at each waypoint.

For the generation of the output matrix, we use Multiple Criteria Decision Making technique (MCDM) method [158] to rank the 5 candidate BSs (5 classes) for each sample (waypoint) and select the highest ranked index which corresponds to the best/optimum BS as the output. MCDM methods aim to provide a systematic framework for evaluating and ranking alternative solutions in decision-making scenarios where there are multiple criteria to consider. In addition, MCDM approaches utilize mathematical and computational models to assist decision-makers in selecting the most suitable option from a set of alternatives. In this chapter, we adopt the Simple Additive Weighting (SAW) method which is one of the MCDM techniques that is based on the weighted average. SAW assigns weights to each criterion based on their perceived importance, and the alternatives are evaluated by summing the normalized scores across all criteria. The mathematical representation of the SAW method can be expressed as follows:

Given p alternatives (candidate BSs) and u criteria, let $X = [x_{ij}]$ be the decision matrix, where x_{ij} represents the performance of alternative i on criterion j. The weight vector $W = [w_1, w_2, \ldots, w_u]$ represents the importance weights assigned to each criterion.

The normalized decision matrix $V = \left[r_{ij}\right]$ is obtained by normalizing each criterion's values, ensuring that different scales do not bias the results. For benefit criteria, higher values are preferred, as they represent better performance or higher desirability. The normalization formula for a benefit criterion x_{ij} is given by:

$$r_{ij} = \frac{x_{ij}}{\max\left(X_j\right)} \tag{6.22}$$

Where X_j is the vector of all values for criterion j, and $\max\left(X_j\right)$ is the maximum value in that vector.

For cost criteria, lower values are preferred, as they represent lower costs. The normalization formula for a cost criterion x_{ij} is given by:

$$r_{ij} = \frac{\min\left(X_j\right)}{x_{ij}} \tag{6.23}$$

Where X_j is the vector of all values for criterion j, and $\min\left(X_j\right)$ is the minimum value in that vector.

The overall score for each alternative i is then calculated by summing the weighted normalized scores:

$$S_i = \sum_{j=1}^{u} w_j \cdot r_{ij} \tag{6.24}$$

The alternative with the highest overall score is considered the most favorable. In our work, we use R to describe the overall score. In order to select the optimum BS for each sample, the BS index with the highest R will be picked as the output of each sample. R can be calculated by the following formula,

$$R = \omega 1 \times RSRP + \omega 2 \times SINR + \omega 3 \times T \tag{6.25}$$

where $\omega 1$, $\omega 2$ and $\omega 3$ represent the weights for the RSRP, SINR, and T respectively. In addition, their range is from [0,1] and $\omega 1 + \omega 2 + \omega 3 = 1$. More specifically, there are 5 classes (5 candidate BSs) to be selected in our model. In terms of the dataset, it includes the input and the output of each sample which are different features (RSRP, SINR, T), and the selected BS index for each sample based on MCDM method.

The selection of weights also has a crucial impact on selecting the candidate BS as the output of each sample. The application of the MCDM method in BS selection relies on the approach used to calculate the weighted network priority criteria. Thus, in this chapter, we adopt the standard deviation method [159] to calculate the weight of each individual criterion as shown in the following equations,

$$\delta_k = \sqrt{\frac{1}{m}\sum_{j=1}^{m}\left(\alpha_{jk} - N_k\right)^2}, \tag{6.26}$$

where m is the number of candidate alternatives and $k = 1, 2, \ldots, s$ and s is the number of criteria. Further, $N_k = \sum_{j=1}^{m} \alpha_{jk}/m$ means the average score. Then, we can obtain the objective weight φ_k of k-th criteria which can be denoted as follows,

$$\varphi_k = \frac{\delta_k}{\sum_{k=1}^{s}\delta_k}, \tag{6.27}$$

If the SD value is larger, the criteria should be assigned a significant weight value as they have a substantial impact on decision making.

B. The SVM-based Predictive Handover Scheme

The proposed SVM based handover prediction method to improve mobility management is presented in this section. In the beginning, a flowchart of SVM algorithm is presented in Fig. 6.6.

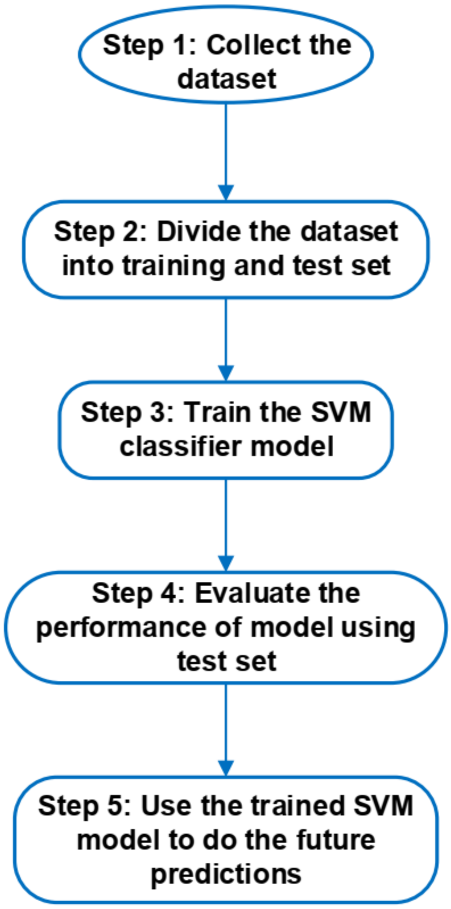

Figure 6.6: The SVM algorithm flowchart

where the SVM mechanism can be described as follows:

Step 1: Data collection: Collect labeled training dataset, where each data point is associated with a class label. Additionally, normalize or standardize the features to bring them to a similar scale.

Step 2: Train/test set split: Split the dataset into training (80%) and testing sets (20%) to evaluate the model's performance.

Step 3: Training the SVM: Input the training data into the SVM algorithm. SVM can find the hyperplane that separates the data into different classes while maximizing the margin between the classes.

Step 4: Testing: Use the testing set to evaluate the performance of the trained SVM model.

Step 5: Prediction: Once satisfied with the model's performance, use it to make predictions on new data.

In this work, handover prediction is considered as a classification problem. The proposed HO scheme is summarized in Algorithm 6.1. The overall scheme comprises

Algorithm 6.1 Proposed HO Algorithm

Initialization:
Input: Dataset
Output: Target BS
Dataset are preprocessed to extract the various features such as RSRP, SINR and time T
The dataset is divided into training set (80%) and testing set (20%)
SVM takes these features as input. Then, SVM classifier is used to choose the best BS based on the values of the criteria. (Refer to Fig. 6.7)
End

two phases: the training phase and the prediction phase. In the training phase, the input to the model is the generated dataset including the different features (RSRP, SINR, T) along with the whole railway track and their labeled next selected BS index $\{p_i\}$. Here, $\{p_i\}$ is the corresponding class which the learning algorithm aims to match the input features of each sample. After the training phase, the prediction phase is to apply the trained classification model in the decision making process to predict the next BS based on future input features. In addition, the SVM classification model was trained using the 'fitcecoc' function, which employs a class error-correcting output codes approach. This function provides a fully trained multi-class error-correcting output model. By utilizing the One-vs-One coding design, 'fitcecoc' creates K(K-1)/2 binary SVM models, where K represents the number of unique class labels. The implementation of error-correcting output codes and the one-vs-one coding design in SVM enhances the performance of the classification model [160].

Thus, the flowchart of the proposed HO scheme can be illustrated in Fig. 6.7.

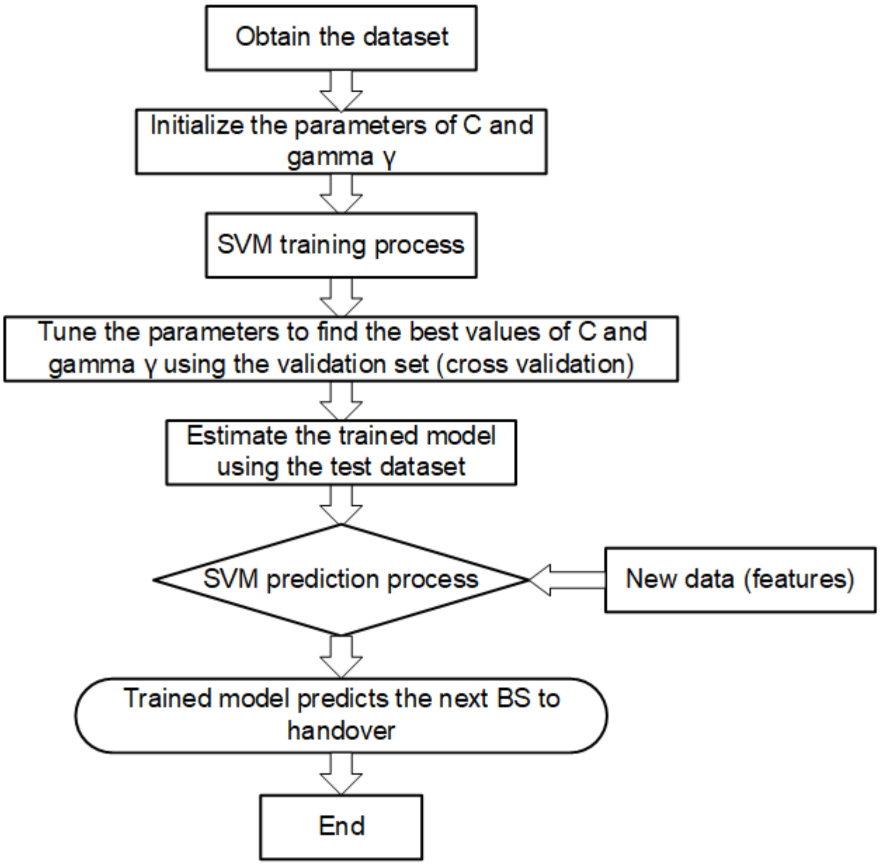

Figure 6.7: The SVM algorithm for HO situation

The objective of the SVM is to find the decision boundary between different classes by constructing a hyperplane in a high-dimensional space [161]. This hyperplane is positioned to have the maximum distance from the nearest training samples of each class. Furthermore, the learning capability and accuracy of SVM according to the chosen parameters and the kernel function applied. The selection of an appropriate kernel function significantly impacts the performance of SVM [162] [163]. In our experiments, four kernels are used: linear, polynomial, sigmoid and RBF. We found that the RBF kernel can achieve the highest accuracy, which maps sample vectors into a higher dimensional space. The RBF kernel parameter γ and the SVM slack variable C are optimized by using K-fold cross-validation [144] on a subset of the training set which is described in Fig. 6.8. This enables us to determine the most suitable values for γ and C that maximize the performance of SVM.

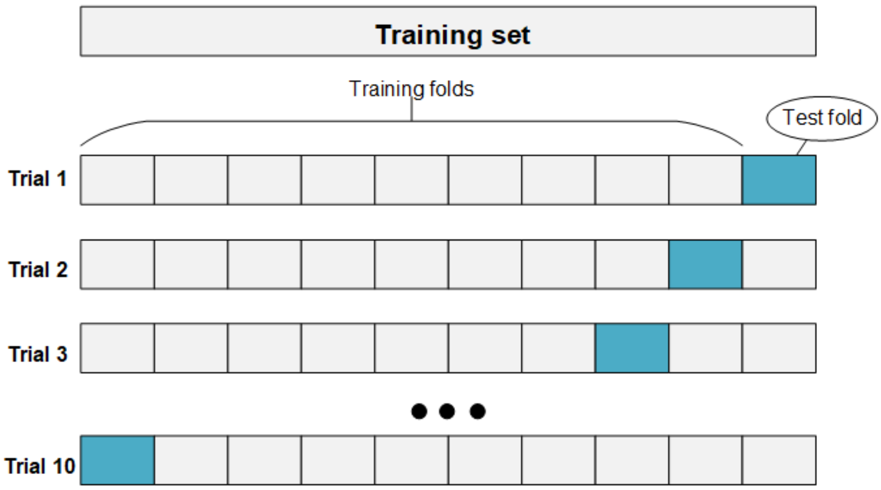

Figure 6.8: K-fold cross-validation for K=10

In the k-fold validation, the training set is divided into k parts. Then, the learning process involving k trials will be performed. During each trial, one of the k parts is designated as the test set which is named as the validation set, while the remaining k-1 parts are utilized for training. Finally, the model structure that achieves the highest average score on the validation sets across the k trials is selected. It is particularly helpful when dealing with limited data to avoid overfitting, which occurs when a model performs well on the training data but struggles to work well on new or unseen data.

In the proposed scheme, the significant steps are the generation of the dataset and the training procedure. Once the trained SVM model is obtained, it can be promptly used in the decision making process to determine the next target BS to handover.

6.5 Performance Evaluation and Simulation Results

This section evaluates the performance of the proposed machine learning based HO scheme.

6.5.1 Experiment Setup

In order to evaluate the handover performance of the proposed algorithm in the HSR system, we firstly set up a HSR communication network model in MATLAB. In this model, we assume that the overall length of the railway track is 15 km and the network consists of 4 macro BSs and 77 small cells, which are distributed linearly along one side of the track as shown in Fig. 6.1. A trained classification model is built as explained in section 6.4.2 during the training phase in the SVM algorithm. After the training phase, the model we obtained can be used to predict (classify) the next BS in HO decision making. Here, we have five different classes (5 neighboring BSs) for the train to handover at position x.

In the simulation, we collect a dataset of size 5000 samples, from which 4000 are used for training purpose (80%), and 1000 samples for test purpose (20%).

6.5.2 Other Machine Learning Algorithms for Comparison

1) **Decision Tree**

   Decision trees [19] [164] [165] stand as a prominent and versatile class of supervised machine learning algorithms that have been widely applied across diverse domains, ranging from classification to regression tasks. Rooted in the realm of decision theory, these algorithms are particularly adept at modeling complex decision-making processes by simulating a tree-like flowchart with a hierarchical structure of nodes and branches.

   A decision tree is a hierarchical structure comprising a root node, branches, internal nodes and leaf nodes that guide the traversal of data to arrive at conclusive decisions or predictions. Each internal node represents a decision point, and each leaf node signifies an outcome as shown in Fig. 6.9.

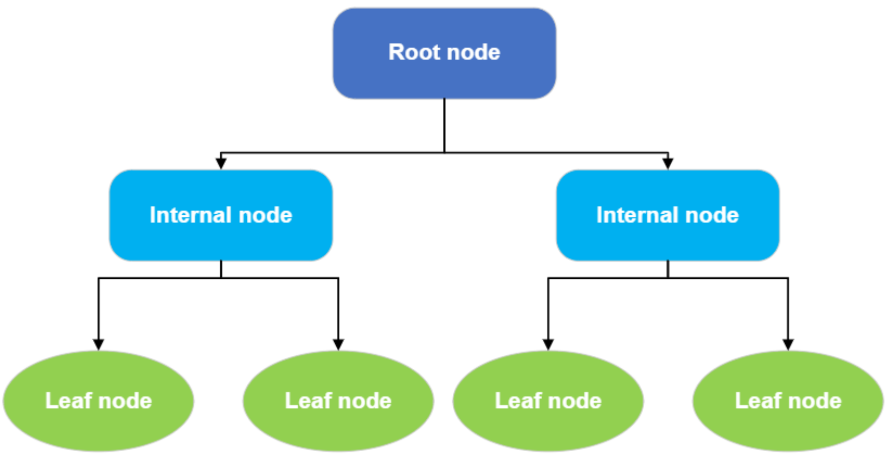

Figure 6.9: Decision tree structure [19]

As shown in Figure 6.9, a decision tree starts with a root node, which does not have any incoming branches. The outgoing branches from the root node then feed into the internal nodes, also known as decision nodes. Based on the available features, both node types conduct evaluations to form homogeneous subsets, which are denoted by leaf nodes, or terminal nodes. The leaf nodes represent all the possible outcomes within the dataset.

Decision tree learning employs a divide-and-conquer strategy by conducting a greedy search to identify the optimal split points within a tree. This process of splitting is then repeated in a top-down, recursive manner until all, or the majority of records have been classified under specific class labels.

There are many popular types of decision tree algorithms, such as ID3, C4.5 and CART.

- **ID3:** Iterative Dichotomiser 3 is an algorithm invented by Ross Quinlan used to generate a decision tree from a dataset. This algorithm leverages entropy and information gain as metrics to evaluate candidate splits [166] [167].

- **C4.5:** This algorithm is considered a later iteration of ID3, which was also developed by Quinlan. It can use information gain or gain ratios to evaluate split points within the decision trees. [167]

- **CART:** The term, CART, is an abbreviation for "classification and regression trees" and was introduced by Leo Breiman. This algorithm typically utilizes Gini impurity to identify the ideal attribute to split on. Gini impurity measures how often a randomly chosen attribute is misclassified. When evaluating using Gini impurity, a lower value is more ideal.

There are multiple ways to select the best attribute at each node including information gain and Gini impurity which act as popular splitting criteria for decision tree models. They help to evaluate the quality of each test condition and how well it will be able to classify samples into a class.

– Entropy

  It is a fundamental concept that stems from information theory and quantifies the impurity in a dataset. In the context of decision trees, entropy is used to measure the uncertainty associated with a particular set of data which can be denoted as

$$\text{Entropy}\,(S) = -\sum_{c \in C} p(c) \log_2 p(c) \tag{6.28}$$

where S represents the data set that entropy is calculated. c represents the classes in the set S. $p(c)$ is the proportion of data points that belong to class c to the number of total data points in the set S. Entropy values are typically in the range between 0 and 1. If all samples in the dataset S belong to one class, then entropy will equal zero. If half of the samples are classified as one class and the other half are in another class, entropy will be at its highest at 1. When there are more than two classes, the entropy value can be greater than 1. This occurs when there is a higher degree of impurity due to a more even distribution of data points among multiple classes. To select the best feature to split on and find the optimal decision tree, the attribute with the smallest amount of entropy should be used.

– Information gain

Information Gain is a vital metric used in decision tree construction, quantifying the difference in entropy before and after a split on a given attribute. The attribute with the highest information gain will produce the best split. Mathematically, information gain for an attribute A is calculated as

$$\text{Information Gain}\,(S, A) = \text{Entropy}(S) - \sum_{\text{v}\in\text{values(A)}} \frac{|S_v|}{|S|} \text{Entropy}\,(S_v) \tag{6.29}$$

where A represents a specific attribute or class label. Entropy(S) is the entropy of the dataset S. In addition, S_v is the subset of data for which attribute A has value v, and values (A) represent the possible values of attribute A. At last, Entropy(S_v) is the entropy of the dataset S_v.

– Gini impurity

It quantifies the likelihood of incorrectly classifying a data point in a dataset. The Gini impurity (Gini(S)) for a set S with respect to classes $c_1, c_2, \ldots, c_k$ is calculated as follows:

$$\mathrm{Gini}(S) = 1 - \sum_{i=1}^{k} \left(p\left(c_i\right)\right)^2$$

where $p(c_i)$ is the proportion of data points belonging to class c_i in set S.

The Gini impurity ranges from 0 to 1, where 0 indicates perfect purity (all data points belong to the same class) and 1 indicates maximum impurity (data points are evenly distributed among different classes). In the context of decision tree construction, the objective is to minimize the Gini impurity by selecting attributes that lead to more homogeneous subsets when the data is split.

2) **K Nearest Neighbor**

The KNN algorithm [168] [169] is a supervised machine learning method and provides a flexible framework for classification and regression tasks across diverse domains. It operates by identifying the K nearest neighbors to a given data point based on a distance metric, such as the Euclidean distance. Subsequently, the class of the data point is determined using majority voting or the average of the K neighbors. This approach enables the algorithm to effectively adapt to different patterns and make accurate predictions based on the local structure of the data.

The Euclidean distance can be calculated as

$$d\left(x, x_i\right) = \sqrt{\sum_{j=1}^{n} \left(x_j - x_{i_j}\right)^2} \tag{6.30}$$

where $d(x, x_i)$ represents the Euclidean distance between two points x and x_i in the n-dimensional space. Moreover, $\sum_{j=1}^{n}$ represents the summation over all dimensions from 1 to n. Furthermore, $\left(x_j - x_{i_j}\right)^2$ calculates the squared difference between the corresponding coordinates of the two points in each dimension. For each dimension j, it computes the squared difference between the j-th coordinate of x and the j-th coordinate of x_i, squares it, sums up these squared differences across all dimensions and then takes the square root of the sum. Thus, the result is the Euclidean distance between the two points in the n-dimensional space.

Additionally, the determination of the parameter K holds significant importance in the KNN algorithm since it decides the quantity of neighboring data points that are considered in the algorithm. The selection of the K value should be based on the input data. If the input data has more outliers or noise, a higher value of K will be better to choose. Moreover, it is recommended to choose an odd value for K to avoid ties in classification. Cross-validation methods can be used to help in selecting the best K value for the given dataset.

The KNN algorithm operates on the principle of similarity, where it predicts the class of a new data point by considering the classes of its K nearest neighbors in the training dataset as shown in Fig. 6.10.

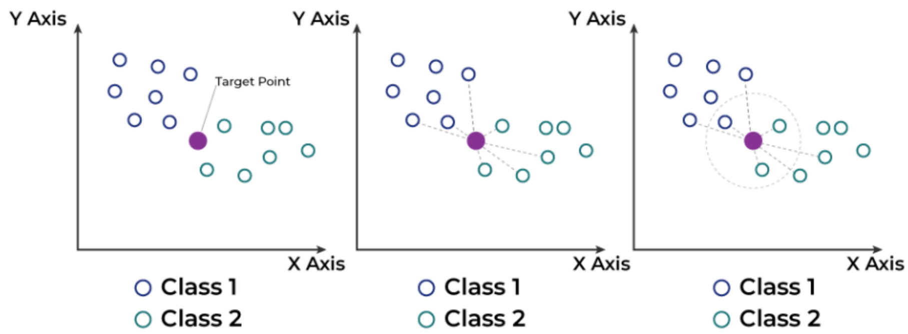

Figure 6.10: KNN algorithm example diagram [20]

Step-by-Step explanation of how KNN works is discussed as follows:

Step 1: Selecting the optimal value of K

K represents the number of nearest neighbors that need to be considered while making the prediction.

Step 2: Calculating distance

Euclidean distance is used to assess the similarity between the target and training data points. The calculation involves determining the distance between the target point and each data point in the dataset.

Step 3: Finding Nearest Neighbors

The K data points with the smallest distances to the target point are the nearest neighbors.

Step 4: Voting for Classification

Each of the k-nearest neighbors“vote” for the class label they belong to. The class labels are determined based on the majority class among these neighbors. The class label that receives the most votes from the k-nearest neighbors is assigned as the predicted class for the new data point.

3) **Artificial Neural Networks**

ANN algorithm refers to a class of machine learning algorithms inspired by the structure and functioning of biological neural networks [170] [171], particularly the human brain. ANN consists of interconnected nodes, often called neurons, organized into layers. These layers include an input layer, one or more hidden layers, and an output layer as shown in Fig. 6.11. The connections between neurons are represented by weights, and each connection is associated with a numerical weight that is adjusted during the learning process.

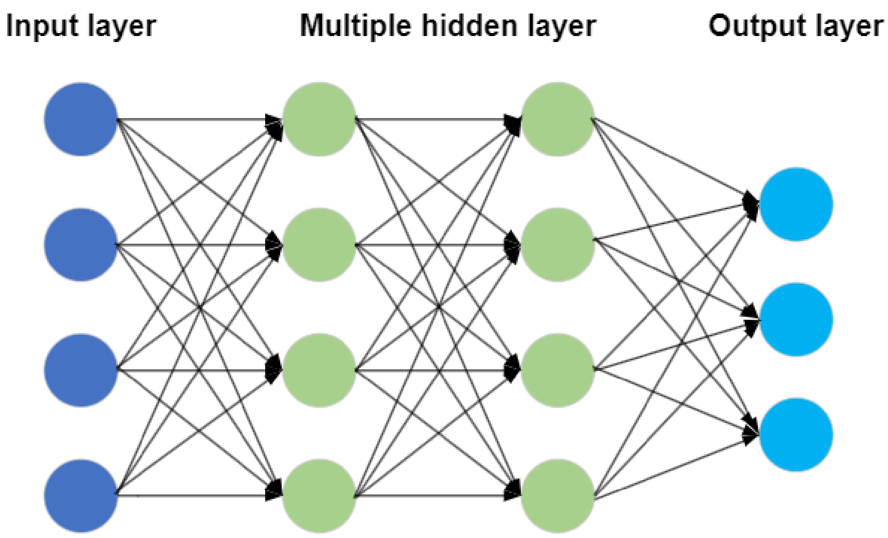

Figure 6.11: The structure of artificial neural network algorithm

Once an input layer is determined, weights are assigned. These weights help determine the importance of any given variable, with larger ones contributing more significantly to the output compared to other inputs. All inputs are then multiplied by their respective weights and then summed. Afterwards, the output is passed through an activation function, which determines the output. The weighted sum of inputs for a neuron (z) is calculated as the dot product of input vector X and weight vector W, including a bias term (b) :

$$z = X \cdot W + b \tag{6.31}$$

Then, the output (a) of the neuron is obtained by applying an activation function (f) to the weighted sum:

$$a = f(z) \tag{6.32}$$

Usually, the activation Function has various types including the sigmoid function (σ), hyperbolic tangent function (tanh), and rectified linear unit function (ReLU):

$$\begin{aligned}
&\sigma(z) = \frac{1}{1+e^{z}} \\
&\tanh(z) = \frac{e^{z}-e^{-z}}{e^{z}+e^{z}} \\
&\operatorname{ReLU}(z) = \max(0, z)
\end{aligned} \tag{6.33}$$

For example, when considering a simple one-hidden-layer neural network, the output (y) is computed as follows:

$$\begin{aligned}
&z_1 = X \cdot W_1 + b_1 \\
&a_1 = f(z_1) \\
&z_2 = a_1 \cdot W_2 + b_2 \\
&y = f(z_2)
\end{aligned} \tag{6.34}$$

In addition, the loss function (L) measures the difference between the predicted output (y) and the true target values (y_{true}):

$$L\left(y, y_{\text{true}}\right) \tag{6.35}$$

Moreover, the backward propagation algorithm is used to calculate the gradient of the loss to the weights and biases and use this information to update the weights and biases which can be denoted as

$$\begin{aligned}
&\frac{\partial L}{\partial W} = \frac{\partial L}{\partial y} \cdot \frac{\partial y}{\partial z_2} \cdot \frac{\partial z_2}{\partial W} \\
&\frac{\partial L}{\partial b} = \frac{\partial L}{\partial y} \cdot \frac{\partial y}{\partial z_2} \cdot \frac{\partial z_2}{\partial b} \\
&W_{\text{new}} = W_{\text{old}} - \alpha \cdot \frac{\partial L}{\partial W} \\
&b_{\text{new}} = b_{\text{old}} - \alpha \cdot \frac{\partial L}{\partial b}
\end{aligned} \tag{6.36}$$

where α is the learning rate.

In our work, the neural network is formed using 15 input neurons and the layer at the output has 5 output neurons. Forward and backward propagation algorithms are utilized to implement ANN. It is trained using the same dataset created for SVM and other machine learning algorithms.

6.5.3 Simulation Results of Handover Performance

The performance of the proposed HO scheme is evaluated in terms of the average number of HOs, HO failure probability, user mean throughput and handover latency.

- Handover failure probability:

  This metric is used to quantify the likelihood of an unsuccessful handover. Handover failure occurs if the received power of users is less than a threshold after a handover is triggered.

- User mean throughput:

  It refers to the effective data rate of a network. It measures how much data can be transmitted from one point to another in a specific amount of time (usually measured in seconds). The throughput can be expressed as:

$$Throughput = B * \log_2(1 + SINR)(\text{ bits/s}) \tag{6.37}$$

  where B is the bandwidth of the channel. Because SINR is one of the criteria we used, selected BS can provide better SINR, which improves the throughput.

- Handover latency:

  The handover latency refers to the duration between receiving (or transmitting) the last packet in the old cell connection and the first packet in the new cell connection. This latency is influenced by the handover initialization, decision and execution phases of the handover process.

The parameters used in the simulation are listed in Table 6.1.

Table 6.1: Simulation parameters

Parameters	Value
Bandwidth (B)	B_{macro} = 20 MHz B_{small} = 500 MHz
Carrier Frequency (F)	F_{macro} = 2 GHz F_{small} = 28 GHz
Transmit Power (P_t)	$P_{t_{macro}}$ = 46 dBm $P_{t_{small}}$ = 23 dBm
Number of BS (N)	$N_{macro} = 4$ $N_{small} = 77$
Noise Power Spectral Density	-174 dBm/Hz
Distance from BS to Rail (D)	D_{macro} = 100 m D_{small} = 2 m
Height of BSs(h)	h_{macro} = 30 m h_{small} = 7 m
Shadowing Standard Deviation (SD)	SD_{macro} = 4 dB SD_{small} = 8 dB

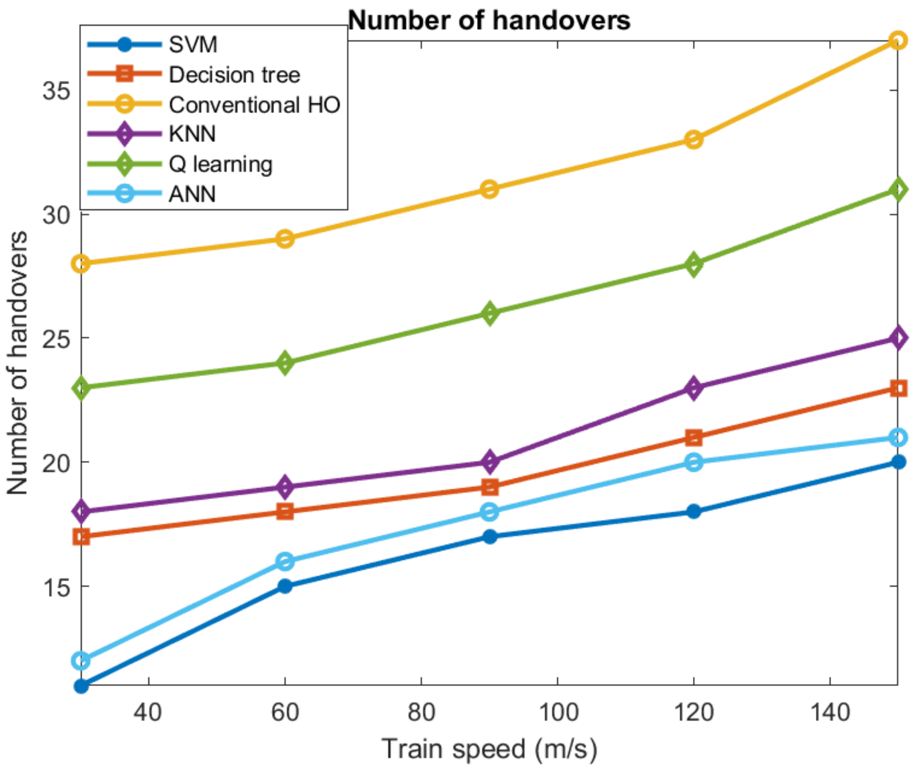

Figure 6.12: Average number of HOs

In Fig. 6.12, it depicts that there is an increasing number of HOs with the increase of train speed. We compare the average number of HOs of the proposed scheme with the conventional HO approach, Q-learning algorithm [131] and other classification algorithms including Decision Tree (DT), K-Nearest Neighbors algorithm (KNN) and Artificial Neural Network algorithm (ANN). The conventional HO method is A3 event and is used as the benchmark. From the figure, it's obvious that the proposed method outperforms other methods and achieves the lowest HO number. Furthermore, among various classification algorithms, ANN performs better than KNN and DT because of its multi-layer connection and various activation functions to deal with nonlinear classification problems. Therefore, the simulation result clearly indicates that the SVM-based HO method is significantly effective in reducing the number of unnecessary HOs.

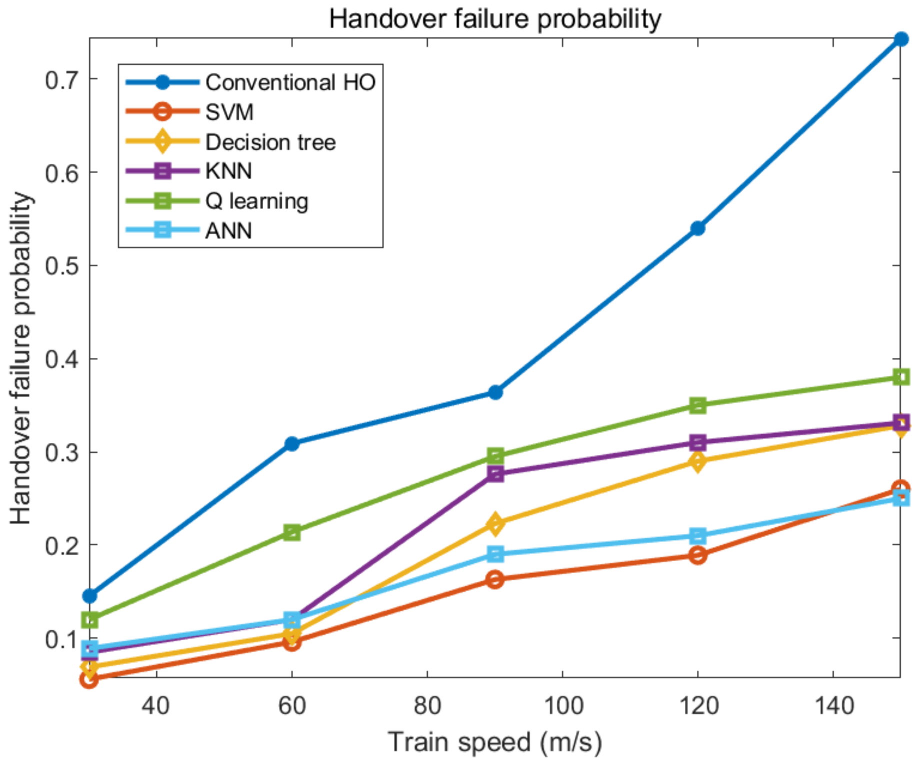

Figure 6.13: Handover failure probability

The probability of HO failure rises with the increasing speed of the train as shown in Fig. 6.13. This is primarily due to the limited time available for the BSs to execute the handover process at the target cell and also the degraded signal. The proposed scheme achieves the best performance when compared with the conventional HO method and Q-learning algorithm [131]. Initially, the HO failure rate is almost the same for the proposed scheme and other classification learning algorithms. However, when the train speed exceeds 60m/s, the proposed SVM-based HO method performs better than other learning algorithms.

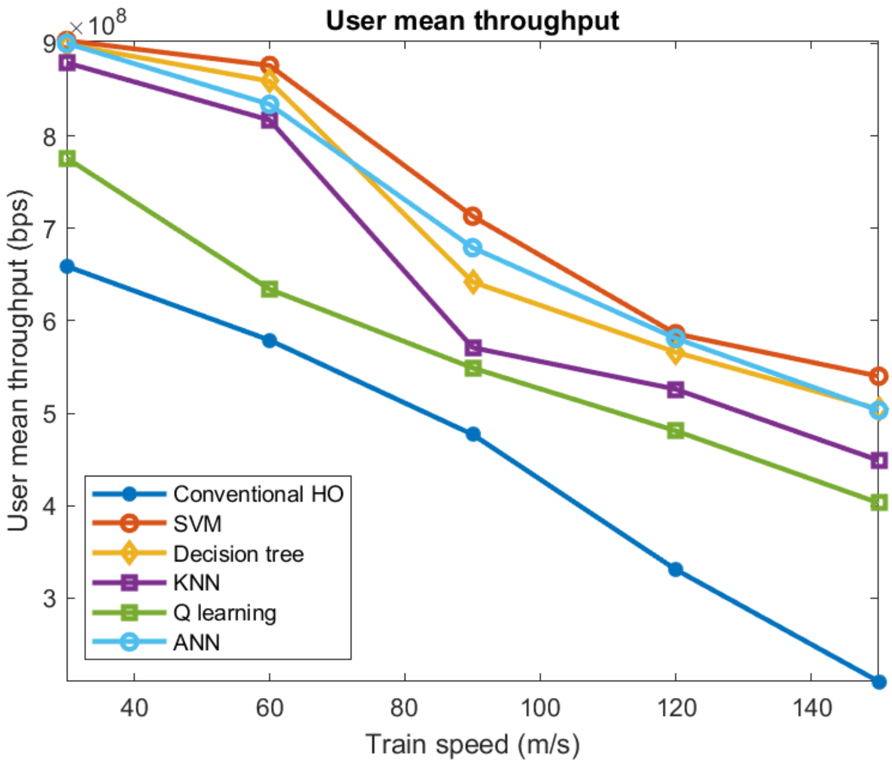

Figure 6.14: User mean throughput

With the increase in train speed, the throughput of users for all methods is decreased as illustrated in Fig. 6.14. The conventional HO method has the worst throughput performance due to the high HO failure rate and the large number of HOs. In addition, the proposed scheme clearly outperforms Q-learning algorithm [131]. When compared with other classification learning algorithms, the proposed method and ANN algorithm can achieve slightly better throughput than DT and obviously better throughput than KNN algorithm. Specifically, when the speed of train is 90m/s, the throughput of the proposed scheme outperforms by about 5%, 11.1%, 24.9%, 29.8% and 49.5% respectively in comparison with ANN, DT, KNN, Q-learning algorithm and conventional HO scheme. The increased number of HO can reduce the throughput. Since frequent handovers can lead to temporary disruptions (handover time) in data transmission. Each handover process involves control signaling and potential latency, which can lower the effective data throughput. Moreover, high handover failure probability can also increase data transmission interruptions, thus negatively impacting throughput.

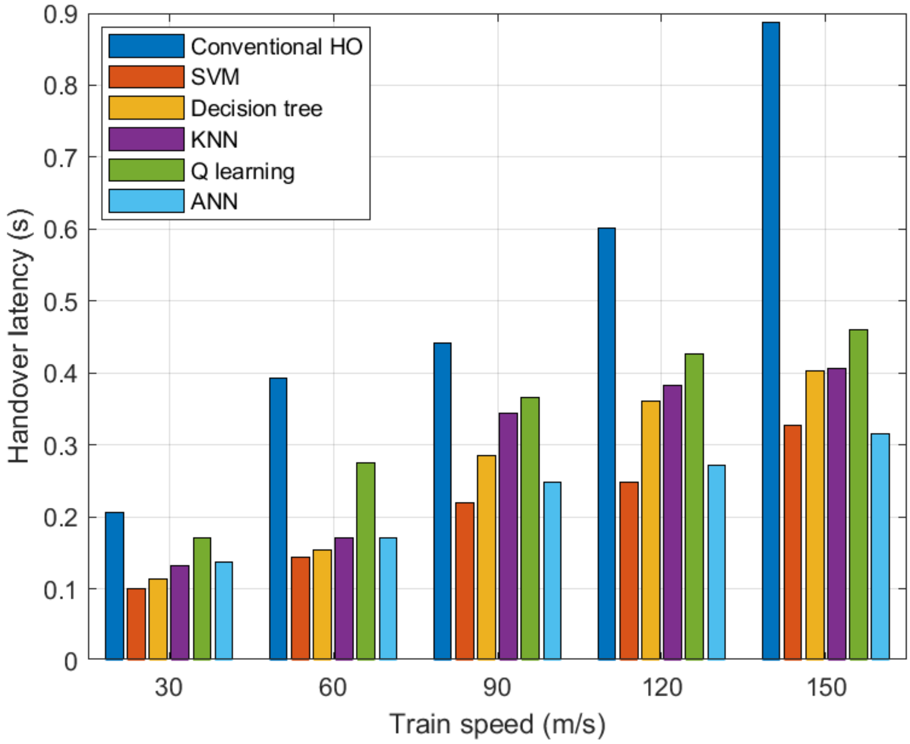

Figure 6.15: Handover latency

Fig. 6.15 shows the handover latency with respect to the train speed, where the proposed scheme can provide a lower handover latency than the conventional HO approach, Q-learning algorithm [131] and other classification algorithms. Specifically, when the train speed is 60m/s, the handover latency of the proposed scheme decreases by about 65.7% in comparison to the conventional HO scheme. The handover latency can be extended as the number of handovers increases because frequent handovers need a longer time for the handover process in data transmission.

6.5.4 Performance Analysis of Classification Algorithms

The performance of the proposed SVM based HO algorithm, DT, KNN and ANN are evaluated by using different metrics such as Precision, Recall, accuracy and F1 score.

1) **Accuracy** The accuracy of classification is defined as,

$$\text{Accuracy} = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}} \tag{6.38}$$

It is used to evaluate the accuracy of the prediction model.

2) **F1 score** The F1 score is a metric that combines precision and recall through a harmonic mean. A higher F1 score, closer to 1, indicates better performance of the ML algorithm. It can be defined as follows,

$$\text{F1 score} = 2 * \frac{(\text{precision} * \text{recall})}{(\text{precision} + \text{recall})}, \tag{6.39}$$

where,

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \tag{6.40}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \tag{6.41}$$

where TP, TN, FP and FN mean true positives, true negatives, false positives, and false negatives respectively. Precision is the ratio of the True Positives to all the Positives and also gives us a measure of the relevant data points. But for the Recall, it is the measure of how accurately our model is able to classify the relevant data. Therefore, F1 scores strike a trade-off between precision and recall by computing the weighted average of the two measures to evaluate the classification model.

The performance metrics obtained during the test phase are summarized in Table 6.2. The

Table 6.2: Performance metrics obtained on the test set of dataset

Algorithm	Accuracy	Precision	Recall	F1 score
Support Vector Machine	98.9%	0.955	0.985	0.97
Decision Tree	96.6%	0.917	0.931	0.924
K-Nearest Neighbors	94.1%	0.946	0.769	0.84
Artificial Neural Network	99%	0.987	0.864	0.92

trained model is assessed on the test set of the dataset. Table 6.2 shows that SVM and ANN-based ML methods can achieve better accuracy and outperform in general when compared with other ML classification methods. From Table 6.2, we can conclude that F1 score for SVM performs better than other algorithms during the test phase.

Confusion matrix is often used to provide valuable insight into the accuracy of our predictions and their alignment with the actual values. It enables us to analyze the correctness of our predictions and evaluate their performance comprehensively. Correct prediction means the predicted class matches the actual class by the classification model. The results of the confusion matrix with the proposed SVM scheme are presented in Fig. 6.16.

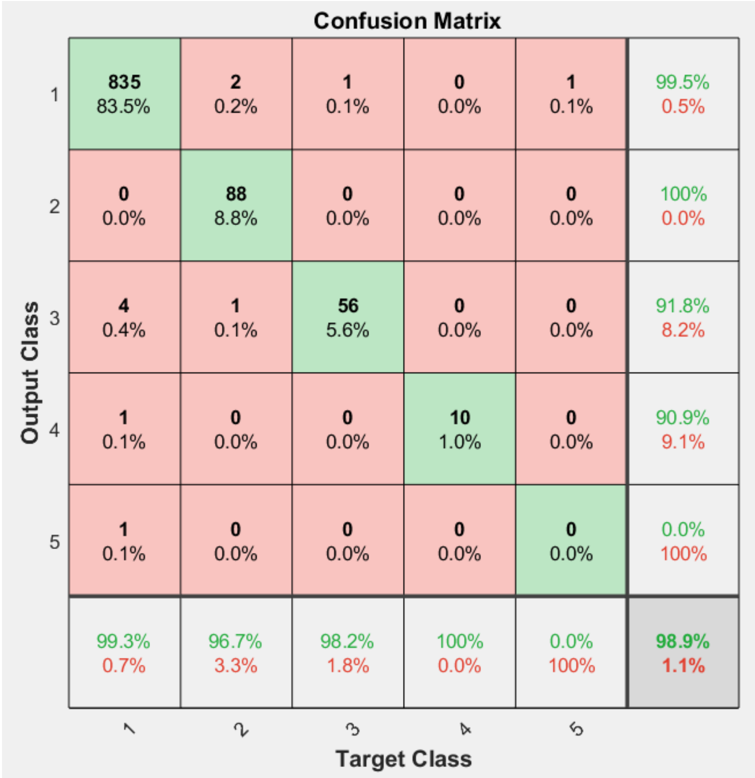

Figure 6.16: SVM-based handover algorithm

In Fig. 6.16, the matrix compares the actual classes (target classes) with the predicted classes (output classes) by the classification model.

- **Green Blocks (Diagonal)**

  These represent correct predictions by the classification model, where the predicted class matches the actual class. They are often referred to as "True Positives" for the respective classes. For instance, the block with number 835 indicates that there are 835 samples where class 1 was correctly predicted as class 1.

- **Red Blocks (Off-Diagonal)**

  These indicate incorrect predictions. The red blocks not on the diagonal show samples where the prediction differs from the actual class. For example, the red block with the number 1 in the first row and third column indicates that there is 1 sample where the actual class was 3, but the model incorrectly predicted it as class 1.

- **White Block(Right Side)**

  The percentages in the last column represent the precision for each class, including both the percentages of correct and incorrect predictions. For instance, for actual target class 1, 99.5% of the predictions were correct and 0.5% were incorrect.

- **White Block(Bottom)**

  The percentages in the bottom row represent the recall for each class. For example, 99.3% of the predictions for class 1 are actually class 1, but 0.7% are not.

- **Gray Block**

  Overall Accuracy: this is the sum of the green block numbers divided by the total predictions. Accuracy should be as high as possible.

The confusion matrices of other ML-based algorithms are shown from Fig. 6.17 to Fig. 6.19.

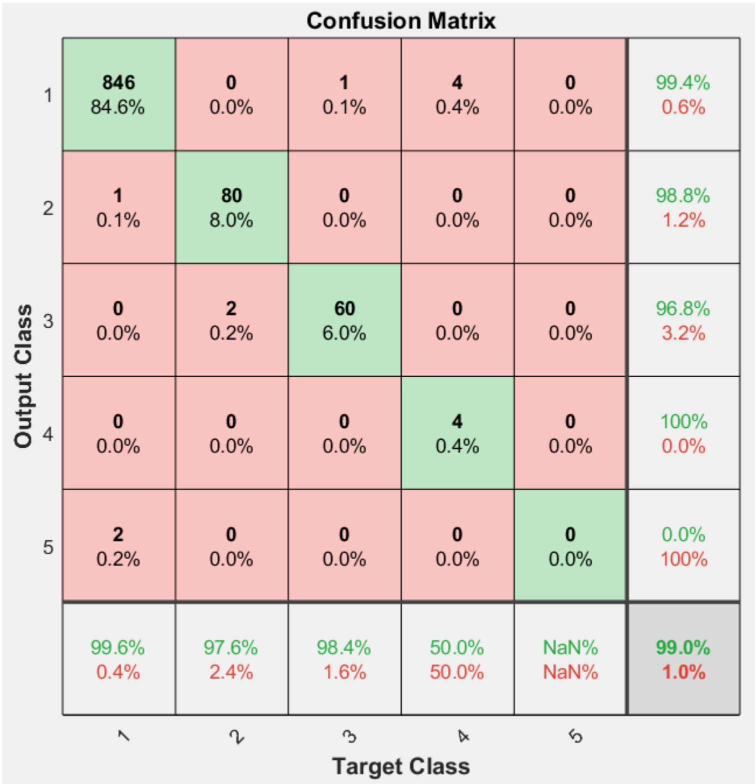

Figure 6.17: ANN-based handover algorithm

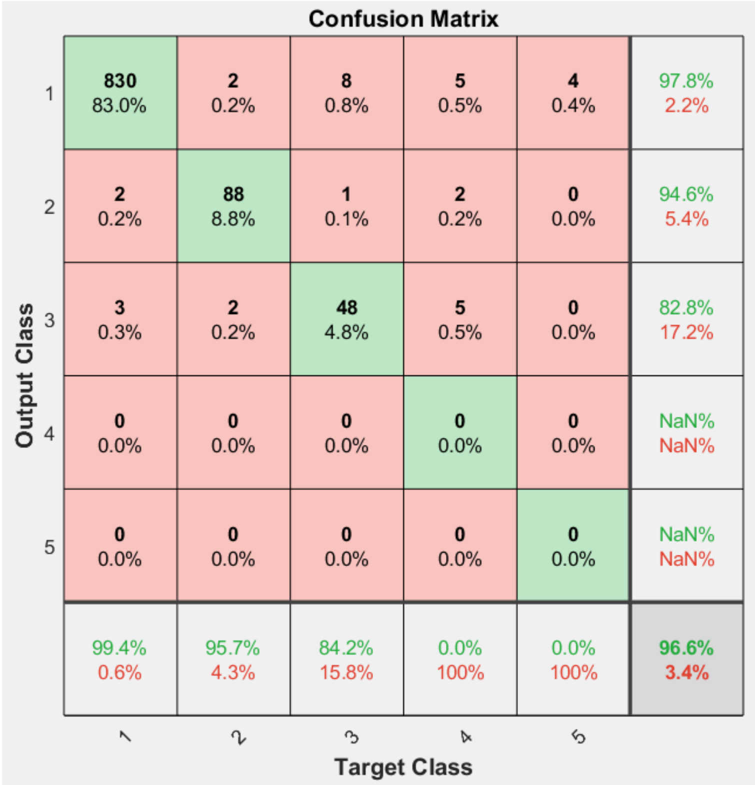

Figure 6.18: DT based handover algorithm

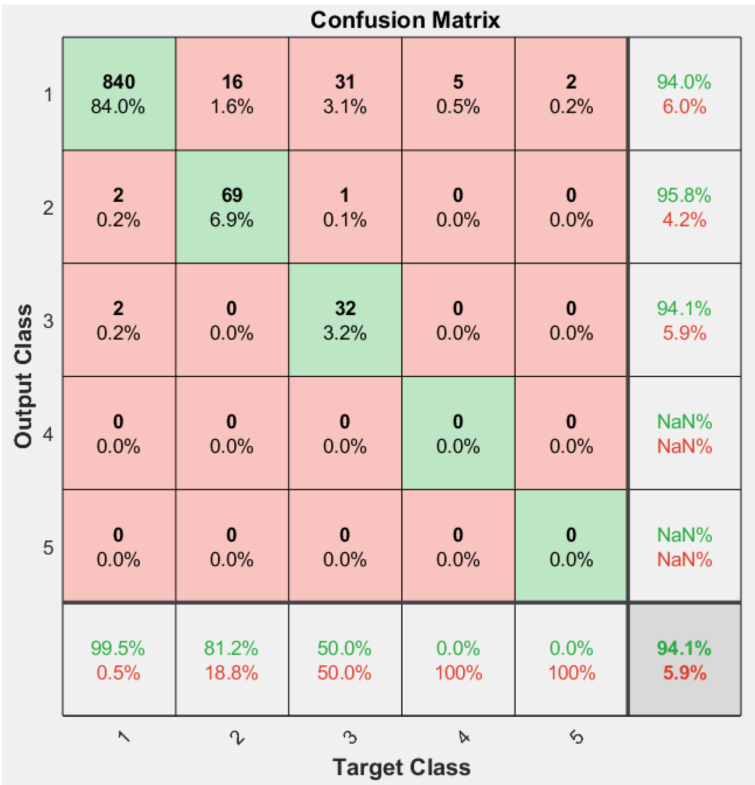

Figure 6.19: KNN-based handover algorithm

From Fig. 6.16 and Fig. 6.19, the confusion matrices demonstrate that the SVM and ANN algorithms have similar prediction performance and both achieve the highest accuracy of the model which are 98.9% and 99% accuracy respectively when compared to the 96.6% and 94.1% accuracy of other ML-based algorithms.

6.6 Summary

This chapter has investigated the HO problems in the HSR wireless communication system. We propose a novel HO scheme based on the Kalman filter and machine learning algorithm to predict the train's next target BS. The prediction task is considered as a classification problem by applying the SVM on available information in the HSR network. This chapter establishes an effective SVM classifier model for improving HO decision making by considering multiple criteria along the railway track. As demonstrated by the simulation results, the proposed scheme outperforms the conventional approach, Q-learning algorithm [131] and other ML-based algorithms in terms of HO number, user mean throughput, HO failure probability and handover latency. In addition, we also use various performance metrics to evaluate their accuracy. The results show that the SVM and ANN algorithms can achieve higher accuracy which are found to be 98.9% and 99%. However, during the actual implementation process (HO decision making process), the SVM outperforms ANN in terms of the average number of HOs, HO failure probability, user mean throughput and handover latency. Therefore, the proposed prediction algorithm holds great promise and deserves further investigation in real-world scenarios.

Chapter 7

Conclusion and Future Work

In this chapter, we summarize the research findings of this thesis and highlight the future directions of this work.

7.1 Conclusions

The work in this thesis focuses on the handover techniques in the HSR wireless communication networks.

In chapter 4, a HO method is proposed to reduce the outage probability and enhance the handover trigger and success probability by improving the handover decision making process. This was done by employing multi-criteria including train speed, location, RSRP and RSRQ based on dual-layer and DAS architecture. The dual-layer involves a TRS positioned on top of the train, facilitating communication between the train and the BSs, thereby reducing penetration loss and mitigating group handover issues. Integrated with DAS architecture, this model employs strategically positioned antennas to enhance signal strength and ensure reliability. Furthermore, in order to evaluate our proposed HO method, three handover performances such as handover trigger probability, handover success probability and outage probability were analyzed theoretically. It was shown that the simulation results coincide with the theoretical analysis perfectly, they both showed that the proposed HO method can improve the handover success probability in comparison with the conventional RSRP based handover decision algorithm.

In chapter 5, we proposed a novel HO method to enhance the throughput and HO success probability, in addition to decreasing the HO number and HO latency. This method was done by applying a reinforcement learning algorithm, namely Q-learning in the HSR scenario. The Q-learning based HO algorithm can dynamically optimize HO decisions to adapt to the fast varying environment in the HSR system. The proposed method leverages the RSRP, SINR and ToS to make efficient handover decisions for seamless connectivity while considering HO signaling overhead. The simulation results demonstrate that the proposed HO method outperforms the conventional approach and existing algorithms in terms of throughput, HO number, HO success probability and HO latency. Moreover, it indicates that the proposed method has the potential to greatly enhance the HO process in the context of 5G, enabling more efficient management of high-speed mobility in the 5G network.
Finally, in chapter 6, we proposed an intelligent HO method by using various machine learning algorithms to choose the optimal target BS. Firstly, the Kalman filter is used to estimate the posteriori of RSRP for the neighboring BSs, taking the measurement report as its primary input. Then, the SVM module considers the output of the Kalman filter, SINR and the traveling time of neighboring BSs as its input to determine the most suitable target BS for a HO. It was shown that the proposed HO method outperforms the conventional approach, existing algorithms and other ML-based algorithms in terms of HO number, user mean throughput, HO failure probability and HO latency. In addition, we also use various performance metrics to evaluate their accuracy. The results demonstrate that the SVM and ANN algorithms can achieve higher accuracy which are found to be 98.9% and 99%. However, during the actual implementation process (HO decision making process), the SVM can perform better in terms of the average number of HOs, HO failure probability, user mean throughput and handover latency.

7.2 Future Works

This section outlines some of the potential future works based on the topics presented in this thesis.

1.) Overloading problem: more complicated scenarios such as the increased number of trains and users, can be considered as one of the future challenges in the HSR network. For instance, some of the BSs might face overloading problems. For this challenge, dynamic data offloading strategies are needed to avoid congested data traffic and should be considered in the HO procedure in HSR. For example, we can consider one more criterion such as capacity when making HO decision to offload the traffic from congested BSs.

2.) Channel modeling: we will consider using ray-tracing techniques to construct to provide a more realistic channel model. Furthermore, the dataset could be also augmented with measured data and empirically derived models based on measurements in order to obtain a more realistic model in HSR. The work of chapter 6 can be extended by improving the dataset construction.

3.) This thesis mainly considers high-speed trains operating in open areas such as rural. In fact, trains can also pass through various complex terrains such as urban areas and tunnels, and the handover algorithm needs to be optimized by comprehensively considering various environmental factors.

4.) More advanced 5G technologies can be used in the HSR network: for instance, due to the high-frequency spectra, the mmWave may suffer from severe path loss and reduced coverage. To overcome this disadvantage, massive MIMO based on beamforming technology can be considered in the HSR mmWave system in the future. Even though beamforming technology is used, the coverage of each beam is limited. When the user leaves the serving beam coverage, the original serving beam will no longer be suitable for it, and a beam handover will be needed to connect to a new beam. Thus, in the high-speed railway scenario, frequent beam handover will happen because of the narrower beam coverage in 5G network.

References

[1] D Lee, G Gil, and D Kim. A cost-based adaptive handover hysteresis scheme to minimize the handover failure rate in 3gpp lte system. *EURASIP J. Wireless Comm. and Networking*, 2010, 12 2010.

[2] B Davaasambuu, K Yu, and T Sato. Self-optimization of handover parameters for long-term evolution with dual wireless mobile relay nodes. *Future Internet*, 7:196–213, 06 2015.

[3] H Kholidy, A Rahman, W Balbuena, A Sayed, M Badr, and F Mustafa. A survey study for the 5g emerging technologies. 04 2023.

[4] M Hosny, M Khashaba, I Khedr, and A Amer. New vertical handover prediction schemes for lte-wlan heterogeneous networks. *PloS one*, 14(4):e0215334, 2019.

[5] Aleksander S and Jose S. An overview of gsm-r technology and its shortcomings. In *2012 12th International Conference on ITS Telecommunications*, pages 626–629, 2012.

[6] R He, B Ai, Z Zhong, M Yang, R Chen, Jn Ding, Z Ma, G Sun, and C Liu. 5g for railways: Next generation railway dedicated communications. *IEEE Communications Magazine*, 60(12):130–136, 2022.

[7] ETSI TR 138 913 V14.2.0 (2017-05). 5g; study on scenarios and requirements for next generation access technologies (3gpp tr 38.913 version 14.2.0 release 14). 2017.

[8] Y Zhou and B Ai. Handover schemes and algorithms of high-speed mobile environment: A survey. *Computer Communications*, 47:1–15, 2014.

[9] J Sheng, X Cai, Q Li, C Wu, B Ai, Y Wang, M Kadoch, and P Yu. Space-air-ground integrated network development and applications in high-speed railways: A survey. *IEEE Transactions on Intelligent Transportation Systems*, 23(8):10066–10085, 2022.

[10] L Yin, T Song, Q Ni, Q Xiao, Y Sun, and W Guo. New signal and algorithms for 5g/6g high precision train positioning in tunnel with leaky coaxial cable. *IEEE Journal on Selected Areas in Communications*, pages 1–1, 2023.

[11] J Zhang, Z Tan, and X Yu. Coverage efficiency of radio-over-fiber network for high-speed railways. In *2010 6th International Conference on Wireless Communications Networking and Mobile Computing (WiCOM)*, pages 1–4, 2010.

[12] B Karimi, Jn Liu, and C Wang. Seamless wireless connectivity for multimedia services in high speed trains. *IEEE Journal on selected areas in communications*, 30(4):729–739, 2012.

[13] L Lu, X Fang, M Cheng, C Yang, W Luo, and C Di. Positioning and relay assisted robust handover scheme for high speed railway. In *2011 IEEE 73rd Vehicular Technology Conference (VTC Spring)*, pages 1–5. IEEE, 2011.

[14] J Wang, X Yang, S Zhao, L Zhang, M Ni, and S Lin. Handover performance improvement for ultra dense network of high-speed railway. In *2017 IEEE 85th Vehicular Technology Conference (VTC Spring)*, pages 1–5. IEEE, 2017.

[15] G Zhong, K Xiong, Z Zhong, and B Ai. Internet of things for high-speed railways. *Intelligent and Converged Networks*, 2(2):115–132, 2021.

[16] R Amiri, H Mehrpouyan, L Fridman, R Mallik, A Nallanathan, and D Matolak. A machine learning approach for power allocation in hetnets considering qos. 03 2018.

[17] G Welch, G Bishop, et al. An introduction to the kalman filter. 1995.

[18] Vijay K. All you need to know about support vector machines, 2022.

[19] R Safavian and D Landgrebe. A survey of decision tree classifier methodology. *IEEE transactions on systems, man, and cybernetics*, 21(3):660–674, 1991.

[20] K-nearest neighbor(knn) algorithm, Jan 2024.

[21] B Ai, X Cheng, T Kürner, Z Zhong, K Guan, R He, L Xiong, W. Matolak, G. Michelson, and C Briso-Rodriguez. Challenges toward wireless communications for high-speed railway. *IEEE Transactions on Intelligent Transportation Systems*, 15(5):2143–2158, 2014.

[22] Wang K, Li S, Lin Y, DOU Z, and Li W. Performance analysis of high-speed railway handover scheme with different network architecture. In *2019 IEEE 8th Joint International Information Technology and Artificial Intelligence Conference (ITAIC)*, pages 1894–1898, 2019.

[23] Rasha B, Hussein E, and Mohamed A. Handover scheme for 5g communications on high speed trains. pages 143–149, 04 2020.

[24] Naina G and Brahmjit S. A novel seamless handover scheme for high-speed railway transport using dual-antenna system. pages 1160–1165, 05 2019.

[25] Y Chen, K Niu, and Z Wang. Adaptive handover algorithm for lte-r system in high-speed railway scenario. *IEEE Access*, 9:59540–59547, 2021.

[26] R El Banna, M ELAttar, and M Abou El-Dahab. Fast adaptive handover using fuzzy logic for 5g communications on high speed trains. In *2021 16th International Conference on Telecommunications (ConTEL)*, pages 10–17. IEEE, 2021.

[27] 3gpp ts 36.331 version 11.19.0, "lte; evolved universal terrestrial radio access (e-utra); radio resource control (rrc); protocol specification,". 2018.

[28] H A Sadrabadi, Nahid A, and Hamidreza B. Performance analysis of the lte handover decision events for the high speed railway application. *Journal of Communication Engineering*, 9(2):184–200, 2020.

[29] Y Chen, J Kang, K Niu, et al. Research on adaptive handover decision algorithm in next-generation high-speed railway scenario. *Mobile Information Systems*, 2022, 2022.

[30] G Yang, X Wang, and X Chen. Handover control for lte femtocell networks. In *2011 International conference on electronics, communications and control (ICECC)*, pages 2670–2673. IEEE, 2011.

[31] H Zhu and Y Peng. Research on adaptive handover scheme based on improved genetic algorithm. *Procedia Computer Science*, 166:557–562, 2020.

[32] Aziz Ur R, Mardeni Bin R, and Tiang J. A survey of handover management in mobile hetnets: Current challenges and future directions. *Applied Sciences*, 13(5):3367, 2023.

[33] A Ulvan, R Bestak, and M Ulvan. The study of handover procedure in lte-based femtocell network. In *WMNC2010*, pages 1–6. IEEE, 2010.

[34] N Rajule, B Ambudkar, and A Dhande. Survey of vertical handover decision algorithms. *International Journal of Innovations in Engineering and Technology (IJIET)*, 2(1):362–368, 2013.

[35] J Márquez-Barja, T Calafate, J Cano, and P Manzoni. An overview of vertical handover techniques: Algorithms, protocols and tools. *Computer communications*, 34(8):985–997, 2011.

[36] M Zekri, B Jouaber, and D Zeghlache. A review on mobility management and vertical handover solutions over heterogeneous wireless networks. *Computer Communications*, 35(17):2055–2068, 2012.

[37] N Elhilali, M Badri, and M Bouami. An overview of vertical handover decision algorithms. In *AIP Conference Proceedings*, volume 2814. AIP Publishing, 2023.

[38] N Nasser, A Hasswa, and H Hassanein. Handoffs in fourth generation heterogeneous networks. *IEEE Communications Magazine*, 44(10):96–103, 2006.

[39] J Zhao, Y Liu, Y Gong, C Wang, and L Fan. A dual-link soft handover scheme for c/u plane split network in high-speed railway. *IEEE Access*, 6:12473–12482, 2018.

[40] W Luo, R Zhang, and X Fang. A comp soft handover scheme for lte systems in high speed railway. *EURASIP Journal on wireless Communications and Networking*, 2012(1):1–9, 2012.

[41] J Li, L Tian, Y Zhou, and J Shi. An adaptive handover trigger scheme for wireless communications on high speed rail. In *2012 IEEE International Conference on Communications (ICC)*, pages 5185–5189, 2012.

[42] Y Lu, K Xiong, Z Zhao, P Fan, and Z Zhong. Remote antenna unit selection assisted seamless handover for high-speed railway communications with distributed antennas. In *2016 IEEE 83rd Vehicular Technology Conference (VTC Spring)*. IEEE, may 2016.

[43] Kristin K. 4.1: Probability density functions (pdfs) and cumulative distribution functions (cdfs) for continuous random variables, 2021.

[44] Dinko Z, Toni M, Ivana Nizetic Kosovic, Mario C, and Josip L. Analyses of ping-pong handovers in real 4g telecommunication networks. *Computer Networks*, 227:109699, 2023.

[45] S Alraih, R Nordin, I Shayea, F Abdullah, and A Alhammadi. Ping-pong handover effect reduction in 5g and beyond networks. In *2021 IEEE Microwave Theory and Techniques in Wireless Communications (MTTW)*, pages 97–101, 2021.

[46] Q Ding, T Fu, S Wang, and J Luo. Precoding-based handover scheme design for high-speed railway communication. *IEEE Wireless Communications Letters*, 12(2):332–335, 2023.

[47] Md. Rajibul P, Md. Rakibul I, Palash R, Md. Abdur R, Ahmad A, Salman A. A, and Mohammad H. Multi-criteria handover mobility management in 5g cellular network. *Computer Communications*, 174:81–91, 2021.

[48] 3GPP TR 36.881 V14.0.0 (2016-06). 3gpp technical specification group radio access network; evolved universal terrestrial radio access (e-utra); study on latency reduction techniques for lte (release 14). 2016.

[49] W Ali, J Wang, H Zhu, and J Wang. An expedited predictive distributed antenna system based handover scheme for high-speed railway. In *GLOBECOM 2017 - 2017 IEEE Global Communications Conference*, pages 1–6, 2017.

[50] R Chen, W Long, G Mao, and C Li. Development trends of mobile communication systems for railways. *IEEE Communications Surveys Tutorials*, 20(4):3131–3141, 2018.

[51] R He, B Ai, G Wang, K Guan, Z Zhong, F. Molisch, C Briso-Rodriguez, and P. Oestges. High-speed railway communications: From gsm-r to lte-r. *IEEE Vehicular Technology Magazine*, 11(3):49–58, 2016.

[52] P Lamas, José Rodríguez-Piñeiro, José A García-Naya, and L Castedo. A survey on lte networks for railway services. In *Proceedings of the IEEE Congreso de Ingenierıa en Electro-Electrónica, Comunicaciones y Computación (ARANDUCON 2012)*, 2012.

[53] Y Zhang, M Wu, S Ge, L Luan, and A Zhang. Optimization of time-to-trigger parameter on handover performance in lte high-speed railway networks. In *The 15th International Symposium on Wireless Personal Multimedia Communications*, pages 251–255, 2012.

[54] T Zhou, H Li, Y Wang, L Liu, and C Tao. Channel modeling for future high-speed railway communication systems: A survey. *IEEE Access*, 7:52818–52826, 2019.

[55] Q Du, G Wu, Q Yu, and S Li. Ici mitigation by doppler frequency shift estimation and pre-compensation in lte-r systems. In *2012 1st IEEE International Conference on Communications in China (ICCC)*, pages 469–474, 2012.

[56] Z Hou, Y Zhou, L Tian, J Shi, Y Li, and B Vucetic. Radio environment map-aided doppler shift estimation in lte railway. *IEEE Transactions on Vehicular Technology*, 66(5):4462–4467, 2017.

[57] R He, B Ai, Z Zhong, M Yang, C Huang, R Chen, J Ding, H Mi, Z Ma, G Sun, and C Liu. Radio communication scenarios in 5g-railways. *China Communications*, pages 1–12, 2023.

[58] B Sun, Y Guo, Y Yu, J Liu, B Ai, Z Zhong, X Sun, and W Wang. Reliability analysis of ctcs-3 train-ground communication system based on 5g-r. *IEEE Transactions on Vehicular Technology*, pages 1–14, 2023.

[59] B Ai, F. Molisch, M Rupp, and Z Zhong. 5g key technologies for smart railways. *Proceedings of the IEEE*, 108(6):856–893, 2020.

[60] E Dahlman, S Parkvall, and J Skold. *5G NR: The next generation wireless access technology*. Academic Press, 2020.

[61] J Zhang, E Björnson, M Matthaiou, H Ng, Dand Yang, and J. Love. Prospective multiple antenna technologies for beyond 5g. *IEEE Journal on Selected Areas in Communications*, 38(8):1637–1660, 2020.

[62] S Alraih, R Nordin, A Abu-Samah, I Shayea, and F Abdullah. A survey on handover optimization in beyond 5g mobile networks: Challenges and solutions. *IEEE Access*, 11:59317–59345, 2023.

[63] C Tao, J Qiu, J Chen, L Yu, W Dong, and Y Yuan. Position-based modeling for wireless channel on high-speed railway under a viaduct at 2.35 ghz. *IEEE Journal on Selected Areas in Communications - JSAC*, 30:834–845, 05 2012.

[64] J Liu and X Wang. Handover analysis on high speed train with doppler frequency spread. *ArXiv*, abs/1703.09869, 2017.

[65] Y Lu, K Xiong, Z Zhao, P Fan, and Z Zhong. Remote antenna unit selection assisted seamless handover for high-speed railway communications with distributed antennas. In *2016 IEEE 83rd Vehicular Technology Conference (VTC Spring)*, pages 1–6, 2016.

[66] B Han, L Liu, J Zhang, C Tao, C Qiu, T Zhou, Z Li, and Z Piao. Research on resource migration based on novel rrh-bbu mapping in cloud radio access network for hsr scenarios. *IEEE Access*, 7:108542–108550, 2019.

[67] J Liu, Y Shi, M. Fadlullah, and N Kato. Space-air-ground integrated network: A survey. *IEEE Communications Surveys Tutorials*, 20(4):2714–2741, 2018.

[68] X Ding, Z Zhang, and D Liu. Low-delay secure handover for space-air-ground integrated networks. In *2020 IEEE 31st Annual International Symposium on Personal, Indoor and Mobile Radio Communications*, pages 1–6, 2020.

[69] H Wang, B Ning, and H Jiang. An experimental study of 2.4ghz frequency band leaky coaxial cable in cbtc train ground communication. In *2011 IEEE 73rd Vehicular Technology Conference (VTC Spring)*, pages 1–5, 2011.

[70] Luo C, Yang X, Zhao H, Lin S, and Wang H. Measurement-based fading model with leaky coaxial cables for urban rail transit in tunnels. In *2020 IEEE International Sym-*

posium on Antennas and Propagation and North American Radio Science Meeting, pages 1285–1286, 2020.

[71] J Wang, H Zhu, and J. Gomes. Distributed antenna systems for mobile communications in high speed trains. *IEEE Journal on Selected Areas in Communications*, 30(4):675–683, 2012.

[72] Abdul Rafay, Sevia Mahdaliza Idrus, Kamaludin Mohamad Yusof, and Siti Hasunah Mohammad. A survey on advanced transmission technologies for high bandwidth and good signal quality for high-speed railways. *Indonesian Journal of Electrical Engineering and Computer Science*, 23(1):293–301, 2021.

[73] A Kanno, PT Dat, N Yamamoto, and T Kawanishi. Millimeter-wave radio-over-fiber system for high-speed railway communication. In *2016 Progress in Electromagnetic Research Symposium (PIERS)*, pages 3911–3915. IEEE, 2016.

[74] D Gavrilovich. Broadband communication on the highways of tomorrow. *IEEE Communications Magazine*, 39(4):146–154, 2001.

[75] Q Huang, J Zhou, C Tao, S Yi, and M Lei. Mobile relay based fast handover scheme in high-speed mobile environment. In *2012 IEEE Vehicular Technology Conference (VTC Fall)*, pages 1–6, 2012.

[76] X Qian, H Wu, and J Meng. A dual-antenna and mobile relay station based handover in distributed antenna system for high-speed railway. In *2013 Seventh International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing*, pages 585–590, 2013.

[77] T Deng, Z Zhang, X Wang, and P Fan. A network assisted fast handover scheme for high speed rail wireless networks. In *2016 IEEE 83rd Vehicular Technology Conference (VTC Spring)*, pages 1–5, 2016.

[78] Sajjad K, Ibraheem S, Mustafa E, and Hafizal M. Handover management over dual connectivity in 5g technology with future ultra-dense mobile heterogeneous networks: A review. *Engineering Science and Technology, an International Journal*, 35:101172, 2022.

[79] Aziz Ur R, Mardeni Bin R, and Tiang J. A survey of handover management in mobile hetnets: Current challenges and future directions. *Applied Sciences*, 13(5), 2023.

[80] J Zhao, Y Liu, Y Gong, C Wang, and L Fan. A dual-link soft handover scheme for c/u plane split network in high-speed railway. *IEEE Access*, 6:12473–12482, 2018.

[81] S Xie, X Yu, and Y Luo. A seamless dual-link handover scheme with optimized threshold for c/u plane network in high-speed rail. In *2016 IEEE 83rd Vehicular Technology Conference (VTC Spring)*, pages 1–5. IEEE, 2016.

[82] S Kang, S Choi, G Lee, and S Bahk. A dual-connection based handover scheme for ultra-dense millimeter-wave cellular networks. In *2019 IEEE Global Communications Conference (GLOBECOM)*, pages 1–6. IEEE, 2019.

[83] L Luan, Wu M, J Shen, J Ye, and X He. Optimization of handover algorithms in lte high-speed railway networks. *International Journal of Digital Content Technology and its Applications*, 6:79–87, 03 2012.

[84] Z Yang, Y Zheng, and J Jing. Optimization of adaptive handover algorithm based on distributed antenna in lte-r. *Journal of Physics: Conference Series*, 1187(3):032094, apr 2019.

[85] Z Chowdhury and M Jang. Handover management in high-dense femtocellular networks. *EURASIP Journal on Wireless Communications and Networking*, 2013:1–21, 2013.

[86] M Chowdhury, B Trung, and Y Jang. Neighbor cell list optimization for femtocell-to-femtocell handover in dense femtocellular networks. In *2011 Third International Conference on Ubiquitous and Future Networks (ICUFN)*, pages 241–245, 2011.

[87] J.-M. Moon and D.-H. Cho. Efficient handoff algorithm for inbound mobility in hierarchical macro/femto cell networks. *IEEE Communications Letters*, 13(10):755–757, 2009.

[88] J Liu, K Xiong, Z Zhao, S Zhong, Z Zhong, and Bo A. The effects of moving speed on handover performances with measurement data. In *2017 15th International Conference on ITS Telecommunications (ITST)*, pages 1–5, 2017.

[89] C Wu, X Cai, J Sheng, Z Tang, B Ai, and Y Wang. Parameter adaptation and situation awareness of lte-r handover for high-speed railway communication. *IEEE Transactions on Intelligent Transportation Systems*, 23(3):1767–1781, 2022.

[90] H Cho, S Shin, G Lim, C Lee, and J Chung. Lte-r handover point control scheme for high-speed railways. *IEEE Wireless Communications*, 24(6):112–119, 2017.

[91] D Deng, M Wu, and Z Chen. An optimized relay handover algorithm in td-lte along high-speed railway. In *2013 IEEE International Conference of IEEE Region 10 (TENCON 2013)*, pages 1–4. IEEE, 2013.

[92] H Chen and Y Cahyadi. A grey prediction based hard handover hysteresis algorithm for 3gpp lte system. In *2012 Seventh International Conference on Broadband, Wireless Computing, Communication and Applications*, pages 590–595. IEEE, 2012.

[93] J Zhao, Y Liu, C Wang, L Xiong, and L Fan. High-speed based adaptive beamforming handover scheme in lte-r. *IET communications*, 12(10):1215–1222, 2018.

[94] H Wu, Y Gu, and Z Zhong. Research on the fast algorithm for gsm-r switching for high-speed railway. *Journal of Railway Engineering Society*, 124:92–96, 2009.

[95] J Bang, S Oh, K Kang, and Y Cho. A bayesian regression based lte-r handover decision algorithm for high-speed railway systems. *IEEE Transactions on Vehicular Technology*, 68(10):10160–10173, 2019.

[96] D Li, D Li, and Y Xu. Machine learning based handover performance improvement for lte-r. In *2019 IEEE International Conference on Consumer Electronics - Taiwan (ICCE-TW)*, pages 1–2, 2019.

[97] Q Wang, C Wu, and J Sheng. Intelligent adaptive mechanism of handover parameters in high-speed railway communication network based on generative adversarial network. In *2022 IEEE 22nd International Conference on Communication Technology (ICCT)*, pages 683–688, 2022.

[98] L Ma, J Wu, and C Li. Localization of a high-speed train using a speed model based on the gradient descent algorithm. *Future Generation Computer Systems*, 85:201–209, 2018.

[99] Y Lu, K Xiong, P Fan, Z Zhong, and B Ai. The effect of power adjustment on handover in high-speed railway communication networks. *IEEE Access*, 5:26237–26250, 2017.

[100] J Wang and L Dai. Asymptotic rate analysis of downlink multi-user systems with co-located and distributed antennas. *IEEE Transactions on Wireless Communications*, 14(6):3046–3058, 2015.

[101] H Zhu. Performance comparison between distributed antenna and microcellular systems. *IEEE Journal on Selected Areas in Communications*, 29(6):1151–1163, 2011.

[102] W Ali, J Wang, H Zhu, and J Wang. Distributed antenna system based frequency switch scheme evaluation for high-speed railways. In *2017 IEEE International Conference on Communications (ICC)*, pages 1–6, 2017.

[103] L Tian, J Li, Y Huang, J Shi, and J Zhou. Seamless dual-link handover scheme in broadband wireless communication systems for high-speed rail. *IEEE Journal on Selected Areas in Communications*, 30(4):708–717, May 2012.

[104] X Qian, H Wu, and J Meng. A dual-antenna and mobile relay station based handover in distributed antenna system for high-speed railway. In *2013 Seventh International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing*, pages 585–590, 2013.

[105] W Ali, J Wang, H Zhu, and J Wang. An optimized fast handover scheme based on distributed antenna system for high-speed railway. In *2017 IEEE 86th Vehicular Technology Conference (VTC-Fall)*, pages 1–5, 2017.

[106] Wan C and Jeffrey G. A. Downlink performance and capacity of distributed antenna systems in a multicell environment. *IEEE Transactions on Wireless Communications*, 6(1):69–73, 2007.

[107] C Yang, L Lu, C Di, and X Fang. An on-vehicle dual-antenna handover scheme for high-speed railway distributed antenna system. In *2010 6th International Conference on Wireless Communications Networking and Mobile Computing (WiCOM)*, pages 1–5, 2010.

[108] Z Liu and P Fan. An effective handover scheme based on antenna selection in ground–train distributed antenna systems. *IEEE Transactions on Vehicular Technology*, 63(7):3342–3350, 2014.

[109] F Yang, H Deng, F Jiang, and X Deng. Handover optimization algorithm in lte high-speed railway environment. *Wireless Personal Communications*, 84(2):1577–1589, 2015.

[110] W Song and W Wu. Research on frequency switching algorithm based on distributed antenna system. *Journal of Physics: Conference Series*, 2383(1):012071, dec 2022.

[111] A Hossein, Nahid A, and Hamidreza B. Performance analysis of the lte handover decision events for the high speed railway application. *Journal of Communication Engineering*, 9(2):184–200, 2020.

[112] 3GPP TS 36.521. Evolved universal terrestrial radio access (eutra); user equipment (ue) conformance specification radio transmission and reception part 1: Conformance testing, v9.3.0. Dec.2010.

[113] W Gheth, M Rabie, B Adebisi, M Ijaz, and G Harris. Communication systems of high-speed railway: a survey. *Transactions on Emerging Telecommunications Technologies*, 32(4):e4189, 2021.

[114] E Gures, I Shayea, A Alhammadi, M Ergen, and H Mohamad. A comprehensive survey on mobility management in 5g heterogeneous networks: Architectures, challenges and solutions. *IEEE Access*, 8:195883–195913, 2020.

[115] M Pan, T Lin, and W Chen. An enhanced handover scheme for mobile relays in lte-a high-speed rail networks. *IEEE Transactions on Vehicular Technology*, 64(2):743–756, 2015.

[116] Hanan K Mohamed S and Mona E. Novel type-2 fuzzy logic technique for handover problems in a heterogeneous network. *Engineering Optimization*, 50(9):1533–1543, 2018.

[117] I Sönmez, Şand Shayea, S Khan, and A Alhammadi. Handover management for next-generation wireless networks: A brief overview. In *2020 IEEE Microwave Theory*

and Techniques in Wireless Communications (MTTW), volume 1, pages 35–40, 2020.

[118] V Yajnanarayana, H Rydén, and L Hévizi. 5g handover using reinforcement learning. In *2020 IEEE 3rd 5G World Forum (5GWF)*, pages 349–354, 2020.

[119] S. Mollel, S Kaijage, M Kisangiri, A Imran, and H. Abbasi. Multi-user position based on trajectories-aware handover strategy for base station selection with multi-agent learning. In *2020 IEEE International Conference on Communications Workshops (ICC Workshops)*, pages 1–6, 2020.

[120] S Khosravi and M Shokri-Ghadikolaei, Hand Petrova. Learning-based handover in mobile millimeter-wave networks. *IEEE Transactions on Cognitive Communications and Networking*, 7(2):663–674, 2021.

[121] Y Chen, X Lin, T Khan, and M Mozaffari. Efficient drone mobility support using reinforcement learning. In *2020 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 1–6, 2020.

[122] Tanu G and Sakshi K. Handover optimization scheme for lte-advance networks based on ahp-topsis and q-learning. *Computer Communications*, 133:67–76, 2019.

[123] D Hegazy, A Nasr, and A Kamal. Optimization of user behavior based handover using fuzzy q-learning for lte networks. *Wireless Networks*, 24:481–495, 2018.

[124] F Hasegawa, A Taira, G Noh, B Hui, H Nishimoto, A Okazaki, A Okamura, J Lee, and I Kim. High-speed train communications standardization in 3gpp 5g nr. *IEEE Communications Standards Magazine*, 2(1):44–52, 2018.

[125] B Ai, F. Molisch, M Rupp, and Z Zhong. 5g key technologies for smart railways. *Proceedings of the IEEE*, 108(6):856–893, 2020.

[126] Z Ding, Y Huang, H Yuan, and H Dong. Introduction to reinforcement learning. *Deep reinforcement learning: fundamentals, research and applications*, pages 47–123, 2020.

[127] S Sutton and G Barto. *Reinforcement learning: An introduction.* MIT press, 2018.

[128] C Watkins. Learning from delayed rewards. 1989.

[129] I Muqattash and J Hu. An ε-greedy multiarmed bandit approach to markov decision processes. *Stats*, 6(1):99–112, 2023.

[130] S Wang and L Zhang. A multi-criteria handover scheme for distributed antenna system based high-speed rail wireless communications. In *2021 24th International Symposium on Wireless Personal Multimedia Communications (WPMC)*, pages 1–6, 2021.

[131] S Wang and L Zhang. Q-learning based handover algorithm for high-speed rail wireless communications. In *2023 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 1–6. IEEE, 2023.

[132] J Wang, X Yang, S Zhao, L Zhang, M Ni, and S Lin. Handover performance improvement for ultra dense network of high-speed railway. In *2017 IEEE 85th Vehicular Technology Conference (VTC Spring)*, pages 1–5. IEEE, 2017.

[133] S Banerjee, M Hempel, and H Sharif. A survey of wireless communication technologies & their performance for high speed railways. 2016.

[134] J Zhao, Y Liu, Y Gong, C Wang, and L Fan. A dual-link soft handover scheme for c/u plane split network in high-speed railway. *IEEE Access*, 6:12473–12482, 2018.

[135] H Zhao, R Huang, J Zhang, and Y Fang. Handoff for wireless networks with mobile relay stations. In *2011 IEEE Wireless Communications and Networking Conference*, pages 826–831. IEEE, 2011.

[136] Li Chen and Y Huang. An enhanced handover scheme adopting mobile relays in a lte-a network for high-speed movements. In *2017 IEEE International Conference on Software Quality, Reliability and Security Companion (QRS-C)*, pages 567–568. IEEE, 2017.

[137] M Pan, T Lin, and W Chen. An enhanced handover scheme for mobile relays in lte-a high-speed rail networks. *IEEE Transactions on Vehicular Technology*, 64(2):743–756, 2014.

[138] H Song, X Fang, and Y Fang. Millimeter-wave network architectures for future high-speed railway communications: Challenges and solutions. *IEEE Wireless Com-*

munications, 23(6):114–122, 2016.

[139] W Ren, J Xu, D Li, Q Cui, and X Tao. A robust inter beam handover scheme for 5g mmwave mobile communication system in hsr scenario. In *2019 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 1–6. IEEE, 2019.

[140] Y Lu, C Zhang, D Chen, W Zhang, and K Xiong. Handover enhancement in high-speed railway 5g networks: A lstm-based prediction method. In *2022 13th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, pages 1–6. IEEE, 2022.

[141] S Alraih, R Nordin, I Shayea, F Abdullah, and A Alhammadi. Ping-pong handover effect reduction in 5g and beyond networks. In *2021 IEEE Microwave Theory and Techniques in Wireless Communications (MTTW)*, pages 97–101. IEEE, 2021.

[142] H Yang, A Alphones, Z Xiong, D Niyato, J Zhao, and K Wu. Artificial-intelligence-enabled intelligent 6g networks. *IEEE Network*, 34(6):272–280, 2020.

[143] R Palas, R Islam, P Roy, A Razzaque, A Alsanad, A AlQahtani, and M Hassan. Multi-criteria handover mobility management in 5g cellular network. *Computer Communications*, 174:81–91, 2021.

[144] Tarciana C de Brito G, Y Dantas, and V Sousa Jr. A machine learning approach for handover in lte networks with signal obstructions. *Journal of Communication and Information Systems*, 35(1):271–289, 2020.

[145] M Manalastas, M Farooq, A Zaidi, A Ijaz, W Raza, and A Imran. Machine learning-based handover failure prediction model for handover success rate improvement in 5g. In *2023 IEEE 20th Consumer Communications & Networking Conference (CCNC)*, pages 684–685. IEEE, 2023.

[146] J Lima, A Medeiros, E Aguiar, S JUNIOR, and T Guerra. User-level handover decision making based on machine learning approaches. *Journal of Communication and Information Systems*, 37(1):104–108, 2022.

[147] Josué V Cervantes-Bazán, Alma D Cuevas-R, L Rojas-Cárdenas, S Lazcano-Salas, F García-Lamont, A Soriano, J Rubio, and J Pacheco. Proactive cross-layer framework

based on classification techniques for handover decision on wlan environments. *Electronics*, 11(5):712, 2022.

[148] J Li, Y Niu, H Wu, B Ai, S Chen, Z Feng, Z Zhong, and N Wang. Mobility support for millimeter wave communications: Opportunities and challenges. *Commun. Surveys Tuts.*, 24(3):1816–1842, jul 2022.

[149] C Zhang, G Wang, M Jia, R He, L Zhou, and B Ai. Doppler shift estimation for millimeter-wave communication systems on high-speed railways. *IEEE Access*, 7:40454–40462, 2018.

[150] D Fan, Z Zhong, G Wang, and F Gao. Doppler shift estimation for high-speed railway wireless communication systems with large-scale linear antennas. In *2015 International Workshop on High Mobility Wireless Communications (HMWC)*, pages 96–100. IEEE, 2015.

[151] Y Li and J Cimini. Bounds on the interchannel interference of ofdm in time-varying impairments. *IEEE Transactions on Communications*, 49(3):401–404, 2001.

[152] Youngjoo K and Hyochoong B. Introduction to kalman filter and its applications. In Felix Govaers, editor, *Introduction and Implementations of the Kalman Filter*, chapter 2. IntechOpen, Rijeka, 2018.

[153] R Diversi, R Guidorzi, and U Soverini. Kalman filtering in extended noise environments. *IEEE Transactions on Automatic Control*, 50(9):1396–1402, 2005.

[154] R Olfati-Saber. Kalman-consensus filter: Optimality, stability, and performance. In *Proceedings of the 48h IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference*, pages 7036–7042. Ieee, 2009.

[155] G Bishop, G Welch, et al. An introduction to the kalman filter. *Proc of SIGGRAPH, Course*, 8(27599-23175):41, 2001.

[156] Z Dong, Y Zhao, and Z Chen. Support vector machine for channel prediction in high-speed railway communication systems. In *2018 IEEE MTT-S International Wireless Symposium (IWS)*, pages 1–3. IEEE, 2018.

[157] Amnon S. Introduction to machine learning: Class notes 67577, 2009.

[158] E Obayiuwana and E Falowo. Network selection in heterogeneous wireless networks using multi-criteria decision-making algorithms: a review. *Wireless Networks*, 23:2617–2649, 2017.

[159] K Yadav and K Singh. The influence of different weighting methods on madm ranking techniques and its impact on network selection for handover in hetnet. In *2023 International Conference on Artificial Intelligence and Smart Communication (AISC)*, pages 959–963. IEEE, 2023.

[160] K Sethy, K Barpanda, K Rath, and K Behera. Deep feature based rice leaf disease identification using support vector machine. *Computers and Electronics in Agriculture*, 175:105527, 2020.

[161] T Dai and Y Dong. Introduction of svm related theory and its application research. pages 230–233, 04 2020.

[162] A Tharwat. Parameter investigation of support vector machine classifier with kernel functions. *Knowledge and Information Systems*, 61:1269–1302, 2019.

[163] S Ghosh, A Dasgupta, and A Swetapadma. A study on support vector machine based linear and non-linear pattern classification. In *2019 International Conference on Intelligent Sustainable Systems (ICISS)*, pages 24–28. IEEE, 2019.

[164] H Swain and H Hauska. The decision tree classifier: Design and potential. *IEEE Transactions on Geoscience Electronics*, 15(3):142–147, 1977.

[165] R Quinlan. Learning decision tree classifiers. *ACM Computing Surveys (CSUR)*, 28(1):71–72, 1996.

[166] B Hssina, A Merbouha, H Ezzikouri, and M Erritali. A comparative study of decision tree id3 and c4. 5. *International Journal of Advanced Computer Science and Applications*, 4(2):13–19, 2014.

[167] S Singh and P Gupta. Comparative study id3, cart and c4. 5 decision tree algorithm: a survey. *International Journal of Advanced Information Science and Technology*

(IJAIST), 27(27):97–103, 2014.

[168] G Guo, H Wang, D Bell, Y Bi, and K Greer. Knn model-based approach in classification. In *On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE: OTM Confederated International Conferences, CoopIS, DOA, and ODBASE 2003, Catania, Sicily, Italy, November 3-7, 2003. Proceedings*, pages 986–996. Springer, 2003.

[169] Z Zhang. Introduction to machine learning: k-nearest neighbors. *Annals of translational medicine*, 4(11), 2016.

[170] F. Darío Baptista, S Rodrigues, and F Morgado-Dias. Performance comparison of ann training algorithms for classification. In *2013 IEEE 8th International Symposium on Intelligent Signal Processing*, pages 115–120, 2013.

[171] I.A Basheer and M Hajmeer. Artificial neural networks: fundamentals, computing, design, and application. *Journal of Microbiological Methods*, 43(1):3–31, 2000. Neural Computting in Micrbiology.