UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

ON ANALYTICAL MODELING OF MOBILITY SIGNALLING IN ULTRA DENSE HETNETS

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY in ELECTRICAL AND COMPUTER ENGINEERING

By

AZAR TAUFIQUE
Norman, Oklahoma
2018

ON ANALYTICAL MODELING OF MOBILITY SIGNALLING IN ULTRA DENSE HETNETS

A DISSERTATION APPROVED FOR THE
SCHOOL OF ELECTRICAL AND COMPUTER ENGINEERING

BY

Dr. Ali Imran, Chair

Dr. Pramode Verma

Dr. Kam Wai Clifford Chan

Dr. Samuel Cheng

Dr. Timothy Ford

Acknowledgments

I would like to express my sincere gratitude to my advisor, Prof. Ali Imran, for his continuous support during my Ph.D. study and research. His expertise, guidance, support, and patience added considerably to my graduate research experience.

I would also like to thank my fellow colleagues in BSON Laboratory namely Hasan Farooq from University of Oklahoma, Tulsa,USA and Dr. Abelrahim Mohamed from 5G Innovation Center (5GIC), University of Surrey, England for their support and providing a conducive research and work environment.

I am extremely grateful to my family; my mother, my sister and grandfather for their prayers, love and moral support.

Table of Contents

List of Figures

List of Tables

Abstract

Multi-band and multi-tier network densification is being considered as the most promising solution to overcome the capacity crunch problem in emerging cellular networks. To this end, small cells (SCs) are being deployed within macro cells (MC) to off-load some of the users associated with the MCs. This deployment scenario gives birth to several new problems. Amongst others, handovers (HOs), signalling overhead and mobility management are becoming increasingly critical challenges. Frequent HOs in ultra-dense SC deployments can lead to a degraded mobility performance and increase signalling overhead significantly. Recently, a new cellular architecture with control/data plane separation has been proposed to overcome these challenges. However, the state of the art analysis of the feasibility of the CDSA remains mostly qualitative. There is dire need for mathematical models to analyze the performance of various aspects of CDSA and quantify its gains, if any, compared to conventional architecture. In this dissertation, we derive several analytical models to compare HO performance in the control/data separation architecture (CDSA) and conventionally deployed networks under various scenarios and configurations. Our developed mathematical framework advances the state of the art by considering HO success, HO failure and no HO scenarios. The proposed models can be used to quantify HO signalling as a function of key cellular system design parameter such as cell density, session duration, velocity, HO duration(s) and intercell overlap coverage factor. Using the developed analytical models, we perform a comparative analysis of HO signalling generated during various HO scenarios in CDSA and conventionally deployed networks. Building on the insights drawn from this analysis, we introduce new parameters for improving the HO execution process in emerging cellular networks viz-a-viz 5G and beyond. These new parameters, when tuned optimally, can significantly reduce the HO signalling load. Closed form expressions are also derived for continuous and continual (intermittent) mobility

scenarios, while considering both HO success and HO failure likelihoods. In addition, we propose an analytical model which enables more radio resource efficient network planning by quantifying HO signalling and success probabilities as function of intercell overlap coverage factor. Analysis indicates that cell density, actual HO time duration and average velocity can be used as the key metrics to optimally plan intercell overlap coverage factor in order to minimize mobility signalling load. Numerical results and analysis based on the developed overall analytical framework indicate that, compared to conventional networks, CDSA offers promising gains in terms HO performance and reduced HO signaling overhead.

List of Key Symbols and Acronyms

HO	Handover
P_{no}	Probability that no HO signalling will be generated
P_f	Probability that mobility signalling will be generated as a result of HO failure
P_s	Probability that mobility signalling will be generated as a result of HO success
c	Intercell Coverage Overlap factor
P_h	Probability that signalling will be generated as a result of HO success in case of continuous mobility
S	Normalized CN mobility signalling load on account of HO failure(s) and HO success(s)
CS	Normalized CN mobility signalling load on account of HOs for continuous mobility
S_f	Normalized CN mobility signalling load on account of HO failure
S_s	Normalized CN mobility signalling load on account of HO success
λ	Session Duration
$\lambda_{,r}$	Residual Session Duration
α	HO failure coefficient
α_i	HO failure coefficient for CBS_i
β	HO success coefficient
β_i	HO success coefficient for CBS_i
γ	No HO coefficient
γ_i	No HO coefficient for CBS_i
θ	Cell residence time
$\theta_{,r}$	Residual Cell residence time
$\theta_{i,r}$	Residual Cell residence time of control base station i
θ_1	Residual Cell residence time of Control base station (CBS)
θ_2	Residual Cell residence time of Data base station (DBS)
T_d	Mobility time duration during which HO takes place
T_p	Time taken for handover completion
ρ_1	Cell density of control base station
ρ_2	Cell density of data base station
v	Average velocity
L_1	Length of the control base station
S_1	Area of control base station
c	Coverage factor
MME	Mobility Management Entity
$TEID$	Tunnel Endpoint Identifer
$E-RAB\ ID$	Evolved Radio Access Bearer Identification
PGW	Packet Data Network Gateway
GTP	GPRS Tunneling Protocol

$GTP-U$	GPRS Tunneling Protocol user plane version
SGW	Serving Gateway
TA	Timing Alignment
CBS	Control Base station
DBS	Database station
LTE	Long term evolution
RRC	Radio Resource Control
$3GPP$	Third generation partnership
HO	Handover
CN	Core Network
CP	Control Plane
DP	Data Plane
UE	User Equipment
mmW	millimeter waves
BS	Base station

CHAPTER 1

Introduction

1.1 Mobility in Cellular Networks

The major cornerstone of cellular networks is the possibility for mobility which enables the users to move anywhere within the coverage area and still receive voice and data services. This century has witnessed an exponential increase in mobile data usage, around 400 million-fold over the past 15 years [1] - thanks to the proliferation of smart devices, rapidly growing volume and a variety of mobile applications. According to the latest visual network index (VNI) report from Cisco [1], global mobile data traffic will increase nearly 8-fold between 2015-2020 reaching 30.6 Exabyte per month by 2020. It has been estimated that 50 million base stations (BSs) will be deployed as soon as 2020 [2]. Moreover, higher densification is consequently going to increase the complexity of the network management [3]. The efficient management of such a convoluted network is a mounting challenge for network operators since Heterogenous networks (HetNets) not only require more adaptations and reconfigurations due to more cells, but also, they feature more adaptations and reconfiguration per cell [4]. Although these estimations are debatable, they give an indication of the situation in the near future. This unprecedented trend has prompted the need for a paradigm shift in future cellular networks to provide much higher capacity and Quality of Service (QoS) than their predecessors, driving the evolution of the 5th Generation (5G) cellular networks [5]. The race to 5G is on, and it is a general consensus among the researchers in both academia and industry that the densification, (e.g., in form of HetNets) is the approach with highest potential gain to achieve the ambitious capacity and QoS goals envisioned for emerging cel-

lular networks. Moreover, the bulk of target 1000x capacity gain in 5G has to stem from network densification, even if massive multiple-input multiple-output (MIMO) and mm Wave are adapted [6, 7, 8]. Network densification by small cells (SCs) is also imperative as approximately 80% of traffic over the entire network is generated indoors, of which 70% to 80% is carried over outdoor macro cellular networks [9].

Such massive deployments raise several problems in terms of signalling overhead, mobility management, energy consumption, capital and running costs, planning and scalability [10, 11, 12, 13, 14, 15, 16, 17, 18]. Most of these issues are tightly coupled to the radio access network (RAN) architecture which constitutes an integral part of cellular systems. With ultra-dense SC deployments, mobility management becomes complex because HOs will happen frequently even for low mobility users [19, 20, 21, 22, 23, 24, 25, 26]. In the conventional RAN architecture, the HO procedure includes transferring all channels (i.e., control and data) from one BS to another with a significant core-network (CN) signalling load. For instance, the results reported in [27, 28, 29] indicate high signalling overhead and call drop rates when the conventional HO mechanisms are applied in dense SC deployment scenarios. To solve this problem, a futuristic RAN architecture with a logical separation between control plane (CP) and data plane (DP) has been proposed in the research community [30, 31, 32, 33, 34, 35].

In the control/data separation architecture (CDSA) shown in Fig. 1.1, a few MCs, known as control base stations (CBSs), provide the basic connectivity services and support efficient control signalling. Within the CBS footprint, on-demand high data rate services are provided by dedicated SCs known as data base stations (DBSs). As shown conceptually in Fig. 1.1, all user equipment (UE) are anchored to the CBS, while the active UEs are associated with both the CBS and the DBS in a dual connection mode [29, 36, 30]. This configuration could offer simple and robust HO procedures because the radio resource control (RRC) connection is maintained

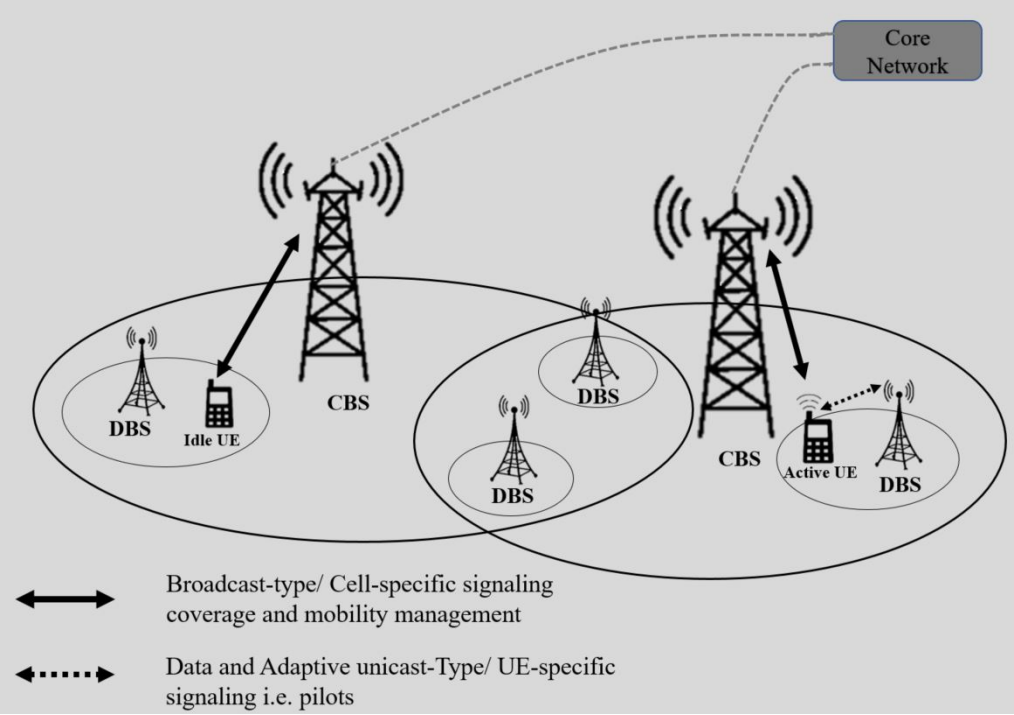

Fig. 1.1: Schematic for Control Data Plane Split Architecture

by the CBS (which is typically a MC). Thus, the UE is anchored to a BS with a large coverage area. As a result, the intra-CBS HOs (i.e., between DBSs under the footprint of the same CBS) will be transparent to the CN. This in turn alleviates mobility signalling and reduces the associated overhead. However, most of the work on CDSA in general and mobility signalling in CDSA in particular [29, 37, 38, 39, 40, 41, 42] provides a qualitative discussion rather than a mathematical analysis with quantitative results.

1.2 Research Objectives

In light of the above discussion in section 1.1, the research presented in this dissertation provides answers to the following questions.

1. How much mobility signalling load is generated in HetNets scenario compared to a conventional Network?

2. If small cell densification is evident for future network growth, how can we reduce mobility signalling load in HetNets?

3. Are there any existing parameters with which can we reduce the mobility signalling load, or can we introduce new parameters to reduce the mobility signalling ?

4. Signalling is generated in case of handover success and failure. How can we quantify the amount generated for both scenarios?

This dissertation addresses the aforementioned research questions by mathematically modeling the mobility Signalling in CDSA under a range of scenarios. Numerical validations are also carried out to find and validate the answers to the above questions. The key contributions of the dissertation are outlined in the following section.

1.3 Contributions

The contributions of this dissertation can be summarized as follows:

- A mathematical framework for handovers probability is derived by incorporating handover success, handover failure and no handover scenario. Current studies on mobility in Heterogeneous networks ignore the effect of handover failure [43] or do not even consider the probability of handover not occurring [29, 37]. On the contrary the handover analysis presented in this dissertation take into account realistic scenario that handover do not succeed with 100% probability and failure may happen as well. Different from prior works [43], the analysis also considers stationary user scenario i.e. when no handover happens.

- Analytical expressions for handovers probability in terms of inter-cell coverage overlap, cell residence time, handover duration and session duration are derived. These analytical expressions thus offer the first key step towards the analytical evaluation of realistic mobility signalling as a function of cell density, speed, session duration, handover time duration and intercell overlap coverage factor.

- Analytical framework to compute the amount of mobility signalling generated in the core network as a result of handovers is developed for both CDSA and conventional network. A recent prior work [43] investigates similar problem, but it does not take the handover failure scenario into account. This dissertation holistically investigates practical considerations which can result in handover failure, and in the case of such occurrence how much signalling is generated in that specific scenario.

- Mobility signalling analytical model quantifies the expected mobility core network signalling generated: when either a non-continuous mobility or continuous mobility scenario takes place. This analytical model for continuous mobility provides an upper bound on amount of mobility related signalling and can be used to model HO signalling generated by cars on highways, airplanes or trains or similar commute systems.

- New handover related parameters such as mobility time duration, time taken for a handover completion and intercell overlap coverage factor are introduced. When configured appropriately these parameters can reduce handover-related core network signalling. Also, for a given topology these parameters can decide the success or failure of a handover.

- Another important contribution we make in this dissertation is the answer to the question, "how intercell overlap coverage factor in emerging cellular can

be planned for minimizing handovers mobility signalling". State of the art networks are planned without much consideration of mobility signalling. This is because there are no models to quantify mobility signalling as function of intercell shared coverage overlap and thus optimally plan cells for minimal mobility signalling exist. In wake of extreme cell densification, this approach is likely to result in extremely poor resource efficiency. This dissertation for the first time quantitatively shows how mobility signalling is dependent upon intercell shared coverage overlap, and how intercell shared coverage overlap can be planned to ensure handover success and minimize handover mobility signalling.

- Finally, this dissertation provides several new and quantitative insights to researchers, vendors and operators whether multi-tier multi-band small cells ought to be deployed using current conventional architecture CDSA approach. The results confirm for the case of ultra-dense heterogeneous networks (HetNets) deployment CDSA ought to be used.

1.4 Dissemination and Publications

Throughout the course of preparation for this dissertation, several dissemination activities were carried out. These activities have resulted in following presentations and (accepted, pending) peer reviewed articles.

1. A. Taufique, M. Jaber, A. Imran, Z. Dawy and E. Yacoub, "Planning Wireless Cellular Networks of Future: Outlook, Challenges and Opportunities" in IEEE Access, vol. 5, pp. 4821-4845, 2017.

2. A. Taufique, A. Mohamed, H.Farooq, A. Imran and R.Tafazolli, "Analytical Modelling for Mobility Signalling in Ultra-Dense HetNets" in IEEE Transac-

tions on Vehicular Technologies , 2018.(Accepted)

3. A. Taufique and A. Imran, "An Analytical Model to Design Optimal Inter-Cell Overlap for Minimizing Handover Signalling" in IEEE Wireless Communication Letters (under review).

4. A. Taufique, A.Rizwan, A. Imran, Kamran Arshad and M.A. Imran, "Big Data Analytics for 5G Networks: Utilities, Frameworks, Challenges and Opportunities" in IEEE Access (Revision submitted).

5. A.Taufique, A.Mohamed, H.Farooq and A.Imran, " An Analytical Model for Handover Performance in Control Data Split Architecture and its Comparison with HetNets" in IEEE Globecomm 2018 (under review)

6. A.Imran, H. Farooq and A.Taufique, "Mobility:How AI can transform this bane into a blessing for future wireless networks" in IEEE Network (under review)

7. S.M.A.Zaidi, A.Taufique, H.Farooq and A.Imran, "Mobility Challenges in 5G Ultra-Dense Heterogeneous Networks: A Survey and Outlook " in IEEE Communication Surveys and Tutorials (under review).

8. S.M.A.Zaidi, A.Taufique and A.Imran, "On the Affect of Mobility and Cell Density on SINR in Emerging Ultra-Dense Cellular Networks " in IEEE Transactions on Vehicular Technology (under review).

9. A.Taufique, S.M.A.Zaidi, Hasan Farooq and A.Imran, "An Analytical Framework to Estimate the Affect of Mobility on SINR and Finite Handovers Signalling Load for Emerging Ultra-Dense Networks " (under review).

10. A.Taufique and A.Imran, "Handover Signalling Reduction using Control Data Separation Architecture (CDSA)" at Graduate Student Research and Cre-

ativity Day, University of Oklahoma, Norman, February 2017. *Nominated for Best Poster Presentation Award*

11. A.Taufique and A.Imran, "Handover for Less: Signalling Reduction using Control Data Separation Architecture (CDSA)" at Tulsa Research Forum, University of Oklahoma, Tulsa April 2017. *Nominated for Best Research in Engineering and Applied Research Award*

12. A.Taufique, " How Control Data Separation Architecture (CDSA) can cure the challenges in Emerging Cellular Networks " at Telecommunication Research Presentation ,University of Oklahoma, Tulsa, April 2016. *Nominated for Best Telecommunication Research Presentation*

13. A.Taufique and A.Imran, "Handover Delay and Signalling Load Reduction Using CDSA in 10 Minutes" at Graduate Student Research and Creativity Day, University of Oklahoma, Norman, February 2017. *Nominated for Best Video Presentation Award*

1.5 Organization

The dissertation is structured as follows. Chapter 2 presents the background and challenges in mobility management, and describes handover procedure and different phases involved in the handover management. This chapter provides background information required to understand the contribution, results and analysis presented later. Chapter 3 introduces the literature review and state-of-the-art work done in mobility management so far. Chapter 4 presents the mathematical model to evaluate probability of handover failure, probability of handover success and probability of no handover. Chapter 5 evaluates the mobility signalling generated during a HO by employing probabilities of handover from chapter 4. Chapter 6 discusses the

mobility signalling load evaluation in case of continuous mobility scenario. Chapter 7 presents the analysis and results on the role of cell overlap coverage area and how it can impact HO signalling in emerging ultra-dense networks. Chapter 8 presents the analytical model for HO signalling in case of finite number of HOs. Chapter 9 discusses the conclusions and future work and thus concludes the dissertation. In this chapter we also outline some possible directions for future work that can be build on the work presented in this dissertation.

CHAPTER 2

Background

2.1 Handover and the state of the art

The distinguishing factor of cellular networks from other networks such as wired networks and fixed wireless networks, is the support of seamless mobility. Cellular network users are mobile, meaning the users are free to roam from one place to the other seamlessly, no strings attached. The cost for this seamless connectivity comes at the price of complex communication protocols and significant signalling overhead among the massive deployment of base stations (BS) and small cells (SC), so users can receive signal coverage on the go [44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54]. The underlying procedure which makes the movement of users from one cell to the other possible is called Mobility Management or simply mobility. Mobility is a broad term as it can mean different terms depending upon the context. It is mainly classified into two categories :

1. Connected Mode Mobility
2. Idle Mode Mobility

Connected mode mobility includes scenarios such as when a user is active on a phone call or data session while moving from one base station tower to another base station tower. In common context, this simple procedure of moving from one base station to another base station is termed as Handover, as shown in Fig. 2.1. On the surface handover seems to be a simple procedure of switching from one tower to another. However, in reality handover is an intricate set of processes running

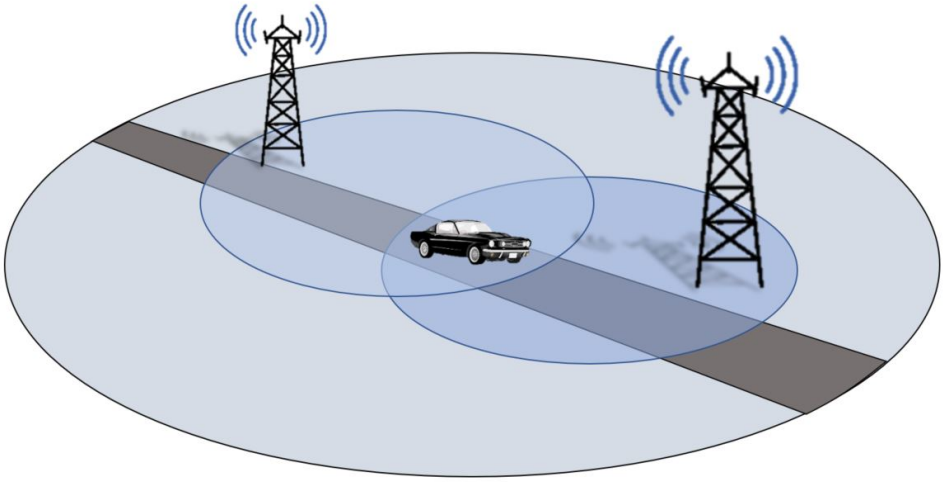

Fig. 2.1: Handover concept illustration on a high level

beyond the radio access network, as a user moves from one cell to another cell. For Connected Mode Mobility, LTE handover consists of three distinct phases [55].

1. Handover Preparation Phase
2. Handover Execution Phase
3. Handover Completion Phase

Handover process in LTE is termed as hard handover, which means that it has to break the wireless connection first, then re-establish the connection after handover to a new cell. Thus it impacts the user experience in the network.

2.2 LTE Entities

To fully explain handover, it's logical to first introduce basic entities in a cellular network that are involved in mobility. In order to do that a high level architectural overview of cellular network considering 4G LTE as an example is shown in Fig. 2.2. The entities invovled in mobility are as follows:

- Base station (eNodeB): It takes cares of user mobility in connected mode. eNodeB is the middle man between user equipment (UE) and the core network (CN). UE is responsible to send mobility related measurement to eNodeB. After reading the measurement reports eNodeB decides if a handover is needed or not.

- Mobility Management Entity (MME): It is responsible for taking care of mobility management of the UE in idle mode, session establishment, setting up of bearers and security procedures. MME is the controlling node in the CN and is responsible for coordinating handover phases in connected mode.

- Serving Gateway (SGW): It is responsible for setting up user plane and also acts as local anchor for mobility for the UE in connected mode. User data is forwarded by SGW to eNodeB and from eNodeB finally to the UE.

- Packet Data Network Gateway (PGW): It is responsible for allocation of IP address to the UE and connection to the external network from LTE network.

- S1: It is the interface between radio access network (RAN) and core side of 4G LTE network. This S1 interface is divided logically between S1-User and S1-MME interface. S1-User interface carries user data between eNodeB and SGW whereas S1-MME carries control signalling messages between eNodeB and MME as shown in Fig. 2.2 .

  If S1 interface is used for sending handover signalling and handover preparation messages between two eNodeBs then handover is termed as S1 Handover.

- X2: It is the interface between two eNodeBs. This X2 interface can carry both user traffic control and control information messages between two eNodeBs as shown in Fig. 2.2. If X2 interface is used for sending handover signalling and handover preparation messages between two eNodeBs then handover is termed as X2 Handover.

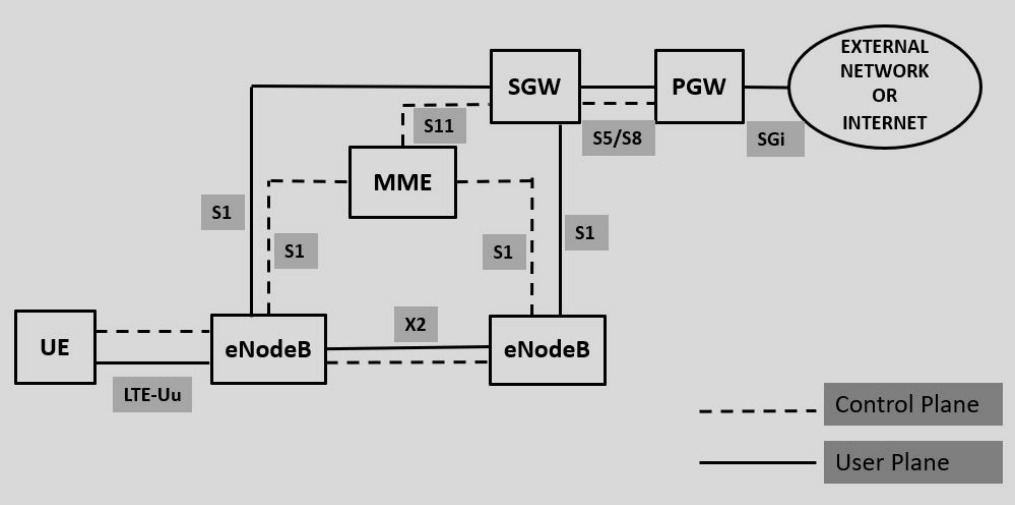

Fig. 2.2: High level overview of Cellular network considering 4G LTE as an example

2.3 Handover Types in LTE

Handovers can be classified by the target system, frequency, or by the method they are performed. In case of LTE, handovers can be divided into intra-LTE handovers and inter-LTE handovers. (Inter LTE mobility also includes inter-working with 2G/3G). These types of handovers are addressed as inter Radio Access Technology (inter-RAT). This dissertation is focused on intra-frequency cases and intra-LTE, so explain intra LTE handovers first. Intra LTE handovers include transitions to the same or different carrier frequency inside an LTE system. These can further be classified into following cases:

- Intra eNodeB handover refers to a case where the source and target cell reside in the same eNodeB. In this case no X2 procedure is required for the handover.
- Inter eNodeB handover depicts a situation where the two target cells are located in two different eNodeBs. This case assumes that MME will not change as a result of handover. S-GW may or may not be relocated. X2 or S1 needs to be initiated.
- Inter eNodeB handover with MME change is another type of handover. X2

handover process cannot handle X2 on MME relocation, so S1 procedure must be used instead. X2 and S1 procedures are discussed later in this chapter.

2.4 X2 and S1 Handover Types based on Signalling Interface in LTE

The handover architecture and implementation has changed radically in LTE compared to legacy 3GPP technologies. For example, in previous technology WCDMA has a radio network controlling element known as radio network controller (RNC) which possesses the necessary intelligence and signalling capabilities to handle the handover. In case of 4G LTE, RNC has been removed and all the intelligence has been pushed down to the eNodeB. In case of LTE, eNodeB is the only element deciding on and implementing handovers. eNodeBs have to signal with each other to perform the handover. This signalling for handover is achieved either through X2 interface or S1 interface. Therefore, based on the interface used, handover could be X2 or S1 based.

2.5 X2 Handover

In the case of X2 handover, the signalling connection requires that the two eNodeBs have X2 interface configured. If X2 interface is missing or not configured between two eNodeBs, it is not considered X2 handover. In case of LTE, the handover is termed as hard handover. This means that the air interface to the source eNodeB is dismantled before the new connection to the target eNodeB is built up. Therefore loss of data during the detach time is a problem. To prevent the packet loss, LTE uses data forwarding from the source eNodeB to the target eNodeB during the handover process. As soon as the source eNodeB has sent the handover command to the UE, it starts to forward the packets received from the S-GW towards the target eNodeB. The target eNodeB buffers the incoming packets, and starts sending

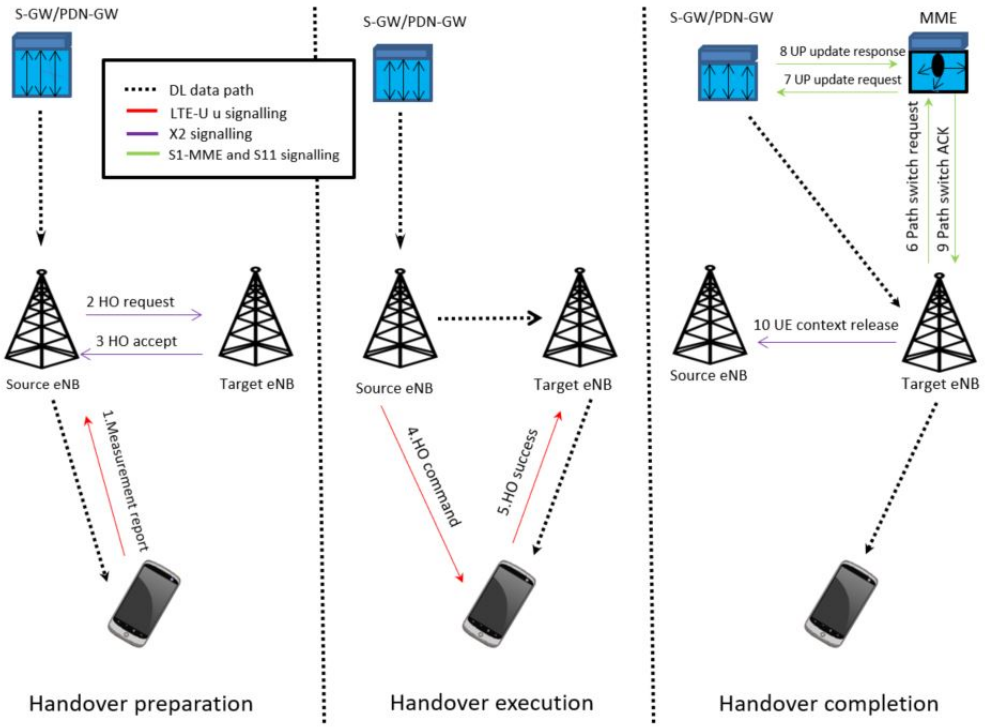

Fig. 2.3: Intra-frequency X2-based handover

them to the UE after it has completed the radio handover. At this point the MME or S-GW are not aware that a handover has occurred. The S-GW is still sending the DL data to the source eNodeB, even though the UE is already connected to the target eNodeB. The UE still gets the data through the forwarding process as shown in Fig. 2.3. In order to change the user plane path to flow directly to the target eNodeB, the target eNodeB sends a path switch request to the MME. The MME then asks the S-GW to change the endpoint of the GTP-U tunnel to the target eNodeB. This is called late path switching, since the actual handover has already been performed before the DL data path is updated. Finally, the target eNodeB informs the source eNodeB that handover and path switching has been successfully completed. Upon this notice, the source eNodeB may drop any context it has still kept for the UE.

2.6 S1 Handover

An S1 handover is necessary if the MME is to be changed because of the handover. Generally this happens only in MME area limits. S1 handover may also be initiated if for some reason an X2 interface is not available. The control signalling will then flow through S1 interface. The S1 handover possibility is useful, since it allows for a handover to complete regardless of possible missing X2 definitions. S1 handover procedure is slightly more complex than the X2 handover, since the MME has to act as an intermediary coordinator and message relay between the source and target eNodeBs. In addition to relaying the messages, the MME also configures the data forwarding process. MME sends the required handover details received from the target eNodeB, as well as the information about the S-GW, to which the source eNodeB is supposed to forward the downlink packets during handover. The S-GW used for forwarding is usually the same, however it could be a different S-GW as well. The MME may decide to use a different S-GW altogether. When the UE has successfully completed the radio handover, the eNodeB notifies the MME about the event. This triggers the path switch procedure, and in the future the data will flow directly to the target eNodeB. At this point a resource timer is also started on the MME. Upon the expiry of the timer, the MME releases the UE context from the source eNodeB and removes the forwarding tunnel from the S-GW as shown in Fig. 2.4.

The UE cannot tell the difference between an X2 based and S1 based handover, since the radio handover is completed alike in both situations. However the user may be able to notice the difference in the data pause. This is because the data forwarding path of an S1 handover can be considerably longer. In the best case, the X2 forwarding route consists of only a single switch. The forwarding path in the S1 handover case flows through an S-GW, which is likely to be further away from

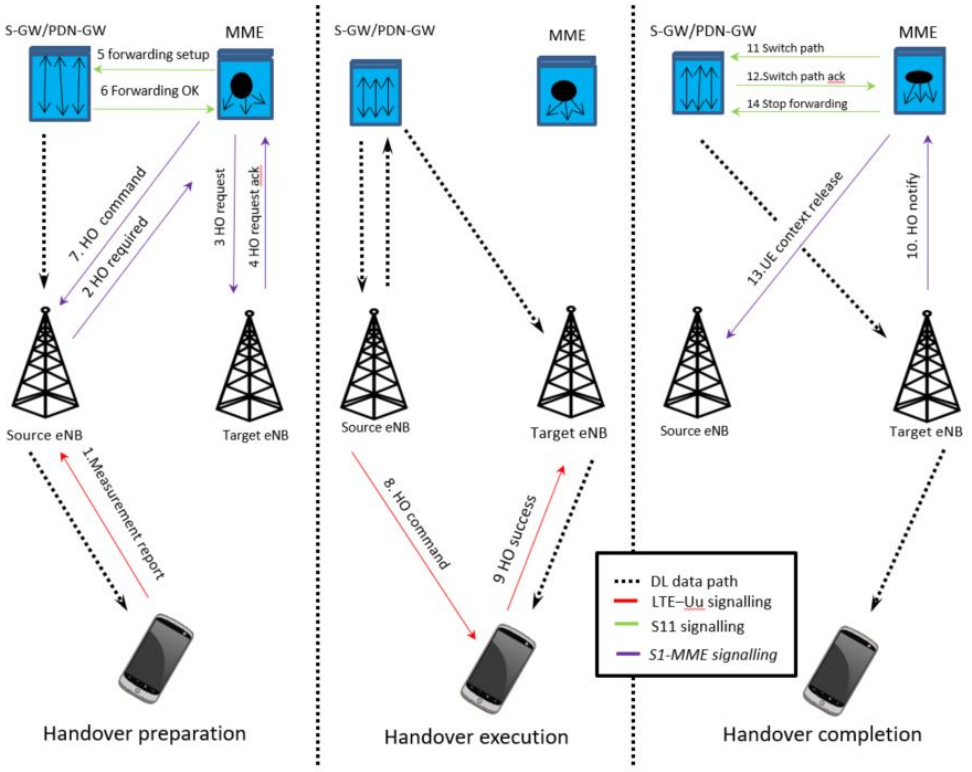

Fig. 2.4: Intra-frequency S1-based handover

the eNodeBs. In this dissertation we focus on S1 handover.

2.6.1 S1 Handover Signalling Call Flow

Fig. 2.5 presents a full signalling flow of an S1 handover. The signalling flow assumes that neither MME nor S-GW are relocated. A more detailed signalling flow with different phases of handover is explained later in the chapter. The enumeration below describes the steps in the Fig 2.5.

1. The UE measurement results trigger a reporting event, and it sends the measurements to the source eNodeB.

2. The source eNodeB decides that handover should be performed. It notices that no X2 interface to the target eNodeB exists, and as a result initiates an MME assisted handover.

3. The source eNodeB sends a message to the MME indicating that a handover is required. This message contains a transparent container meant to be forwarded by the MME to the target eNodeB. The message also includes information on whether the X2 interface is available for data forwarding, and the identities of the target eNodeB and the target tracking area identity (TAI). The target TAI is used by the MME to determine whether the MME needs to be changed or not. The existence of X2 is useful in cases with MME relocation, since the data forwarding may then be done through X2.

4. The MME forwards the transparent container to the target eNodeB along with information about the needed bearers for data and signalling, as well as possible handover restrictions.

5. If the target eNodeB deems that it has necessary resources for the handover, it establishes a context for the UE. The target eNodeB then sends an acknowledgement to the MME. The acknowledgement includes information about the successfully setup bearers and possible forwarding parameters. The message also includes mobility control information in a transparent container, which is sent to the UE at a later stage.

6. The MME sets up the data forwarding function with an indirect data forwarding tunnel request with necessary transport layer identifiers.

7. The S-GW acknowledges the forwarding, and sends the identifiers of its own to the MME.

8. The MME sends the Handover Command to the source eNodeB. This command contains information about the bearers which are to be forwarded during handover. This information along with the target to source transparent container, including information the UE uses to attach to the target cell.

9. As with X2, the source eNodeB sends the *RRCConnectionReconfiguration* with *mobilityControlInformation* to the UE.

10. After sending the Handover Command to the UE, the source sends the eNodeB status transfer to the MME. As in the X2 case, this message includes the PDCP status, preserving the sequence numbering to prevent unnecessary retransmissions.

11. The source eNodeB status is forwarded to the target eNodeB.

12. The UE synchronizes to the target cell.

13. To notify the eNodeB that the handover has been completed, the UE sends a *RRCConnectionReconfigurationComplete* to the target eNodeB.

14. As the eNodeB realizes that the UE has successfully attached itself, the eNodeB notifies the MME.

15. Upon the reception of the Handover Notify, the MME starts a resource release timer. Upon the expiration of this timer, the MME will release the context from the source eNodeB, as well as dismantle the forwarding setup from the S-GW. This happens in steps 19-22, which are not elaborated here for the sake of conciseness.

16. The MME sends a bearer modification request to the S-GW. The purpose of this message is to switch the path towards the target eNodeB.

17. The S-GW switches the path and acknowledges the bearer modification.

2.7 Handover Phases and Associated Procedures

As discussed in previous sections the handover procedure consists of 3 phases.

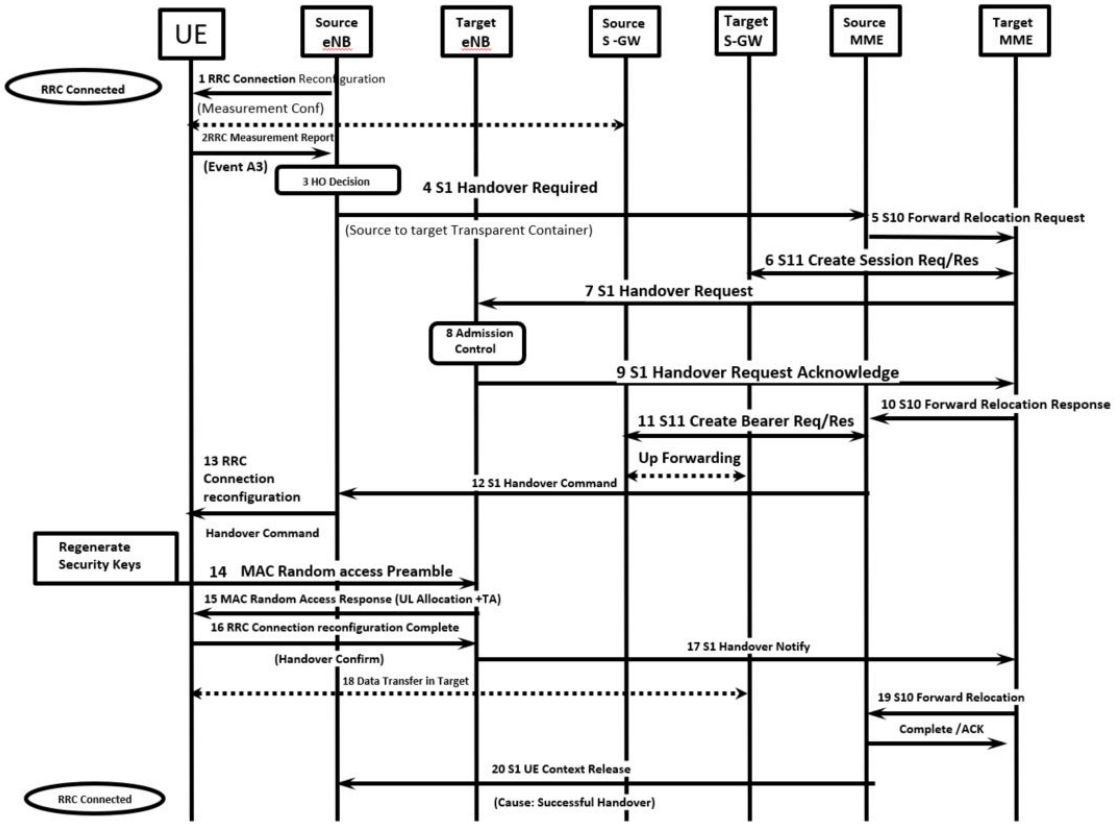

Fig. 2.5: Ladder Diagram for Inter eNodeB S1 handover with MME relocation

- Handover Preparation Phase
- Handover Execution Phase
- Handover Completion Phase

2.7.1 Handover Preparation Phase

During handover preparation phase, an active UE looks for a neighboring cell to find out if it is better than the serving cell. This is done by measuring signal strength, either through reference signal received power (RSRP) or reference signal received quality (RSRQ). The neighbor is better than the serving cell, the UE will trigger a measurement report to the serving cell including a target cell identification , (i.e., Physical Cell Identity (PCI) and signal strength measurement). Once the measurement report is received the handover decision is made by the serving base station. Once a decision is made, a handover request is sent to the target base station. If the target base station is already a neighbor of the serving cell via X2 interface, this request will be sent via X2 interface. Otherwise, the request will be sent via S1 interface, which will result in core network signalling load. Within this handover request security information, UE related information such as UE security capabilities, Quality of service Class Identifier (QCI) value, Evolved Radio Access Bearer Identification (E-RAB ID), S-SGW Tunnel Endpoint Identifier (TEID) and related information is included. Once the target cell receives this information, it will perform admission control, such as verifying if it has capability to accommodate the incoming handover or not. If the target cell can accommodate the handover, it will send a handover acknowledgement to the serving cell via X2 or S1 interface, whatever the case may be. The target cell reserves RRC resources and allocates a new cell Radio network temporary identifier (C-RNTI). Once the handover request acknowledgement is sent within the acknowledgement, it also includes E-RAB

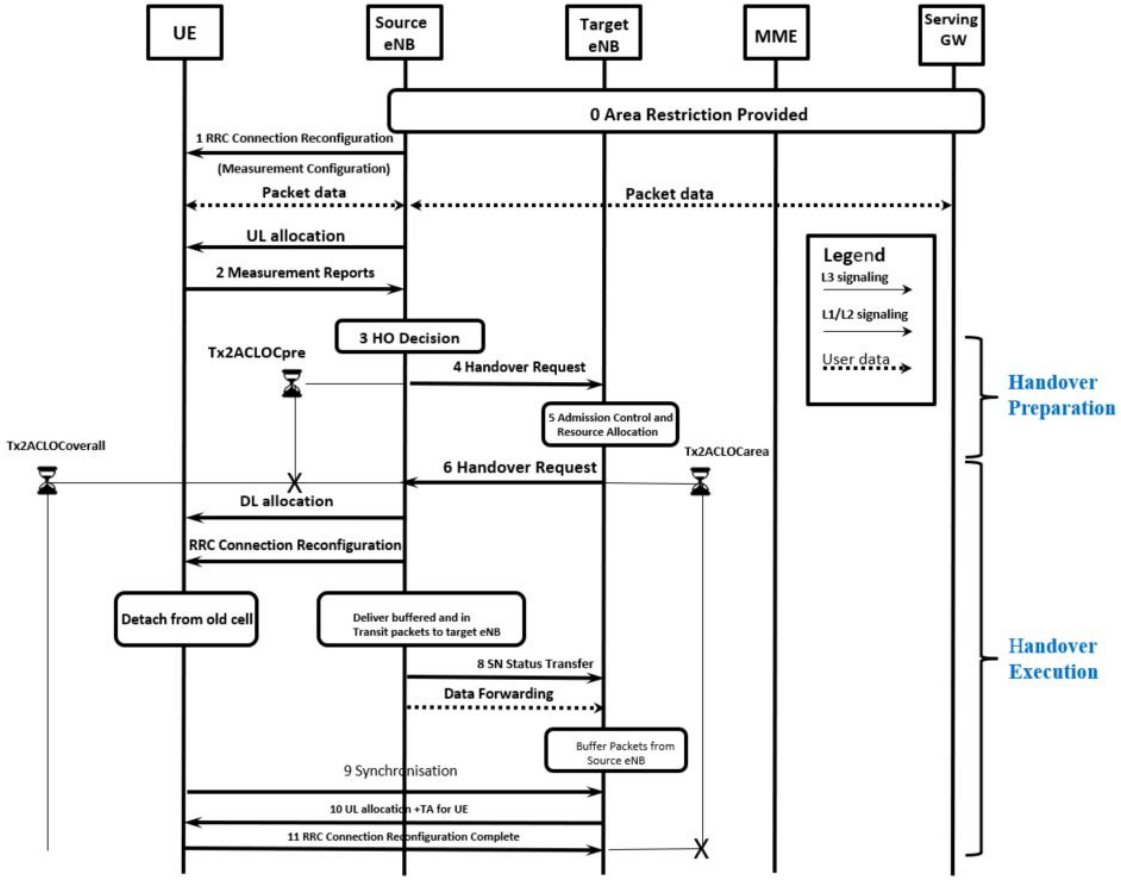

Fig. 2.6: Intra-frequency X2-based handover (Part 1)

ID, Target eNodeB TEID and target base station security algorithm related information. In case of an X2 handover, upon receiving the acknowledgement, an X2 Transport bearer downlink (DL) establishment is setup between the source and target base station. During this handover transition phase, data can be forwarded from the source to the target BS while the UE is not completely connected to the target base station. In case of S1 handover, there will be no data transfer over X2, unless it is indirect forwarding. However, if there is support of direct forwarding, data will be forwarded from source SGW to target SGW. Direct forwarding will result in CN signalling load during handover preparation phase. Handover preparation phase is shown in Fig. **??**

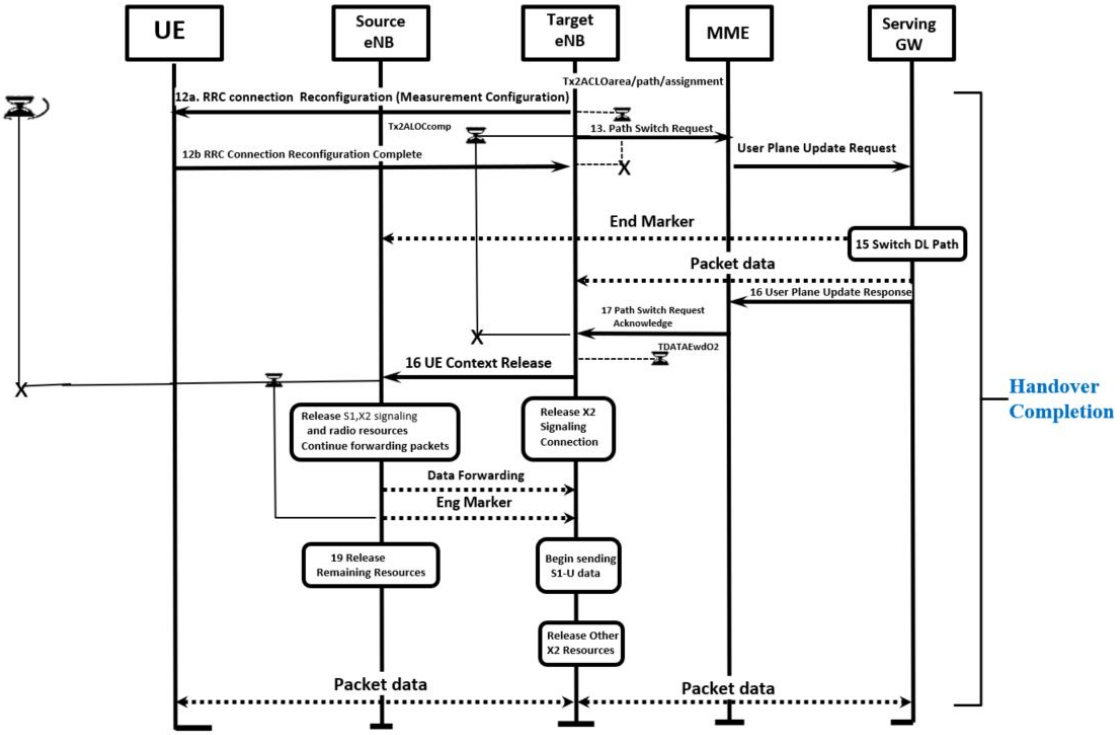

Fig. 2.7: Intra-frequency X2-based handover (Part 2)

2.7.2 Handover Execution Phase

Once preparation phase is successful, handover execution phase has to take place. In handover execution phase, UE is given a command to detach from source base station and get connected with the target base station. For S1 handover scenario, this command will come from MME towards the source base station. Source base station will forward this command to the UE. UE will detach from source cell and moves towards the target cell. While UE is in the process of getting connected to the target base station, it has to go through uplink synchronization and random-access procedure. Once UE gets connected to the target base station, it starts to receive data which was stored in the buffer during handover transition. Even though UE gets connected to the target base station, handover is not complete yet.

2.7.3 Handover Completion Phase

Handover completion phase starts when *RRCConnectionReconfiguration* message is sent by the target eNodeB to the UE. After path switch request is complete, UE will send RRC Connection Reconfiguration Complete message back to the target eNodeB. MME sends the update user plane request to the serving gateway. After downlink path switch request is completed by sending End Market packet to the source eNodeB, target eNodeB will send a UE Context Release command to the source eNodeB. The procedure for HO Completion phase and signalling is shown in Fig. 2.7

Signalling flow and procedure of X2 handover is shown in Fig. 2.6 and Fig. 2.7 respectively whereas S1 handover with MME relocation procedure is shown in Fig. 2.5. In order to identify the overall signalling load on account of handover in the core network for S1 handover, we use the information shown in Fig. 2.8. The details of the messages are shown in Table 2.1. *We derive, evaluate and analyze an analytical framework to evaluate this signalling load in this dissertation.*

2.8 Possible Approaches for Reducing Handover Signalling

In the previous section handover phases are explained. Even if a handover is successful or not, it still adds a burden of signalling load in the core network on account of handover preparation, execution and completion phases respectively. If a handover is unsuccessful, it results in additional signalling compared to a successful handover. Handover is unsuccessful if any of the handover preparation, execution or completion phase is unsuccessful. A handover preparation phase can be unsuccessful for the following reasons.

- RF configuration issues

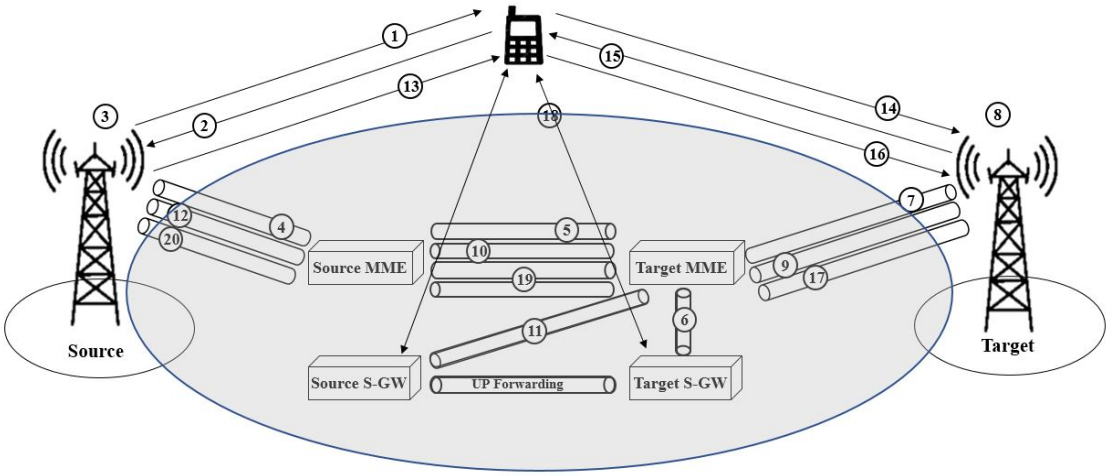

Fig. 2.8: Signalling messages exchanged during a typical S1 Handover

- Missing handover links
- Handover to target cell is not possible (No X2 established)
- Miscellaneous

Handover execution can be unsuccessful for the following reasons.

- Over shooter cell in the neighborhood
- Poor tuning of handover parameters
- RF issues
- Miscellaneous

One way to avoid the excessive signalling wastage in the case of an unsuccessful handover is to optimize and configure the network to avoid the possibility of the above reasons to occur. Hence, handover preparation and execution phase takes place successfully by avoiding any chances of failure. The other approach is to quantify how much signalling is generated when a handover takes place. Once we can see how much signalling is generated during handover success or failure, it can

Table 2.1: HO Messages

Number	Description
1	RRC Connection Reconfiguration
2	RRC Measurement Report
3	HO Decision
4	S1 Handover Required
5	S10 forward Relocation Request
6	S11 Create Bearer Request/Response
7	S1 handover request
8	Admission Control
9	S1 Handover Request Acknowledge
10	S10 Forward Relocation Response
11	S11 Create Bearer Request/Response
12	S1 Handover Command
13	RRC Connection Reconfiguration
14	Random Access Preamble
15	Random Access Response (UL Allocation + TA)
16	RRC Connection Reconfiguration Complete
17	S1 Handover Notify
18	Data Transfer in Target
19	S10 Forward Relocation Complete/ACK for control base station
20	S1 UE Context Release Command

be a good gauge to compare CDSA architecture with conventional architecture. In addition, the signalling quantification model can also be used to introduce new parameters and design changes in the HO protocol to reduce the HO signalling as well HO failure probability.

2.9 Factors determining Signalling Load in Handover Failure

The total CN signalling load generated during a handover failure depends upon a number of factors such as:

- UE speed and mobility
- BS density

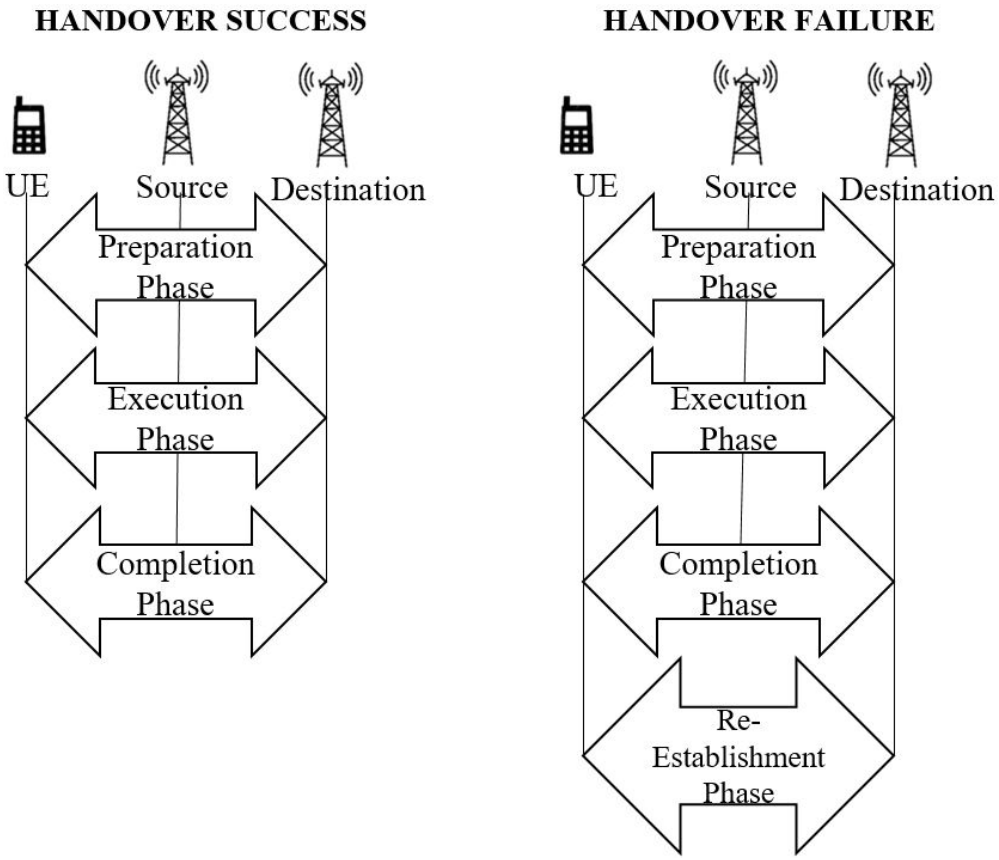

Fig. 2.9: Comparison of HO Success and HO Failure Signalling and Phases

- Session duration
- Transport network reliability (stability)
- Miscellaneous

2.10 Handover Failure Signalling versus Handover Success Signalling

In this section we compare the HO failure and HO success procedure as shown in Fig. 2.9. The associated procedure and respective messages are shown for HOs side by side.

2.11 HO Failure generates more signalling than HO success

Looking at Fig. 2.9, the signalling for handover success and failure can be compared, assuming same amount of signalling is generated in preparation, execution and completion phase of a handover. In order for a handover to fail, it can fail at any of the three phases independently. For a UE to remain in the system after HO failure it has to perform RRC connection re-establishment as shown in Fig. 2.9.

Let us say:

- S_f = Mobility Signalling generated in case of HO failure
- S_s = Mobility Signalling generated in case of HO Success
- S = Total Signalling generated as a result of HO Success and Failure
- $S = S_f + S_s$

Assuming all phases and messages carry equal weight. From Fig. 2.9, RRC re-establishment is equal to 0.25 of the complete HO procedure.

$S_f = S_s + \mathbf{0.25}\ S_s$

$S_f = \mathbf{1.25}\ S_s$

$S = \mathbf{1.25}\ S_s + S_s$

$S = \mathbf{2.25}\ S_s$

2.12 Assumptions

In order to derive the analytical model(s) of mobility signalling, the following assumptions are made in this dissertation.

- The user remains in the system upon handover (HO) success or HO failure. In case of HO failure, UE will remain connected to the CBS but will require RRC connection re-establishment with the DBS.

- Different amounts of signalling are generated in CN in the case of HO failure and HO success. Specifically, a HO failure event generates more signalling than a HO success event as shown in Fig. 2.9.

- A user can remain within the same CBS and not perform an inter-CBS HO with probability P_{no}. It can perform an inter-CBS HO from one CBS to another CBS. Inter-CBS HO can be successful with probability P_s, or it could be a failure with probability P_f.

- An LTE system with equal number of low and high mobility distributed users are considered. The term *sector* is used with the same meanings as a cell.

- HO failure is considered on account of too late HO. These assumptions are valid for CDSA in ultra-dense networks, as with densification, more HO failures may take place because of too late scenario if HO parameters are not tuned accordingly.

- HO failure can take place due to various reasons other than too late HO, such as transport network reliability i.e., S1 interface is down, poor RF conditions, radio link failure and partial HO etc. An analytical model for HO failure due to reasons other than too late HO can be derived accordingly. For the case of HO failure caused by poor RF conditions, radio link failure is triggered when the downlink signal to interference noise ratio (SINR) is below a certain threshold (Qout = - 8 dB) and stays below $-$ 6 dB for at least 1 sec [56]. Using this approach, probability of SINR greater than the threshold can be computed. If SINR threshold is less than the threshold for a given time, it will be a HO failure and vice versa. Similarly, HO failure caused by partial

HO can be characterized by calculating the probability of whether all bearers get transferred completely or not. By computing the probability of all bearers transferred or not, we can compute probability of HO success or HO failure respectively. For a more detailed discussion on possible HO failure scenario, reader is referred to [56].

- Regardless of the HO failure reasons, all factors contribute to the same amount of CN signalling load.

2.13 Summary

In this chapter, details and background about the handover procedure in the cellular network was provided. It was shown how many types of handovers exist in LTE. Different types of handovers based on signalling interface were also explained. Comprehensively detailed discussions were provided for X2 and S1 type handover, including signalling flow and ladder diagrams.

After providing the necessary theoretical foundation to the reader regarding handover in LTE, three important phases of a handover were discussed. In order for a handover to be successful, all three phases need to be successful. If any of the phase is not successful, it results in handover failure. Therefore, we have compared the difference between Handover success and failure signalling.

The conclusion of the comparison was that HO failure results in more Signalling than the HO success. This conclusion is a key motivation behind this dissertation as a few recent studies do attempt to quantify HO signalling load, do not consider HO failure as we do in this dissertation. Finally, the chapter was concluded by clearly spelling out the assumptions made behind the analysis presented in this dissertation.

CHAPTER 3

Literature Review

3.1 Introduction

In order to make the HO procedure seamless and have the least data interruption to the user, there have been a number of approaches proposed in literature. In this chapter we review a number of approaches proposed in prior literature about HO optimization. We will conclude this chapter by performing literature review on HO signalling optimization.

3.2 Handover Optimization

The high throughput requirement, heterogeneity of UEs and BSs, and security awareness of upcoming 5G environments appeal for a fast, distributed and privacy preserved HO scheme. Traditional HO optimization approach [57, 58, 59, 60, 61, 62, 63, 64, 65, 66] is inappropriate in 5G networks due to densification, dynamic channel conditions caused by higher frequency bands such as mmWave and different mobility patterns in wake of autonomous cars and intellignent transport system. On one hand, a large number of BSs tightly packed in the coverage area makes it extremely difficult to configure and maintain HO parameter optimization using existing manual techniques. On the other hand, small-cells can be frequently powered on/off [67, 68, 69, 70, 71], channel condition changes drastically for mobile UEs, and neighboring cell list changes accordingly. Authors in [72] use optimization theory to achieve maximum data rate with minimum blocking probability in a heterogeneous environment. They suggest a user centric approach is more suitable in hybrid

5G environment than network centric approach. Users are divided into HO and non-HO users and ensure minimum throughput of 2 Mbps is achieved.

3.2.1 Make Before Break Handover

HO in LTE is hard, where UE has to break the connection with the serving BS before resuming the new connection with the target BS, and in this process undesirable yet intrinsic service interruption is experienced. Authors in [73] propose a 3G like soft handover approach where multiple serving cells are represented by an Active Set (AS). They point out that fixed AS window can prevent radio link failures to a great extent, however throughput degradation is observed as radio resources of the weaker cells are unnecessarily wasted by the user. To counter this, they proposed dynamic AS window where add/remove parameters are linearly adapted based on slope of the linear curve that creates the dependency between add/remove offset and size of AS. Author in [74] discusses the pros and cons of make-before-break HO, and conclude that they are unsuitable for 5G networks. Authors of [75] pointed out that the inherent principles of make-before-break approach is to assume that the link to the source cell will be robust upon triggering the HO. However, realistically, HO takes place when radio condition is degrading, in addition, interference from target cell results in low reliability and an increase in number of retransmissions. Hence, early HO should be devised to benefit from the multi-connectivity situation. Another drawback of make-before-break is the complexity at UE side to process multiple RF and processing chains in parallel, resulting in handling two or more protocol stacks simultaneously. Even though this approach helps us to attain 0 ms latency requirement of 5G, but at the cost of substantially increased complexity. This in turn will impact the cost of UE in addition to the impediment of the deeper penetration

through IoT. Owing to the aforementioned intricacies, 3GPP RAN WG2 during its meeting number 94 (May 2016) decided to discard make-before-break like procedures from the scope.

3.2.2 Pagingless Approach

Authors in [76] presented a novel frame structure with sub-millisecond subframe duration operating in Time Division Duplex (TDD) mode aimed for 5G networks. The frame structure carries UL beacon resources constituting the basis of mobility and user tracking, which enables a pagingless system. However, this approach can lead to excessive amount of uplink messages which, in turn, may yield UE battery drainage, which is contradictory to one of the main 5G prerequisites.

3.2.3 Handover in mmWave Band

Traditional HO is Received Signal Strength (RSS) based, whereas pilot signal strength dictates cell-edge and offers the assistance to perform HO to target cell. Such a rule appears rudimentary and ineffective for addressing the unique challenges of the emerging 5G networks with millimeter wave (mmWaves). The mmWaves small cells pose an unprecedented challenge to mobility with the addition of narrow beams due to high penetration and propagation losses. The RF reception changes drastically with UE speed, hence it is difficult to provide stable service to high speed users. Authors in [77] suggest a novel Inter-Beam HO Class (IBHC) concept combined with HO control and radio resource management functionalities. Initially user is assigned into a mobility class, and corresponding HO frequency is defined such that pedestrians will observe more HOs than higher velocity UEs. The mobile user is assigned a group of beams as per mobility class, load conditions and expected path of

UE. Each beam in the group contains similar resource allocation to improve reception quality. HO is thus performed only at the edge of beam-group. The assumption that the individual signals of each beam are perfectly synchronized can be true for low speed users, however it may not suffice high speed users' requirement. RSS-based association leads to an unbalanced load condition where users tend to concentrate on the strongest cell. Moreover, overly frequent HOs between adjacent BSs may be observed, which in turn increase the overhead and delay of re-association. Angela Sara in [78] demonstrates a mobility-aware user association strategy for mmWave network, to overcome the limitations of the conventional RSS-based association. These authors' proposal exhibits several attractive features, such as: a) tracking channel condition after mobility, b) considering load condition in small-cells to prevent HO to high load cells, thus providing load-balance feature, c) preventing recurrent HOs, d) applying fully distributed approach (i.e., each mobile user associates to a BS independently of other UEs). Smart mobility as required by 5G networks can be achieved through the proposal, however, the HO procedure aimed for mmWave band has not been addressed. In addition, accuracy of dedicated positioning system, e.g. Global Positioning System (GPS), is a big concern especially for indoor situations.

High Speed Train Users: A novel concept for train communication using 60 GHz involves the concept of Moving Cell [79]. To avoid a large number of HOs in high speed train, the authors propose to employ Radio over Fiber (RoF) technique to make the serving cells move together with the train, and ultimately provide smooth uninterrupted transmission to passengers. However, for this scheme to be practical, the train's velocity and direction needs to be known beforehand in order to obtain synchronization and provide adaptation to the passenger's speed. Furthermore, due to inability to cope up with

randomness, this concept is not appropriate to indoor environments.

3.2.4 SON Based Optimization

Conventional RSS based HOs take effect with Time-To-Trigger (TTT) fulfillment to counter fast fading and ping-pong HOs. In one study, sensors were deployed at critical crossing and self-organization techniques were used to improve HO efficiency using vehicular traffic data gathered in London [80]. HO is triggered for HetNets when received power (P_{rx}) of Femto cell plus signal to noise to ratio (SNR) greater than received power of of Macro cell, satisfies over a period of time i.e., time to trigger (TTT). HO decision is further assisted with pathloss, based bias and an additional bias to counter fading. The technique uses knowledge of traffic speed to adapt femto bias and mitigate inefficiency caused by TTT while preventing ping-pong. The proposed method can dynamically manage MME loads to reduce HO completion times and prevent ping-pong. Performance benefits can increase as integration of smart city data in network optimization progresses: e.g., data about public transport networks such as buses/subway can be exploited to improve accuracy.

3.2.5 Improving Mobility Robustness Optimization (MRO)

Authors in [81] argued that optimizing HO Margin (HOM) is more effective than TTT. They proposed Dynamic Fuzzy Q-Learning algorithm based on fuzzy logic controller for mobility robust optimization (MRO) in HetNet. Q-learning is reinforcement learning to provide optimal HOM using inputs of a) number of HO success cases, and b) number of dropped calls to estimate user experience. The proposed algorithm can decrease HO ratio while maintaining call-drop ratio low. However, the procedures and priorities for HO towards

small-cells are not defined. Context-aware MRO is evaluated in [82] by taking into account speed attributes of a UE and speed estimation error. Firstly, UEs are grouped based on their speed, and then MRO is applied on individual groups. Context-aware mobility robustness considerably reduces HO failures, but it has limitations in supporting ultra-high reliability applications. On the contrary, Multi-Connectivity (MC) supports ultra-high reliability applications at the expense of increased signalling overhead.

3.2.6 Base Station Clustering

Understanding HO behavior of cells is a rewarding, yet challenging task. Authors in [83] study HO behavior of cells and propose a clustering model using machine learning to group cells with similar HO behavior. Further evaluation was done on actual HO KPI of close to 2000 WCDMA cells. The idea is to forecast number of HOs, detect abnormal HO, and respective optimization to be done thereafter. Centralization of RAN e.g. using Cloud-RAN offers numerous advantages including better mobility management. Uladzamir et al. [84] recently proposed mobility aware hierarchical clustering approach (HIER) to group Virtual Base Stations (VBSs) in two steps. The first step is location aware VBS clustering based on the location of Radio Resource Head (RRH). The second step is traffic aware placement of the cluster defined in previous step. Simulation results show that HIER can achieve up to 34.8% better QoS when using 5.8% additional RRHs. An Extended Cell (EC) concept is proposed in [85] to dynamically form groups of several adjacent cells. HO performance improvement is rendered by increasing the overlapping area between two adjacent cells in RoF indoor networks. Simulation results show that number of HOs are reduced and the call drop probability during HO is decreased by 70%. Although proven effective, the scheme lacks dynamic pro-

cedures to define ECs to optimize network resources. This shortcoming was addressed by authors in [86] by extending the idea and coming up with a proposal on Moving Extended Cell (MEC). Here, each mobile UE is covered by 7-cell EC where each transmits the same user data at every instance which in turn reduces HO latency through early preparation. Evaluation results show the proposed architecture can totally avoid call drop and packet loss for UEs with velocity of up to 40 m/s. Authors added that MEC is very efficient in tackling HO for mmWave cells but is vulnerable to throughput inefficiency as all seven cells in the cluster are transmitting for a single user. Virtual Cell (VC) is a concept discussed by Hossain et al. in [87] as one possible solution to increase the throughput efficiency of 60 GHz radio frequency (RF) network. VC is a central part of an actual cell, where the boundary is divided into multiple tiles. A wireless sensor network keeps track of UE location and periodically sends report to a centralized controller. Multiple Antenna Terminals (AT) cover a single cell, and only a single antenna terminal is activated at an instant. When the UE steps on the tile, controller activates respective neighbor AT to transmit similar data. Maximum of only two ATs can be activated for HO preparation in contrast to 6 in MEC. End results of using VC concept show an increase of 33% throughput efficiency in comparison to MEC. Authors in [72] uses optimization theory to achieve max data rate and min blocking probability in a heterogeneous environment. They suggest user centric approach is more suitable in hybrid 5G environment than network centric. By dividing users into HO and non-HO users their scheme ensures that a minimum throughput of 2 Mbps is achieved.

3.3 Handover Signalling Reduction

There exists a large body of research work on other areas of HetNets such as energy efficiency [88, 89, 90, 91] etc. However, very little work has been done on reducing mobility signalling in Handovers for HetNets. The work presented in [91] focuses on optimization of HO procedure in HetNets by incorporating context information such as user speed, channel gains and traffic load in the cells. This work proposes a Markov chain-based framework to model the HO process for the mobile user and derives an optimal context-dependent HO criterion. This work clearly demonstrates that context-awareness can indeed improve the HO process and significantly increase the performance of mobile UEs in HetNets.

In contrast to [91], this dissertation aims to address the question of how much signalling is generated in case of HOs for CDSA based HetNets. One recent study in [92] does provide HO analysis for CDSA based HetNets. This work provides the first tractable mobility aware model for a two-tier downlink cellular network with ultra-dense small cells and Control plane / User plane split architecture. The work performs in depth HO analysis and sheds light on HO costs in terms of number of HOs taking place per unit length. However, widely differing from the scope of this dissertation, [92] does not compare quantitatively how much amount of signalling load is generated in case of HO success and HO failure in CDSA for HetNets. The work in [93] focuses on HO problem in two tier networks which arises in HetNets due to network densification. The solution to the problem is specified in terms of HO skipping based on velocity of the user, so that connection can be maintained for longer duration without causing any connection interruption. HO cost is defined based on the delay incurred on account of HO interruption which takes place

during a HO. Though [93] considers the two tier HetNet model, unlike our study it does not consider a CDSA specifically. In addition, it does not take into account mobility signalling load considerations.

According to recent report by Nokia Siemens Networks, in current network deployments signalling is growing 50 percent faster than data traffic [94]. Previously published works [29, 37, 38, 39] on mobility signalling claims that mobility signalling is reduced as long as UE's mobility is within the coverage area of CBS. As a result, signalling channel is not changed and mobility signalling is reduced. However, this is not the case when the UE moves from one CBS to another CBS. Other studies [40, 41, 42] analyze the dual connectivity and HO failure rate of the CDSA using simulations, without providing a concrete analytical framework. In order to evaluate the HO signalling cost, [95] and [96] propose HO management schemes and evaluate the signalling cost for femtocells. However, these analyses assume that the HO is successful for 100 percent of the time, which is not the case in real networks. In order to assess the mobility and signalling reduction benefits of CDSA, a framework is direly needed that quantifies the mobility related signalling load in realistic settings. A first attempt towards this framework is reported in [43]. While this attempt provides the first few building blocks of the required framework, in contrast to the complete framework developed in this dissertation, it misses out two important facts: 1. It does not consider HO failures scenario. 2. It does not take into account quality of service (QoS) requirements such as HO time, shared coverage factor and HO duration related parameters which are essential for QoS requirements of time sensitive applications in emerging ultra-dense HetNets.

3.4 Summary

In this chapter we provided a review of the state of the art research activities on handovers in cellular networks. The review of pertinent literature shows that a lot of research is being done in various areas of mobility management and handovers. However, only a limited number of studies investigate mobility Signalling. In the few studies that investigate mobility Signalling, the focus remains on idle mode, whereas this dissertation investigates account the active mode Signalling that constitutes the bulk of HO Signalling. We identify one very recent study that happens to be most related to the scope of this dissertation as it does address mobility signalling in CDSA in connected mode. The framework presented in this dissertation advances the state of the art substantially beyond this study by making two key contributions. 1. Unlike prior most relevant work that unrealistically assumes the handover to be successful 100% of times, our framework considers realistic scenario where handover can fail or be successful. This is a significant, because HO failure causes more signalling than HO success. 2. In contrast to prior relevant works, the framework developed in this dissertation takes into account quality of service requirements such as handover time, intercell shared coverage overlap factor and handover duration related parameters which are essential for quality of service requirements of time sensitive applications in emerging ultra-dense HetNets.

CHAPTER 4

Handover Signalling Probability Model

In order to evaluate the CN signalling load as a result of HO, we need to find out the probabilities for handover success (P_s), handover failure (P_f) and in case when no handover takes place (P_{no}). In order to compute these values, we need to model the handover procedure between two base stations in terms of a timing diagram shown in Fig. 4.1.

The symbols in the timing diagram are shown in Table 4.1. In order to compute the signalling probabilities, we need to define a system model.

4.1 System Model

Consider a CDSA cellular network where the CBSs are modelled as a Poisson Point Process (PPP) with density ρ_1 , while DBSs are modelled as another PPP with density ρ_2 , where $\rho_2 \geq \rho_1$.

Assume a session duration λ with probability density function (pdf) $f_S(\lambda)$ and mean E[λ]. The CBS residence time is modeled as a random variable θ_1 with pdf $f_{R_1}(\theta_1)$ and mean E[θ_1], while the DBS cell residence time is modelled as a random variable θ_2 with pdf $f_{R2}(\theta_2)$ and mean E[θ_2].

Fig. 4.1 provides a timing diagram that illustrates the definition of all the parameters. Without loss of generalization, we follow [43] and assume that users move in random directions with a random velocity. Under this assumption, E[θ_1] can be approximated by the ratio between the number of UEs in a CBS and the number of UEs leaving a CBS per unit time [97]. According to [97]

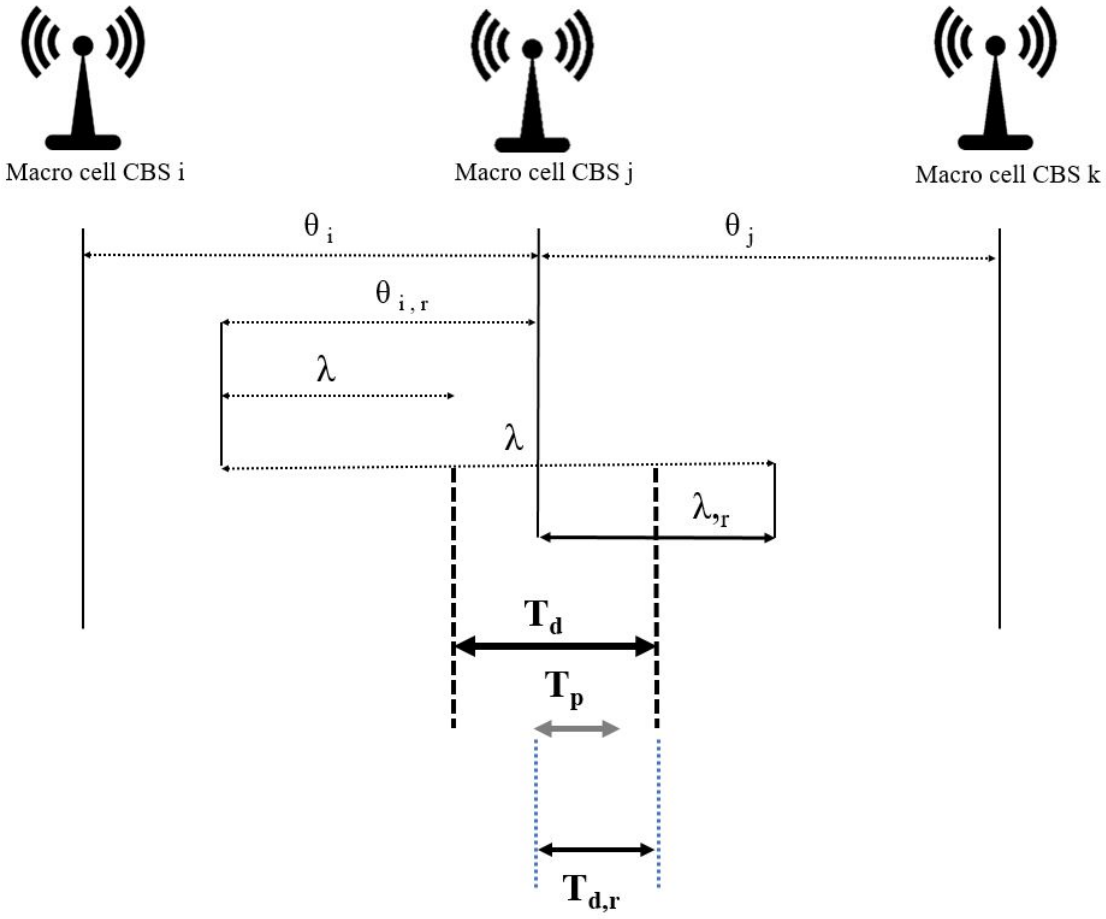

Fig. 4.1: Timing diagram of handover model parameters

Table 4.1: Symbol Description

Symbol	**Description**
S_f	Normalized CN mobility signalling load on account of HO failure
S_s	Normalized CN mobility signalling load on account of HO success
λ	Session Duration
λ_r	Residual Session Duration
θ	Cell residence time
θ_1	Cell residence time of CBS
θ_2	Cell residence time of DBS
θ_r	Residual Cell residence time
$\theta_{i,r}$	Residual Cell residence time of control base station i
θ_2	Residual Cell residence time of data base station
T_d	Mobility time duration during which HO takes place
$T_{d,r}$	Residual Mobility time duration during which HO takes place
T_p	Time taken for handover completion
ρ_1	Cell density of control base station
ρ_2	Cell density of data base station
v	Average velocity
L_1	Length of perimeter of the control base station
S_1	Area of control base station
c	Coverage factor

$$\mathrm{E}[\theta_1] = \frac{\text{Number of UEs in a CBS}}{\text{Number of UEs leaving a CBS}} \tag{4.1}$$

Following derivations in [97], $\mathrm{E}[\theta_1]$ can be approximated as:

$$E[\theta_1] \approx \frac{\pi * S_1}{E[v]L_1} \tag{4.2}$$

where, the symbols S_1 in equation (4.2) indicates area of the CBS and L_1 represents length of the perimeter of CBS as given in Table 4.1. As we are considering a PPP model, according to [98]

$$S_1 = \frac{1}{\rho_1} \tag{4.3}$$

and

$$L_1 = \frac{4}{\sqrt{\rho}_1} \tag{4.4}$$

Substituting (4.3) and (4.4) into (4.2). Equation (4.2) can be re-written as

$$E[\theta_1] \approx \frac{\pi}{4E[v]\sqrt{\rho}_1} \tag{4.5}$$

Similarly, the expected cell residence time for DBS can be listed as

$$E[\theta_2] \approx \frac{\pi}{4E[v]\sqrt{\rho}_2} \tag{4.6}$$

4.2 Mobility Time Duration (T_d)

If a user is in a moving state, due to mobility, it may move from one CBS to another CBS. For a successful HO, the HO procedure must be completed within a certain time duration. Otherwise, UE may move out of coverage of the serving CBS and HO failure occurs. As the user approaches the edge of CBS, it starts receiving signal coverage from a neighboring CBS. Ideally, the HO ought to take place when a user is receiving signal from both neighboring

and serving CBS, while it is moving in the direction of the neighboring CBS. This duration while UE is receiving signal from serving CBS and moving in the direction of neighboring CBS, receiving signal above a fixed threshold from neighboring CBS is termed as *mobility time duration* and abbreviated as T_d. In other words T_d is the time taken to traverse the intercell overlap coverage area.

In order to define T_d mathematically, we proceed as follows: a UE stays in a cell for a given time equal to average cell residence time (θ). T_d is a function of average cell residence time. Cell residence time is dependent upon cell density and user velocity. Therefore, in order to derive a relation between mean cell residence time and mobility time duration (HO duration), we model it as:

$$T_d = E[\theta] * c * 10^{-3} \tag{4.7}$$

Where T_d is the HO time duration (HO time) in the above equation (4.7). HO time depends upon the intercell coverage overlap area characterized by c. For larger intercell coverage overlap area, HO duration is longer, and for small shared coverage area it is shorter. Therefore, the coverage parameter c ranges between 0.1 and 0.9. The shared coverage factor c is dimensionless in our model. Mobility time duration is taken in milliseconds, whereas cell residence time is in seconds (depending upon the cell radius), therefore a factor of 1/1000 is used for conversion.

4.3 Time Taken For a Handover Completion (T_p)

The HO procedure starts from the instant measurement report is sent by the UE to the source CBS, and concludes once UE receives RRC connection reconfiguration message from the target CBS. The HO procedure consists of

three phases: preparation, execution and completion phase as shown in Fig. 2.9. For a successful HO all three phases need to be completed successfully. The time taken to complete all the phases of HO successfully is termed as T_p. In [99], authors have studied the HO failure rate and delay of the overall delay of HO as well. Their results include overall HO duration which is around 83-95 ms depending upon the UE speed and physical layer error rate. HOs can take place sooner than this duration as well. In order to compute the effect of signalling load in case of both HO success and failure scenarios we use the value of T_p as 100 ms, an upper bound to meet the HO delay requirements in this dissertation.

4.4 Probability of Handover Failure

For the CDSA system model, it is known that CN signalling is generated in inter CBS HOs only [30]. Expressed differently, all the DBS HOs do not generate CN signalling as long as the CBS anchor point remains the same. The definition of P_f is equivalent to the UE attempting to change the serving CBS, while doing so it is not successful. With reference to Fig. 4.1, P_f is equivalent to the probability that T_p occurs beyond residual mobility time duration $T_{d,r}$. The session started when UE was associated with CBS_i and failed to finish and the session is dropped when the UE tries to attempt a HO in order to associate with CBS_j, where j > i and $T_p > T_{d,r}$. Considering Fig. 4.1, we can write P_f as:

$$P_f = Prob.(T_p > T_{d,r}) * Prob.(\lambda > \theta_{1,r}) \tag{4.8}$$

where Prob.( ) means probability of an event and $\theta_{1,r}$ is the residual cell residence time of a CBS. The probability that session duration (λ) is greater

than the residual cell residence time, is computed as:

$$Prob.(\lambda > \theta_{1,r}) = 1 - Prob.(\lambda < \theta_{1,r})$$

When session duration is less than residual cell residence time, it is computed as:

$$Prob.(\lambda < \theta_{1,r}) = \int_{x=0}^{\infty} f_{\theta_{1,r}}(x) \int_{y=0}^{x} f_{\lambda}(y)\ dy\ dx \tag{4.9}$$

$$Prob.(\lambda > \theta_{1,r}) = 1 - \int_{x=0}^{\infty} f_{\theta_{1,r}}(x) \int_{y=0}^{x} f_{\lambda}(y)\ dy\ dx \tag{4.10}$$

Similarly, $T_{d,r}$ is the residual mobility time duration during which HO takes place as shown in Fig. 4.1. The probability that time taken for an inter-CBS HO completion (T_p) is greater than residual mobility time duration is computed as:

$$Prob.(T_p > T_{d,r}) = 1 - Prob.(T_p < T_{d,r})$$

When time taken for HO completion is less than residual mobility time duration, it is computed as:

$$Prob.(T_p < T_{d,r}) = \int_{z=0}^{\infty} f_{T_{d,r}}(z) \int_{v=0}^{z} f_{T_p}(v)\ dv\ dz \tag{4.11}$$

$$Prob.(T_p > T_{d,r}) = 1 - \int_{z=0}^{\infty} f_{T_{d,r}}(z) \int_{v=0}^{z} f_{T_p}(v)\ dv\ dz \tag{4.12}$$

Plugging the values from equations (4.12) and (4.10) in equation (4.8), we get probability of failure as:

$$P_f = \Big(1 - \int_{z=0}^{\infty} f_{T_{d,r}}(z) \int_{v=0}^{z} f_{T_p}(v)\ dv\ dz\Big) * \Big(1 - \int_{x=0}^{\infty} f_{\theta_{1,r}}(x) \int_{y=0}^{x} f_{\lambda}(y)\ dy\ dx\Big) \tag{4.13}$$

4.5 Probability of Handover Success

The definition of P_s is equivalent to the UE attempting to change the serving CBS, and it is successful in doing so. With reference to Fig. 4.1, P_s is equivalent to the probability that T_p instant occurs before $T_{d,r}$ duration. In other words, the session started when UE was associated with CBS_i and finished successfully in the next CBS when the UE tried to attempt a HO in order to associate with CBS_j , where j > i and $T_p < T_{d,r}$. Considering Fig. 4.1, we can write P_s as:

$$P_s = Prob.(T_p < T_{d,r}) * Prob.(\lambda > \theta_1, r) \tag{4.14}$$

Plugging values from equations (4.10) and (4.11) into equation (4.14), we get:

$$P_s = \Big(\int_{z=0}^{\infty} f_{T_{d,r}}(z) \int_{v=0}^{z} f_{T_p}(v)\ dv\ dz \Big) * \Big(1 - \int_{x=0}^{\infty} f_{\theta_{1,r}}(x) \int_{y=0}^{x} f_{\lambda}(y)\ dy\ dx \Big) \tag{4.15}$$

4.6 Probability of No Handover

The probability of no HO (P_{no})is the probability that the UE does not attempt to change the serving CBS. With reference to Fig. 4.1, P_{no} is equivalent to the probability that the session started when UE was associated with CBS_i and finished successfully in the same CBS and UE did not try to attempt a HO in order to associate with CBS_j , where j > i. Considering Fig. 4.1, we can write P_{no} as:

$$P_{no} = Prob.(\lambda < \theta_1, r) \tag{4.16}$$

Plugging value from equation (4.9) into equation (4.16), it becomes:

$$P_{no} = \int_{x=0}^{\infty} f_{\theta_{1,r}}(x) \int_{y=0}^{x} f_{\lambda}(y) dy dx \tag{4.17}$$

The probabilities of HO failure, HO success and no HO considering general distributions are shown in equations (4.13), (4.15) and (4.17) respectively. In order to have closed form expression for these probabilities we consider exponential distribution as follows.

4.7 Exponential Distribution for Session Duration, Mobility time duration and Cell Residence Time

The expressions for P_f, P_{no} and P_s computed earlier in equations (4.13), (4.15) and (4.17) are given for general distribution. In order to have a closed form solution, we consider the scenario where the session duration, cell residence time and mobility time duration are exponentially distributed. Exponential distribution has been considered in this dissertation as it represents the worst-case scenario from signalling load perspective. The model(s) in [43] show that the HO-related signalling load is memoryless under exponential distribution and the signalling probability is independent of the previous case. Consequently, we consider the exponential distribution to model the worst-case scenario in both the CDSA and the conventional architecture. This is to evaluate the upper bound of the signalling load that corresponds to insights into the worst case scenario.

According to [43] when session duration and the cell residence time are exponentially distributed, the residual session duration and the residual cell residence time are also exponentially distributed such that:

$$f_{\lambda}(t) = f_{\lambda,r}(t) = \frac{e^{-\frac{t}{E[\lambda]}}}{E[\lambda]} \tag{4.18}$$

$$f_{\theta_1}(t) = f_{\theta_{1,r}}(t) = \frac{e^{-\frac{t}{E[\theta_1]}}}{E[\theta_1]} \tag{4.19}$$

The mobility time duration is derived from cell residence time. Therefore, if cell residence time is assume exponential, probability density function (pdf) of mobility time duration is given as:

$$f_{T_d}(t) = \frac{e^{-\frac{t}{E[T_d]}}}{E[T_d]} \tag{4.20}$$

Similarly, using Lemma 1 in [43] the pdf of residual mobility time duration in case of exponential distribution is given as:

$$f_{T_d}(t) = f_{T_{d,r}}(t) = \frac{e^{-\frac{t}{E[T_d]}}}{E[T_d]} \tag{4.21}$$

Substituting (4.18), (4.19) and (4.20) into equations (4.13), (4.15) and (4.17) respectively. After mathematical simplification, we get P_f, P_s and P_{no} closed form expressions to be as follows:

$$P_f = \left(\frac{E[T_p]}{E[T_{d,r}] + E[T_p]}\right) * \left(\frac{4E[\lambda]E[v]\sqrt{\rho}_1}{\pi + 4E[\lambda]E[v]\sqrt{\rho}_1}\right) \tag{4.22}$$

$$P_s = \left(\frac{E[T_{d,r}]}{E[T_{d,r}] + E[T_p]}\right) * \left(\frac{4E[\lambda]E[v]\sqrt{\rho}_1}{\pi + 4E[\lambda]E[v]\sqrt{\rho}_1}\right) \tag{4.23}$$

$$P_{no} = \frac{\pi}{\pi + 4E[\lambda]E[v]\sqrt{\rho}_1} \tag{4.24}$$

The closed form expressions for P_f and P_s indicate that they depend upon cell density, user velocity, session duration, mobility time duration and time taken for a HO completion. Mobility time duration in turn depends upon cell residence time and intercell shared coverage factor. Therefore, from a design perspective a larger value of intercell shared coverage factor and high cell

residence time result in higher successful HO probability. This insight can help cellular network designers to plan better ultra-dense networks, which can result in more successful HOs and less mobility signalling compared to conventional networks.

4.8 Numerical Results

Probability of HO Signalling and Coverage factor

This subsection evaluates the probability of signalling in case of Handover success, failure and no Handover versus velocity for different values of coverage factor. The evaluation is based on exponential distribution for session duration, cell residence time and mobility time duration. The evaluation is also based on normalized densities w.r.t the CBS. The value of c influences overall mobility signalling. As in Fig. 4.2 , for low value of c, probability of failure signalling is high and decreases with an increase in shared coverage factor. For low values of shared coverage c, the HO boundary shrinks resulting in smaller values of T_d. On account of smaller values of mobility duration T_d, probability of failure signalling increases. Whereas for high value of c, the HO boundary region is suitable for HO success. Therefore the probability of HO success signalling increases with shared coverage factor c as shown in Fig. 4.3 . Also its worthy to note, for reasonable coverage factor values, at very low speeds (less than 20 Km/hr), probability of success (P_s) increases because gradual increase in speed results in correct handover trigger timing thereby not causing a too late HO. However with increase in mobility at higher speeds, probability of failure (P_f) increases as a result of increase in number of too late handovers, while probability of success starts to decrease, as evident in Fig. 4.3 and Fig. 4.2 respectively.

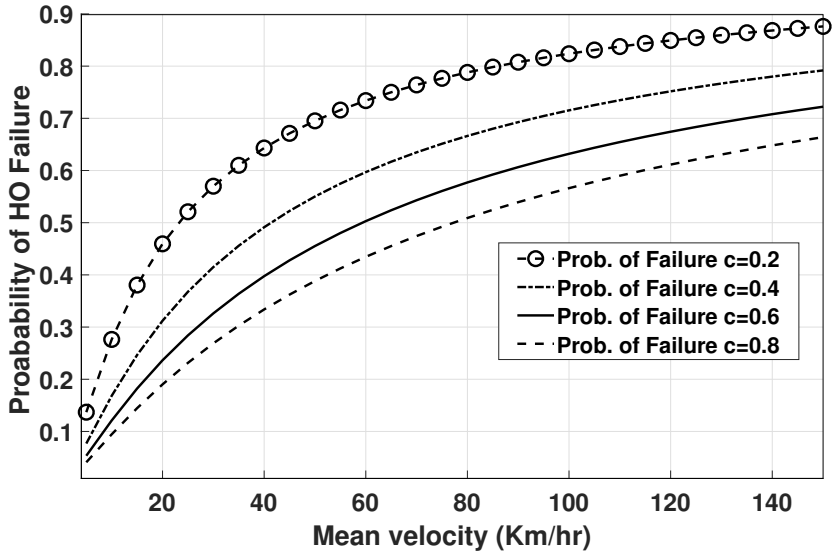

Fig. 4.2: Probability of HO failure vs mean velocity for different values of coverage factor. E[λ] = 5 mins and E[ρ] = 10

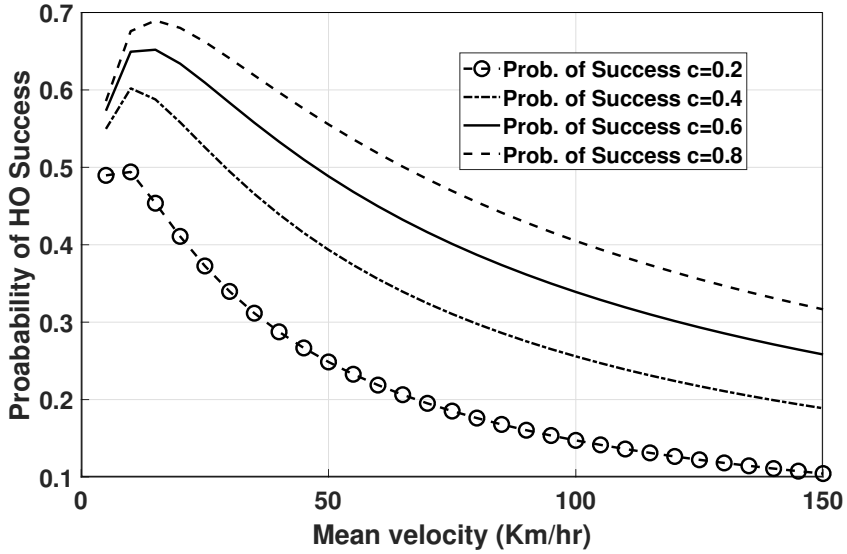

Fig. 4.3: Probability of HO success vs mean velocity for different values of coverage factor. E[λ] = 5 mins and E[ρ] = 10

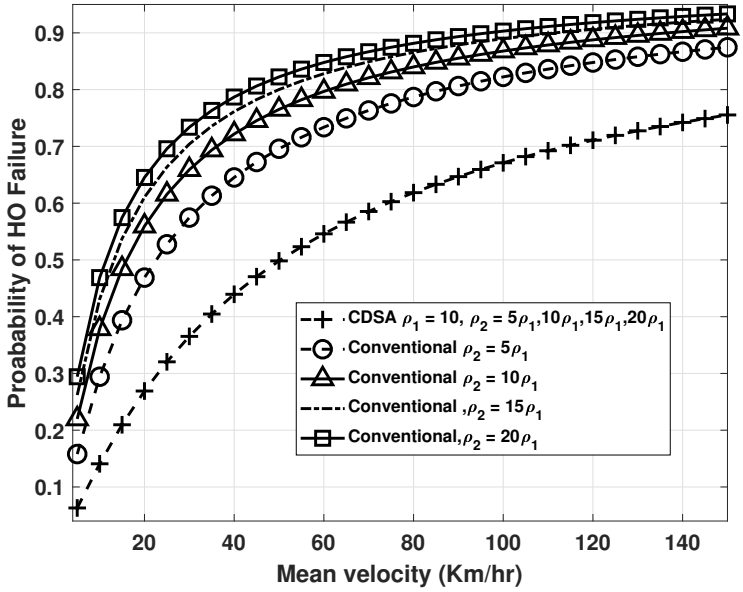

Fig. 4.4: Probability of HO failure vs mean velocity for different values of cell density. E[λ] = 5 mins and c = 0.5

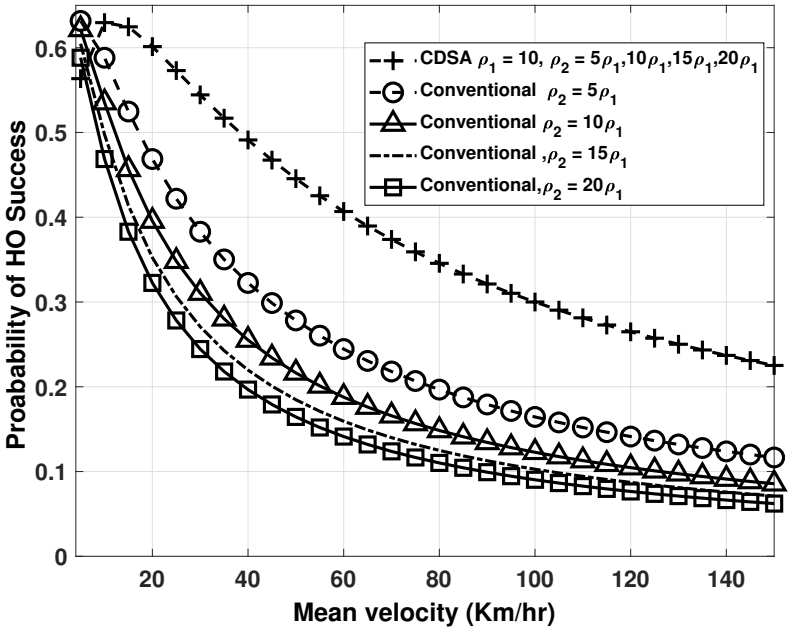

Fig. 4.5: Probability of HO Success vs mean velocity for different values of cell density, session duration E[λ] = 5 mins and intercell coverage overlap c = 0.5

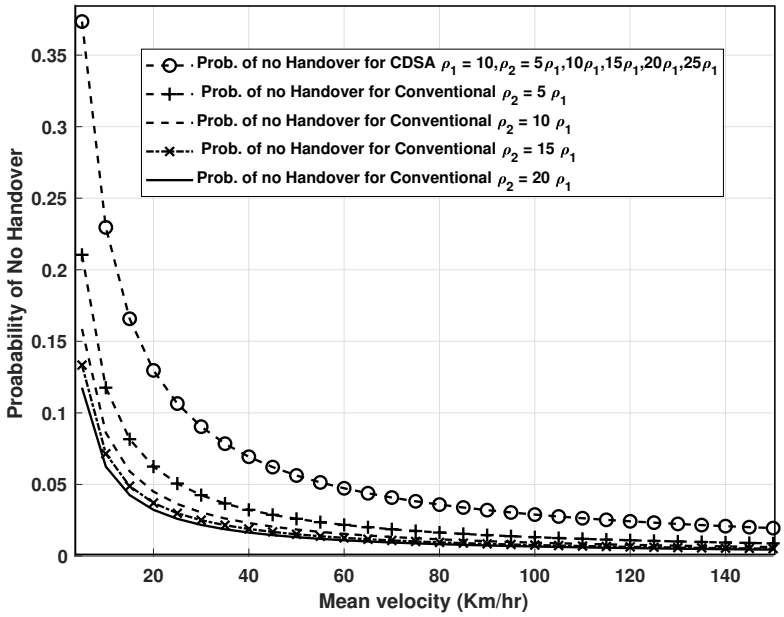

Fig. 4.6: Probability of no handover signalling vs mean velocity for CDSA and different cell densities of conventional network with session duration E[λ] = 5 mins

The probability of no HO signalling is the same for all coverage factor values and does not depend on the value of c but changes with velocity and cell density. In order to observe (P_{no}) for different cell densities, the value of P_{no} is shown in Fig. 4.6. Probability of no HO signalling has highest value for CDSA and decreases with increase in cell density for conventional architecture. Probability of failure

text

is lowest for CDSA versus conventional architecture while probability of success starts low for CDSA but with an increase in mean velocity it becomes higher than conventional architecture, as shown in Fig. 4.4 and Fig. 4.5 respectively for c = 0.5.

4.9 Quantification of Handover Failure Signaling

A typical HO procedure consists of three phases: preparation, execution and completion phases [100]. In this study for the current system model, the user gets connected back to the system through RRC connection re-establishment in case of HO failure. It must be kept in mind, HO failure can take place in any of the preparation, execution and completion phase(s). When a HO failure takes place, the UE has to go through connection re-establishment procedure once again in order to get connected with a DBS. Numerical computation of HO signalling considering each message and processing at different nodes is computed in [95], [96], [101] and [102]. In order to approximate, how much additional signalling is generated in case of HO failure, consider that if HO failure takes place during HO completion phase, then RRC re-establishment will take place to keep the user in the system after HO failure. This procedure results in more signalling messages compared to HO success alone as shown in Fig. 2.9. HO failure signalling is normalized with S_f and HO success signalling is normalized with S_s. The total HO signalling (failure and success signalling) normalized by S is given as follows:

$$S = S_f + S_s \tag{4.25}$$

Looking at Fig. 2.9, in chapter 2 we can write S_f in terms of S_s

$$S_f = S_s + 0.25 * S_s$$
$$S_f = 1.25 * Ss$$

With reference to Fig. 2.9, therefore total signalling is:

$$S = 1.25 * S_s + S_s$$

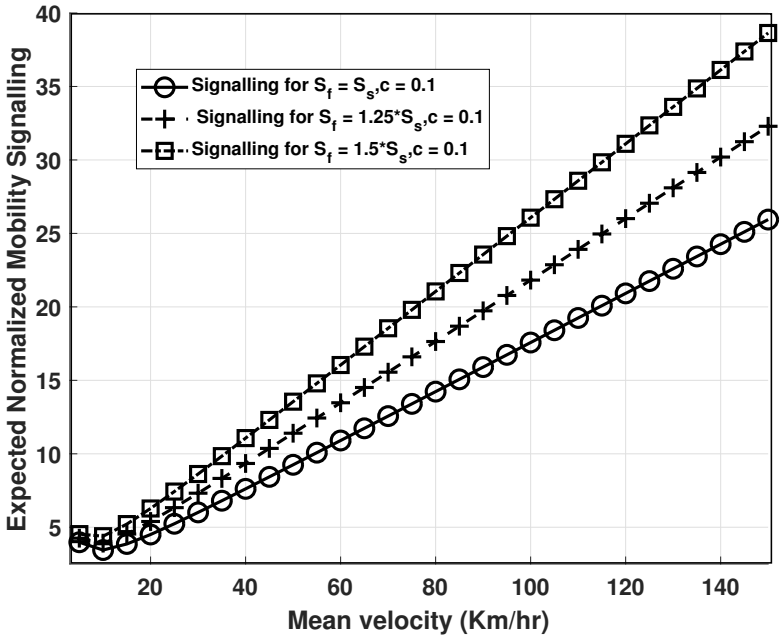

Fig. 4.7: Comparison of handover failure and success for c = 0.1, cell density E[ρ] = 10 and session duration E[λ]= 5 mins, in terms of expected mobility signalling normalized with S_f and S_s indicating how handover failure results in more mobility signalling

$$S = 2.25 * S_s \tag{4.26}$$

$$\tag{4.27}$$

This indicates that HO failure signalling load has different quantitative evaluation than HO success signalling load. Differences in HO failure signalling load depend upon the scenario(s) considered. Fig. 4.7 provides information about an increase in expected signalling load for different S_f values for a intercell shared coverage factor $c = 0.1$. It shows that for the given scenario considered, HO failure signalling load results in almost 1.6 times more normalized expected signalling load.

4.10 Summary

In this chapter we derived a framework that can allow to compute the probabilities of handover. In order to keep the analysis applicable to a range of scenarios, specific handover probabilities were computed, i.e., probability of handover success (P_s) , probability of handover failure (P_f) and probability of no handover (P_{no}). The computed probabilities depend upon cell density, user velocity, handover time duration, session duration, cell residence time and inter-cell coverage overlap.

The presented analytical framework thus offers the first key step towards the analytical evaluation of realistic mobility signalling as a function of cell density, speed, session duration, handover time duration and shared coverage overlap. Building on the results in this chapter, in next chapter we derive the framework for quantifying the signalling load associated with the HO success, failure and no HO scenarios.

CHAPTER 5

Mobility Signalling Model

The total CN mobility signalling load generated during a HO depends upon a number of factors such as:

- UE speed and mobility type i.e., continuous or continual mobility
- BS density
- Session duration
- Transport network reliability (stability)
- Intercell Coverage overlap
- Miscellaneous

The user(s) is assumed to be RRC connected and active in the network. We model the HO scenario and CN signalling generated as a result of probability of HO failure, HO success and no HO signalling (dervied in Chapter 4) using Markov chain as shown in Fig. 5.1. In the CDSA approach shown in Fig.1.1, each inter-CDSA HO success or HO failure generates CN signalling. We will denote this CN signalling as $S_{i,j}$, while intra-CBS HOs i.e., DBSs do not generate CN signalling. The coefficients α, β and γ in Fig. 5.1 are HO coefficients. In a cellular network, the probability to HO from one cell to another is not the same for all the sectors. This difference is on account of various factors described above. These HO coefficient values represent the difference in probabilities for HO from one cell to another cell. Building on the signalling quantification framework presented in [43] we can model the

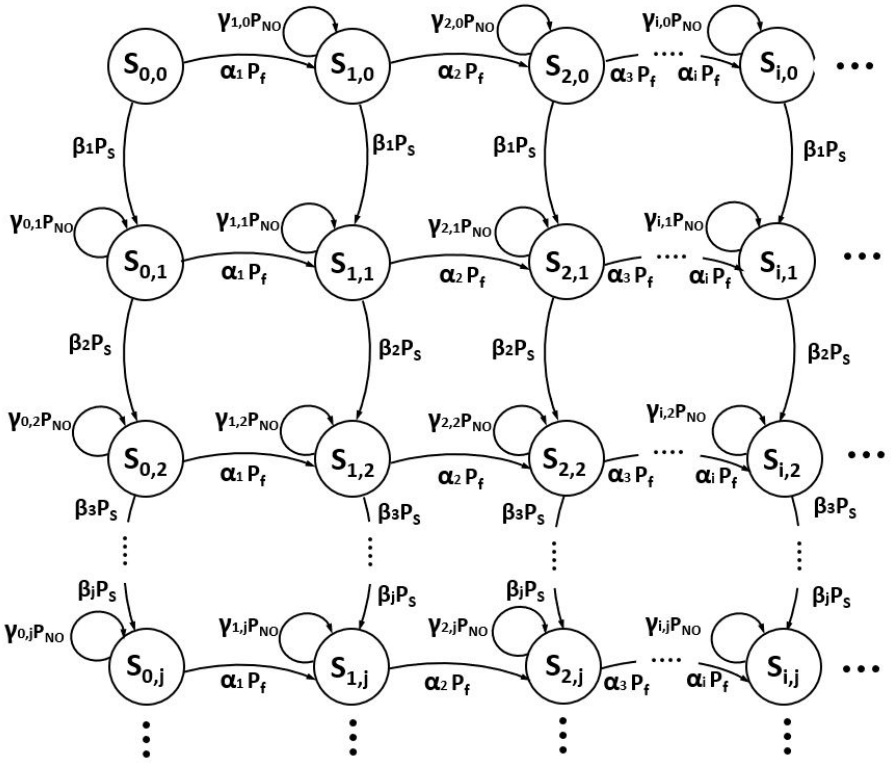

Fig. 5.1: Markov chain modeling of no-handover, handover failure and handover success related core-network mobility signaling

CN signalling on amount of HO success, failure and no HO as shown in Fig. 5.1 as follows:

- P_f = Probability that signalling is generated as a result of HO failure
- P_s = Probability that signalling is generated as a result of HO success
- P_{no} = Probability that HO attempt will not be made
- $S_{i,j}$ = CN mobility signalling load generated on account of $\boldsymbol{i}$ handover failure(s) and $\boldsymbol{j}$ handover success(s)

5.1 Computation of Expected Mobility Signalling

The goal is to find out the average or expected amount of mobility CN signalling which is generated in case of HO success(s) and HO failure(s), including

how no HO attempts will influence the aggregate signalling. From Fig. 5.1 it is clear that the user will always generate mobility signalling starting from state $S_{0,0}$ and will not stay in that state.

The expected value of RRC CN mobility signalling load E[$S_{i,j}$] generated by a UE in the CDSA can be calculated as:

$$E[S_{i,j}\,] = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} S_{i,j} * Prob.(S_{i,j}\,) \tag{5.1}$$

The Prob. ($S_{i,j}$) can be calculated by solving the Markov chain shown in the Fig. 5.1. Since the amount of signalling generated by the user(s) movement increases with time, a transition from CN signalling state $CS_{i,j}$ to $CS_{m,n}$ has zero probability when *i,j* > *m,n*. Based on this Markov chain, Prob.($S_{i,j}$) can be formulated as:

$$Prob.(S_{i,j}) = \begin{cases} \frac{\alpha_i P_f^i}{(1-P_{no})^i} * P(S_{0,0}\,) & , for\ i > 0, j = 0 \\ \frac{\beta_j P_s^j}{(1-P_{no})^j} * P(S_{0,0}\,) & , for\ i = 0, j > 0 \\ \frac{(i+j)\alpha_i P_f^i \beta_j P_s^j}{(1-P_{no})^{i+j}} * P(S_{0,0}\,) & , for\ i > 0, j > 0 \\ P(S_{0,0}\,) & , for\ i = 0, j = 0 \end{cases} \tag{5.2}$$

From the Markov chain in Fig. 5.1 the values of HO coefficients α, β and γ are such that the following conditions are true:

$$\beta_1 P_s + \alpha_1 P_f = 1, for\ i = 0, j = 0$$
$$\beta_1 P_s + \alpha_{i+1} P_f + \gamma_{i,0}\, P_{no} = 1, for\ i > 0, j = 0$$
$$\beta_{j+1} P_s + \alpha_1 P_f + \gamma_{0,j}\, P_{no} = 1, for\ i = 0, j > 0$$
$$\beta_{j+1} P_s + \alpha_{i+1} P_f + \gamma_{i,j}\, P_{no} = 1, for\ i > 0, j > 0$$

$$\alpha_1 \geq \alpha_2 \geq \alpha_3 \geq \alpha_4 \geq ... \ \geq \alpha_i$$
$$\beta_1 \geq \beta_2 \geq \beta_3 \geq \beta_4 \geq ... \ \geq \beta_j$$
$$\gamma_{1,0} \leq \gamma_{2,0} \leq \gamma_{3,0} \leq \gamma_{4,0} \leq ... \ \leq \gamma_{i,0}$$

$$\gamma_{0,1} \leq \gamma_{0,2} \leq \gamma_{0,3} \leq \gamma_{0,4} \leq \ldots \leq \gamma_{0,j}$$
$$\gamma_{1,1} \leq \gamma_{2,1} \leq \gamma_{3,1} \leq \gamma_{4,1} \leq \ldots \leq \gamma_{i,1}$$
$$\gamma_{1,2} \leq \gamma_{2,2} \leq \gamma_{3,2} \leq \gamma_{4,2} \leq \ldots \leq \gamma_{i,2}$$

Similarly,

$$\gamma_{1,j} \leq \gamma_{2,j} \leq \gamma_{3,j} \leq \gamma_{4,j} \leq \ldots \leq \gamma_{i,j}$$

5.2 Lemma

For exponential distribution of cell residence time and session duration, the values of α, β and γ are :

$$\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 =, ..., \alpha_i = 1$$
$$\beta_1 = \beta_2 = \beta_3 = \beta_4 =, ..., \beta_j = 1$$
$$\gamma_{1,0} = \gamma_{2,0} = \gamma_{3,0} = \gamma_{4,0} =, ..., \gamma_{i,0} = 1$$
$$\gamma_{0,1} = \gamma_{0,2} = \gamma_{0,3} = \gamma_{0,4} =, ..., \gamma_{0,j} = 1$$
$$\gamma_{1,1} = \gamma_{2,1} = \gamma_{3,1} = \gamma_{4,1} =, ..., \gamma_{i,j} = 1$$

Proof

Preliminary: Given the session duration and the CBS residence time are exponentially distributed, the residual session duration and the residual CBS residence time will also be exponentially distributed [43]. Consequently, the probability of not generating signalling is memoryless and independent of the state, i.e., independent of the state whether signalling has been generated previously or not.

This implies that $P_{no} = \gamma_{1,0}P_{no} = \gamma_{2,0}P_{no} \ldots = \gamma_{i,0}P_{no}$ resulting in $\gamma_{1,0} = \gamma_{2,0} \ldots = \gamma_{i,0} = 1$. Similarly, for other $\gamma_{0,j} = \gamma_{i,j} = 1$. Since at any given state, $\alpha_i P_f + \beta_j P_s + \gamma_{i,j} P_{no} = 1$. When $\gamma_{i,j} = 1$ for all states, then the term

$\alpha_i P_f + \beta_j P_s$ remains the same in all the states. As the residual session duration and the residual cell residence time have exactly the same distribution as the session duration and the cell residence time, respectively, $\alpha_i P_f$ and $\beta_j P_s$ remain constant in all states. This condition can only be satisfied when $\alpha_1 = \alpha_2 = ... = \alpha_i = 1$ and $\beta_1 = \beta_2 = ... = \beta_j = 1$.

5.3 Mobility Signalling Mathematical Model Derivation

As the probabilities of the signalling states in Markov chain for Fig. 5.1 are shown in equation (5.2), these probabilities depend on state $S_{0,0}$ i.e., $P(S_{0,0})$. Once we compute the probability of this state, we can compute probabilities for other states as well. For a Markov chain we know that:

$$\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} Prob.(S_{i,j}) = 1$$

Using Fig. 5.1 we can sum up all the signalling states such that:

$$Prob.(S_{0,0}) + \sum_{i=1}^{\infty} Prob.(S_{i,0}) + \sum_{j=1}^{\infty} Prob.(S_{0,j}) + \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} Prob.(S_{i,j}) = 1$$

After simplifying the equation above, we can write the Prob.($S_{0,0}$) = P($S_{0,0}$) as:

$$P(S_{0,0}) = \frac{1}{1 + \sum_{i=1}^{\infty} Prob.(S_{i,0}) + \sum_{j=1}^{\infty} Prob.(S_{0,j}) + \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} Prob.(S_{i,j})} \quad (5.3)$$

After mathematical manipulation and solving (5.3). We get

$$P(S_{0,0}) = \frac{P_s P_f}{P_s P_f + P_f^2 + P_s^2 + P_f(1 - P_{no}) + P_s(1 - P_{no})} \quad (5.4)$$

$$E[S_{i,j}] = S_{0,0} * Prob.(S_{0,0}) + \sum_{i=1}^{\infty} S_{i,0} * Prob.(S_{i,0}) + \sum_{j=1}^{\infty} S_{0,j} * Prob.(S_{0,j}) + \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} S_{i,j} * Prob.(S_{i,j}) \quad (5.5)$$

where ,

$$S_{0,0} = [0 * S_f] + [0 * S_s]$$

$$S_{i,0} = [i * S_f] + [0 * S_s]$$

$$S_{0,j} = [0 * S_f] + [j * S_s]$$

$$S_{i,j} = [i * S_f] + [j * S_s]$$

After plugging the values from equations (5.2) and (5.4) in equation (5.5), mathematical simplification results in expected signalling. The expected signalling as a result of HO failures and HO success can be computed as:

$$E[S_{i,j}] = (\frac{P_f(1-P_{no})}{P_s^2} + \frac{(P_f - P_{no} + 1)(1-P_{no})}{P_s^2} + \frac{(1-P_{no})^2}{P_s P_f})S_f P(S_{0,0}) + (\frac{P_s(1-P_{no})}{P_f^2} + \frac{(P_s - P_{no} + 1)(1-P_{no})}{P_f^2} + \frac{(1-P_{no})^2}{P_s P_f})S_s P(S_{0,0}) \quad (5.6)$$

Equation (5.6) can be used to quantify the RRC CN mobility signalling load for a mobile user. The expected signalling load can be computed by substituting the values of P_f, P_s and P_{no} derived in Chapter 4.

5.4 Mobility Signalling for Conventional Networks

In order to assess the mobility signalling load generated in conventional networks, the modeling approach used earlier for CDSA can be adapted to model the conventional network mobility signalling as well. By doing so, we can com-

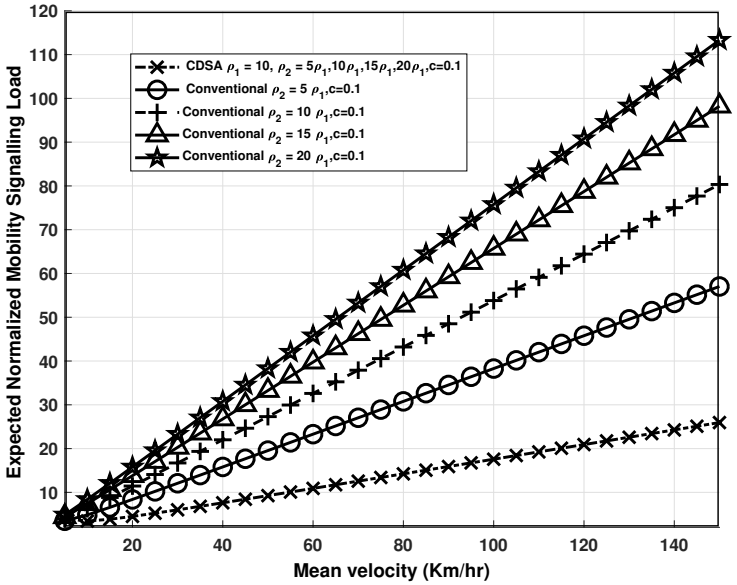

Fig. 5.2: Normalized expected signalling load vs. mean velocity for intercell shared coverage factor $c = 0.1$ and session duration $\mathrm{E}[\lambda] = 5$ mins

pare the advantage of CDSA over conventional cellular network deployment. The evaluation and comparison approach for CDSA vs. conventional network is proposed in [43] which can be adapted for this scenario as well.

5.5 Numerical Results

Signalling in CDSA versus Conventional networks

In this subsection we evaluate how much expected CN mobility signalling is generated in case of CDSA versus conventional networks, as proposed in the analytical model derived earlier in this chapter. We consider exponential distribution for session duration, cell residence and mobility duration time. The evaluation is based on normalized densities w.r.t CBS. In addition, the RRC signalling load (in terms of expected value) is normalized with S (more specifically S_f for HO failures and S_s for HO success). Fig. 5.2 shows the normalized expected mobility signalling load vs. mean velocity for intercell

coverage overlap factor $c = 0.1$, while Fig. 5.3 provides this information for intercell coverage overlap factor c value of 0.6. With low values of intercell coverage overlap factor c there is a high probability of HO failure and an increase in CN mobility signalling load, even at slow speeds. Fig. 5.2 indicates CDSA generates $\{31, 54, 72, 87\}$*S times less signalling load compared to the conventional network with different cell densities. Expected mobility signalling load reduces in the case of a high intercell coverage overlap factor c. With an increase in medium and high mobility speeds, probability of failures increase, so expected mobility signalling is supposed to increase along with increase in mobility. Fig. 5.3 shows that CDSA results in $\{5, 9, 12, 14\}$*S times less signalling load vs. conventional networks even at high velocity and intercell coverage overlap factor c respectively. CDSA is a clear winner for generating less mobility signalling load. These plots suggests that CDSA performs equally better at greater mobility and high speed scenarios. This proves our initial hypothesis that in case of ultra dense networks, CDSA deployment is beneficial, whereas conventional networks results in excessive mobility signalling load.

5.6 Summary

In this chapter, we utilized the probabilities of handover already derived in Chapter 4 to evaluate mobility signalling in RRC connected mode (active mode). A two-dimensional Markov chain was used to represent the overall RRC mobility signalling in the core network. The state transitions from one Markov state to another Markov state was decided based on probability of handover failure and probability of handover success respectively. Similarly, remaining in the same Markov state corresponds to probability of no handover.

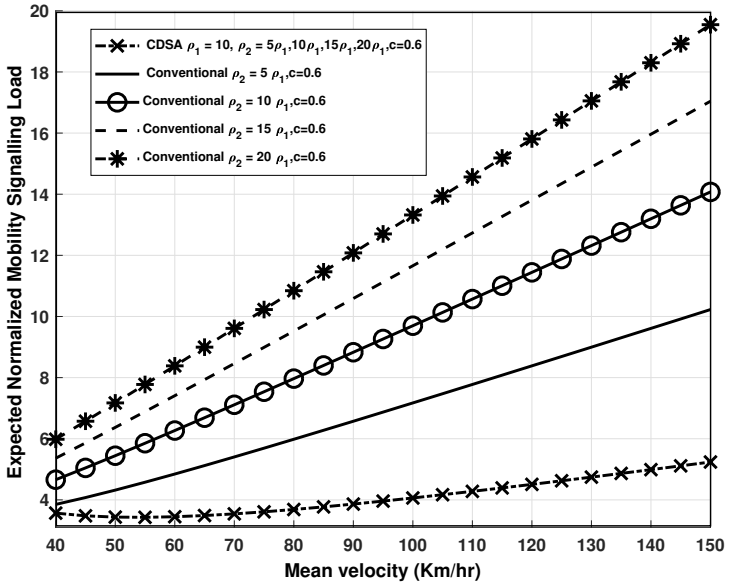

Fig. 5.3: Normalized expected signalling load vs mean velocity for intercell shared coverage overlap $c = 0.6$ and session duration $\mathrm{E}[\lambda] = 5$ mins at high speeds

Mathematical computations and algebraic simplifications were used to solve the Markov chain so that expected mobility signalling can be computed.

To compute the expected mobility signalling in CDSA with a fair benchmark, the analytical framework was adapted to quantify the mobility Signalling in conventional HetNets. The numerical results show that as expected, in ultra-dense network the mobility signalling reaches to prohibitive level as the cell density and user speed increases. Comparative results show that, CDSA results in far less core network Signalling compared to HetNets in same deployment and mobility scenarios. This analysis quantifies the gain of CDSA in terms of mobility signalling there by determining its need and viability for real deployment.

CHAPTER 6

Continuous Mobility Signalling Model

During a mobility HO scenario, one of the two cases can happen. Either the HO is successful or the HO is not successful. The user remains in the system even in case of HO failure. For the case of continuous mobility, a user generates mobility signalling as a result of HO success and HO failures while the session duration is continuous. It is assumed that the user remains RRC connected with the CBS in the system even in case of HO failure and gets connected back to DBS through RRC connection re-establishment. As the session is continuous, the probability of no HO signalling is zero in this case. The probability of failure and probability of success are complementary of each other in this case. The Markov chain for this special scenario is shown in Fig. 8.1. Looking at the 2D Markov chain in Fig. 8.1, it can be inferred:

- P_f = Probability that CN mobility signalling will be generated as a result of HO failure
- P_h = Probability that CN mobility signalling will be generated as a result of HO success in case of continuous mobility
- $CS_{i,j}$= Aggregate CN mobility signalling load on account of i HO failures and j HO successes in case of continuous mobility

The goal is to find out the average or expected amount of CN signalling which is generated in case of continuous mobility as a result of continuous HO success and failures respectively. This probability $\mathrm{P}(\mathrm{CS}_{i,j})$ can be calculated by solving the Markov chain shown in Fig. 8.1. Since the amount of signalling generated by the users' movement (HO failures and success) increases with

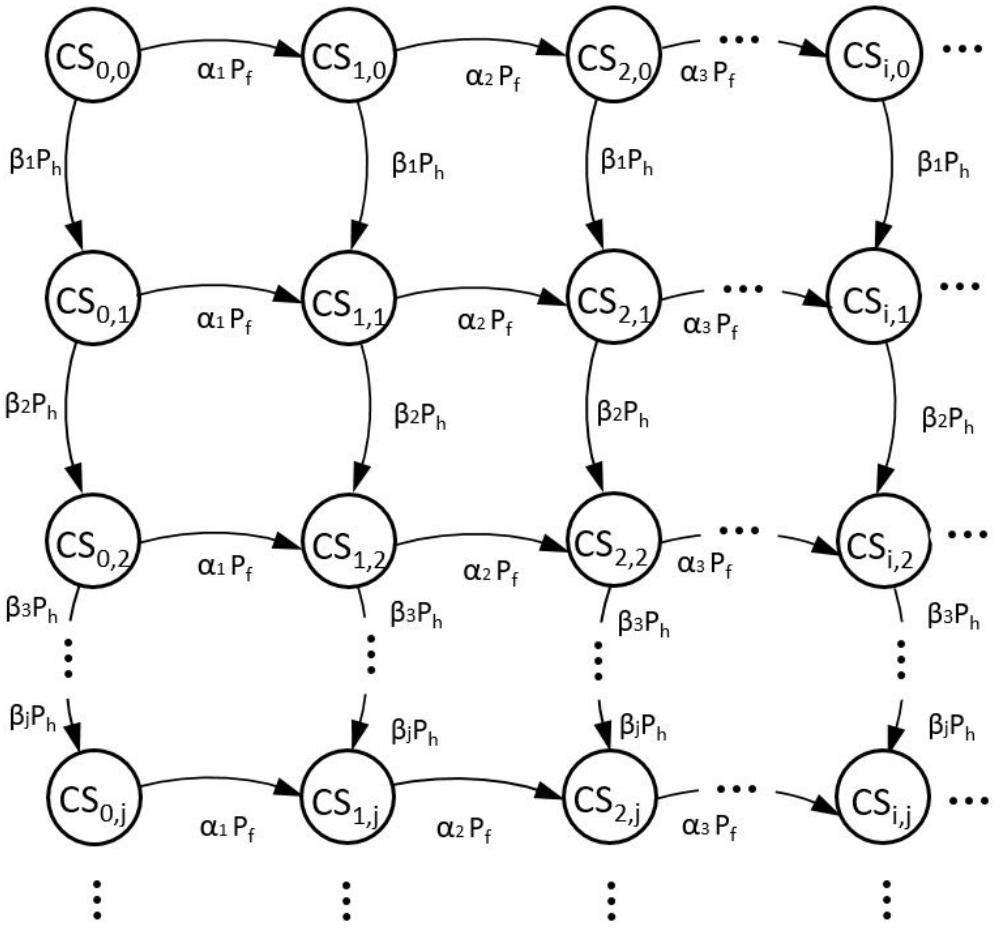

Fig. 6.1: Markov chain representing continuous handover success and failure signalling scenarios

time, a transition from CN signalling state $CS_{m,n}$ to $CS_{i,j}$ has zero probability when $m,\ n > i,\ j$.

Based on this model, Prob ($CS_{i,j}$) can be formulated as:

$$Prob.(CS_{i,j}) = \begin{cases} P(CS_{0,0}) & i = 0, j = 0 \\ \alpha_i P_f^i P(CS_{0,0}) & i > 0, j = 0 \\ \beta_j P_h^j P(CS_{0,0}) & i = 0, j > 0 \\ (i+j)\alpha_i P_f^i \beta_j P_h^j P(CS_{0,0}),\ i > 0,\ j > 0 \end{cases} \tag{6.1}$$

In Lemma 1 of Chapter 5 it was shown that for exponential cell residence and session duration the values of α and β are:

$$\beta_1 = \beta_2 = \beta_3 = \beta_4 =, ..., = \beta_j = 1$$
$$\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 =, ..., = \alpha_i = 1$$

6.1 Derivation of Continuous Mobility Probability of Success (P_h)

The probability of failure in the case of continuous mobility is the same as computed in the Chapter 4 equation (4.21) for a non-continuous scenario.

$$P_f = \left(\frac{E[T_p]}{E[T_{d,r}] + E[T_p]}\right) * \left(\frac{4E[\lambda]E[v]\sqrt{\rho}_1}{\pi + 4E[\lambda]E[v]\sqrt{\rho}_1}\right)$$

Considering a continuous mobility scenario, the probability of success is the complement of probability of failure. If HO failure does not take place then it will be probability of success. The probability of HO success is given as:

P_h = 1 - P_f

The probability of HO success is computed as:

$$P_h = 1 - \left[\left(\frac{E[T_p]}{E[T_{d,r}] + E[T_p]}\right) * \left(\frac{4E[\lambda]E[v]\sqrt{\rho}_1}{\pi + 4E[\lambda]E[v]\sqrt{\rho}_1}\right)\right] \tag{6.2}$$

6.2 Handover Signalling for Continuously Mobile Users

In the case of the continuous mobility scenario, a large number of HOs take place. When we consider, there are a lot of HOs successes and failures happening consistently and users have a high mobility, it requires us to derive another expression for CN signalling generated as a result of continuous HOs scenario. The number of HO failures is denoted by $\boldsymbol{i}$ and number of HO successes is denoted by $\boldsymbol{j}$. The expected CN signalling is given as:

$$E[CS_{i,j}\,] = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} CS_{i,j} * Prob.(CS_{i,j}\,) \tag{6.3}$$

where ,

$$CS_{0,0} = [0 * S_f] + [0 * S_s]$$

$$CS_{i,0} = [i * S_f] + [0 * S_s]$$

$$CS_{0,j} = [0 * S_f] + [j * S_s]$$

$$CS_{i,j} = [i * S_f] + [j * S_s]$$

To compute the expected signalling in case of continuous mobility, we need to find out the probability of state $CS_{0,0}$ i.e., $P(CS_{0,0})$. Looking at Fig. 8.1 and we know that for a Markov chain:

$$\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} Prob.(CS_{i,j}\,) = 1$$

Expanding the expression using Fig. 8.1

$$Prob.(CS_{0,0}\,) + \sum_{i=1}^{\infty} Prob.(CS_{i,0}\,) + \sum_{j=1}^{\infty} Prob.(CS_{0,j}\,) + \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} Prob.(CS_{i,j}\,) = 1$$

Resolving the mathematical expression to compute the value of $P(CS_{0,0})$

$$P(CS_{0,0}) = \frac{1}{1 + \sum_{i=1}^{\infty} Prob.(CS_{i,0}) + \sum_{j=1}^{\infty} Prob.(CS_{0,j}) + \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} Prob.(CS_{i,j})}$$

Simplifying the mathematical procedures of equation above, the probability of state $CS_{0,0}$ is given as:

$$P(CS_{0,0}) = \frac{P_h P_f}{P_h P_f + P_f^2 + P_h^2 + P_f + P_h} \tag{6.4}$$

Now in order to compute the expected mobility signalling, in case of continuous mobility we plug values from equations (6.1) and (6.4) into equation (6.3) and solve :

$$E[CS_{i,j}] = CS_{0,0} * Prob.(CS_{0,0}) + \sum_{i=1}^{\infty} CS_{i,0} * Prob.(CS_{i,0}) + \sum_{j=1}^{\infty} CS_{0,j} * Prob.(CS_{0,j}) + \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} CS_{i,j} * Prob.(CS_{i,j})$$

After solving the CN mobility signalling for continuous mobility users in case of HO successes and failures turns out to be:

$$\begin{aligned} E[CS_{i,j}] = &\left(\frac{P_f}{P_h^2} + \frac{(P_f+1)}{P_h^2} + \frac{1}{P_h P_f}\right) S_f P(CS_{0,0}) \\ &+ \left(\frac{P_h}{P_f^2} + \frac{(P_h+1)}{P_f^2} + \frac{1}{P_h P_f}\right) S_s P(CS_{0,0}) \end{aligned} \tag{6.5}$$

The CN signalling for high mobility users depends upon P_f and P_h. After substituting the values of P_f, P_h and $P(CS_{0,0})$ from equations (4.21), (6.2) and (6.4) respectively, mobility signalling for continuous mobility users can be evaluated.

Comparison of Normal and Continuous Mobility Signalling

After comparing the two analytical equations (5.6) and (6.5) respectively (normal (continual) and continuous mobility), it is evident that in case of the

continuous mobility scenario the probability of no HO is zero ($P_{no} = 0$). If we substitute the value of $P_{no} = 0$ in the normal mobility expected signalling equation (5.6), the two equations apparently become equal. Even though the two equations look equal for $P_{no} = 0$, it is not true mathematically as the values of P_f, P_h and P_s are different, including the values of $S_{i,j}$ and $CS_{i,j}$.

6.3 Numerical Results

Continuous Mobility versus Non-Continuous Mobility

In this subsection we compare the expected non-continuous CN mobility signalling versus continuous CN mobility signalling as derived analytically in Chapters 5 and in this Chapter 6. Fig. 6.2 indicates non-continuous and continuous mobility signalling for $c = 0.1$ and $c = 0.5$ respectively. In both scenarios, continuous mobility signalling is much higher compared to non-continuous mobility at low speeds. For non-continuous mobility, as velocity increases, probability of no HO signalling approaches zero. From the numerical comparison of expected normal and continuous mobility signalling in the Fig. 6.2, it is evident that continuous mobility signalling provides the upper bound for expected signalling generated as a result of HO. P_{no} is zero at all times for the continuous mobility signalling scenario. For a normal scenario, at low speeds P_{no} is not equal to zero. However, with increase in velocity, P_{no} starts approaching zero. This is evident from expected signalling generated at high velocities, and is the same both in the case of normal and continuous mobility scenarios. This confirms that in order to compute upper limit for mobility signalling in any case, the continuous mobility model can be used.

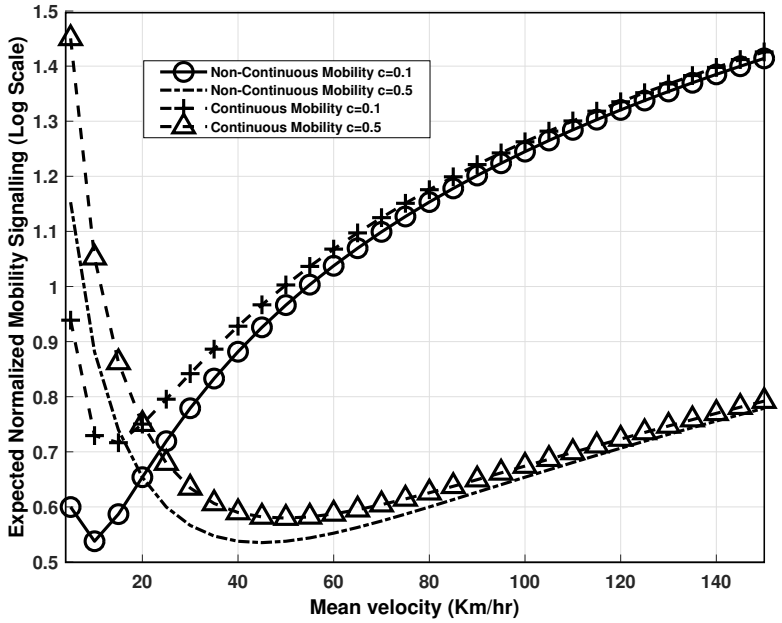

Fig. 6.2: Expected mobility signalling vs mean velocity in case of normal and continuous mobility for Intercell shared Coverage Overlap values $c = 0.1$ and $c = 0.5$, in terms of normalized with S_f and S_s

6.4 Summary

In this chapter, we relaxed the stationary condition by stating that user is in a continuous mobility state i.e., $P_{no} = 0$. Therefore, user will perform handovers continuously as a user moves from one base station to another. Either a handover will be successful or a handover failure will occur. In order to find out the expected mobility signalling, a two dimensional Markov chain is used with probabilities of handover failure and handover success respectively.

Mathematical analysis and algebraic simplifications were used to solve the Markov chain so that expected mobility signalling could be derived in case of the continuous mobility scenario.

The significance of the numerical results computed through the derived expression in this chapter is three fold:

1. The results confirm the intuition that mobility signalling for users with

continual or intermittent mobility is far less than that for users that are continuously moving.

2. The results provide an upper bound on amount of mobility related signalling.

3. The derived framework can be used to model and analyze HO signalling generated by cars on highways, airplanes or trains and similar commute systems.

CHAPTER 7

Intercell Coverage Overlap and Mobility Signalling

The current cell planning paradigm simply leaves out mobility considerations for post deployment optimization. At best, mobility related issues are considered during planning phase in the form of ensuring some overlap between adjacent cells, just enough to allow sufficient time for HO to take place, but not too much to cause excessive interference. However, no systematic method exists to determine the optimal area of overlap between cells, particularly from the perspective of minimizing HO signalling overheads. The cell overlap design determines HO signalling overheads as it dictates the ratio of successful and failed HOs. Despite its significance in determining quality of user experience (QoE) in legacy networks, the cell overlap is generally designed heuristically. This approach has worked in legacy networks but is not adequate for emerging ultra-dense networks for two reasons.

1) In emerging networks, drastically shrinking cell size means HO will become much more frequent, and mobility signalling and its impact on QoS cannot be neglected, even in planning phase.

2) With increased number of cells, post-deployment optimization of a myriad of overlapping regions can become extremely expensive task in terms of time, complexity and labor. Therefore, it will make sense to taken into account mobility signalling, even during planning phase. To the best of our knowledge, no existing study addresses this question i.e., how cell coverage overlap can be optimized for minimizing mobility signalling.

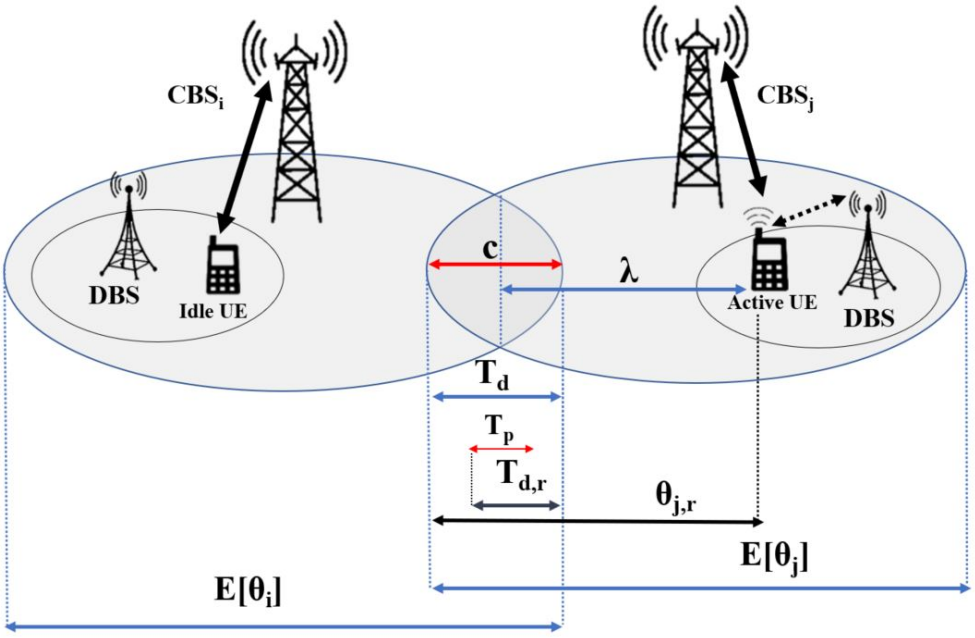

Fig. 7.1: Schematic of a typical CDSA, coverage overlap boundary and HO Modeling

7.1 System Model for Shared Coverage Area and Mobility Signalling

This analysis focuses on radio resource control-related core network mobility signalling exchange, which takes place during HO procedure in a cellular system. The system model similar to continuous mobility users scenario shown earlier is used. The scenario for shared coverage area is shown in Fig. 7.1

7.2 Probability of HO Success

The probability of HO success in this case is similar to what is computed previously in the case of the continuous mobility scenario. The probability of HO (P_h) success in this case is given as:

$$P_h = 1 - [\frac{E[T_p]}{E[T_{d,r}] + E[T_p]} * \frac{4E[\lambda]E[v]\sqrt{\rho_1}}{\pi + 4E[\lambda]E[v]\sqrt{\rho_1}}] \tag{7.1}$$

Where the definition of T_p, T_d, $T_{d,r}$ and λ are given in Fig.7.1

Probability of HO Failure

The probability of HO failure in case of continuous mobility users is similar to the one we computed previously in case of continuous mobility and is given as:

$$P_f = \frac{E[T_p]}{E[T_{d,r}] + E[T_p} * \frac{4E[\lambda]E[v]\sqrt{\rho_1}}{\pi + 4E[\lambda]E[v]\sqrt{\rho_1}} \tag{7.2}$$

Expected Mobility Signalling for Continuous Mobility Users

The expected normalized mobility signalling for continuous mobility users is derived in Chapter 6 and is given as follows:

$$E[CS_{i,j}] = (\frac{P_f}{P_h^2} + \frac{(P_f+1)}{P_h^2} + \frac{1}{P_h P_f})S_f P(CS_{0,0}) + (\frac{P_h}{P_f^2} + \frac{(P_h+1)}{P_f^2} + \frac{1}{P_h P_f})S_s P(CS_{0,0})$$

7.3 Coverage Overlap and Handover Related Mobility Signalling

Earlier we computed the probability of HO success, HO failure and normalized expected mobility signalling. Now we discuss the opportunity of how shared coverage area is correlated with normalized mobility signalling. In order to make the effect of shared coverage area applicable to a range of deployment scenarios, we proceed as follows by analyzing the previously derived signalling equations. The expected normalized signalling load is a function of HO probabilities as follows:

$$\textbf{Signalling load} = \boldsymbol{f}\ (P_f \textbf{ and } P_h)$$

The probability of HO success and HO failure is a function of the following factors:

$$P_f \textbf{ and } P_h = \boldsymbol{f} \textbf{ (}T_p\textbf{, }T_d\textbf{, cell density, velocity)}$$

The mobility time duration in return is a function of:

$$T_d = \boldsymbol{f} \textbf{ (Cell residence time, cell overlap fraction)}$$

The above analysis indicates that shared coverage area indirectly affects the mobility signalling load. In order to show the effect of shared coverage area on mobility signalling, we analyze the mobility signalling for different values of shared coverage factor in the following numerical results section.

7.4 Numerical Results

7.4.1 Expected Mobility Signalling and Coverage Factor

This subsection evaluates the affect of coverage factor c on expected CN mobility signalling. Fig. 7.2 shows the expected normalized CN mobility signalling vs. mean velocity for different values of coverage factor. For a given velocity and cell density, intercell coverage factor c determines which configuration would result in least amount of mobility signalling. Fig. 7.2 indicates the coverage factor of 0.1 results in least amount of signalling at lowest speed for average cell density E[ρ] = 10. However, the same coverage factor value of 0.1 does not result in the same least amount of mobility signalling for average cell density E[ρ]=200 as shown in Fig. 7.3. The general trend is that small values of c result in low mobility signalling at low speeds vs. large values of

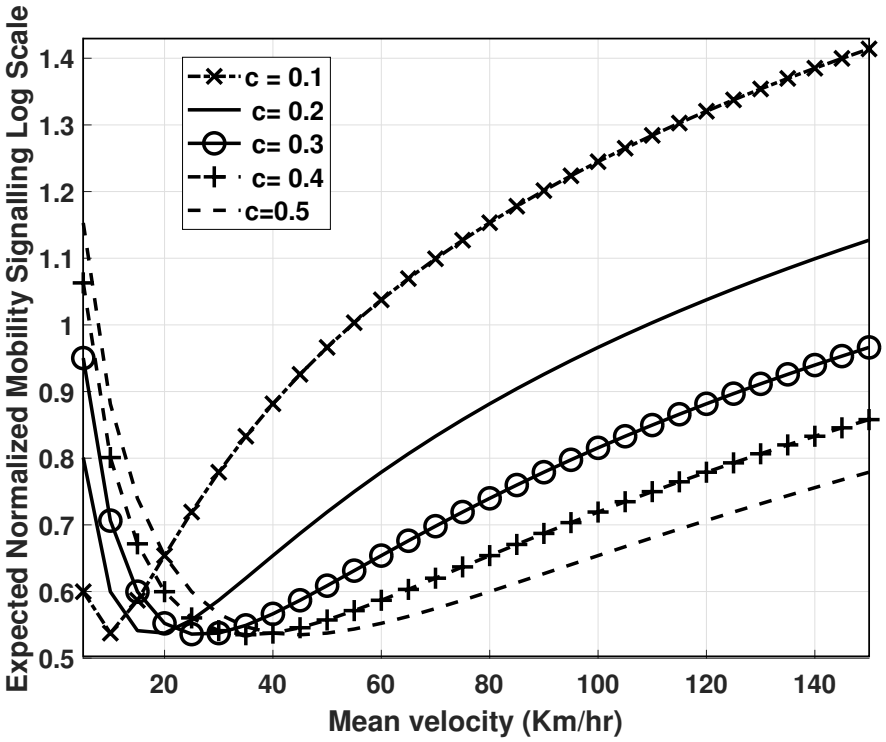

Fig. 7.2: Expected normalized mobility signalling vs. mean velocity for cell density E[ρ] = 10 and session duration E[λ]= 5 mins. Expected mobility signalling is normalized with S_f and S_s

c at higher speed. This tells us, in order to find out the appropriate value of coverage factor for cellular network planning which can result in optimum amount of signalling, analytical model(s) from this dissertation can be used. The analytical model described can be used to evaluate signalling for a given scenario of allowed number of HO failures, HO successes, velocity and cell density.

7.4.2 Expected Signalling vs. Mobility Time Duration

In order to observe the behavior of expected normalized mobility signalling vs. mobility time duration, results in Fig. 7.4 can be used. It indicates signalling for different cell densities for a velocity of 60 Km/hr. Analysis shows that signalling is reduced for large T_d values and it is increased for higher cell

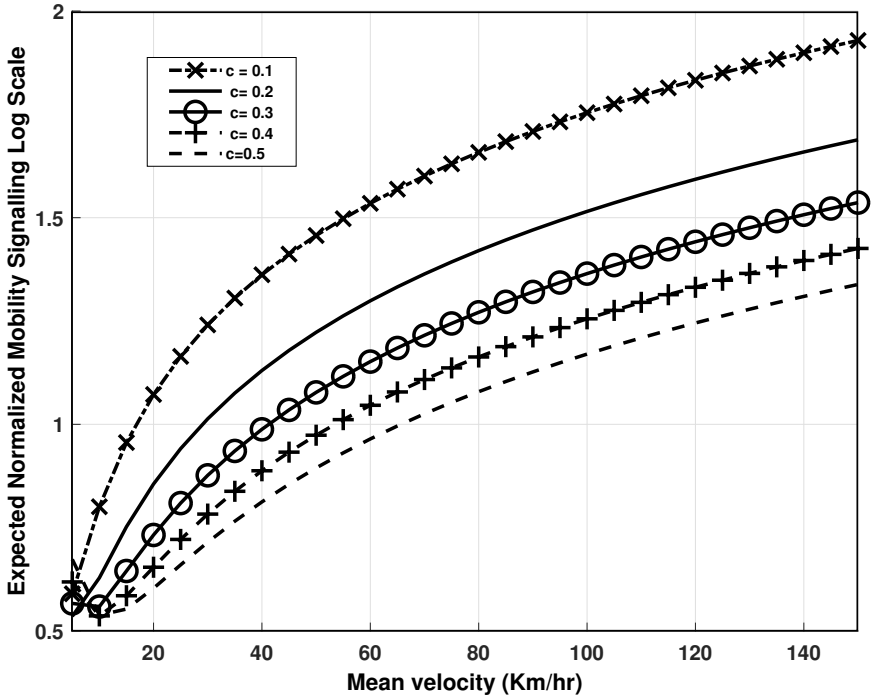

Fig. 7.3: Expected normalized mobility signalling vs. mean velocity for cell density E[ρ] = 200 and session duration E[λ]= 5 mins. Expected Mobility signalling is normalized with S_f and S_s

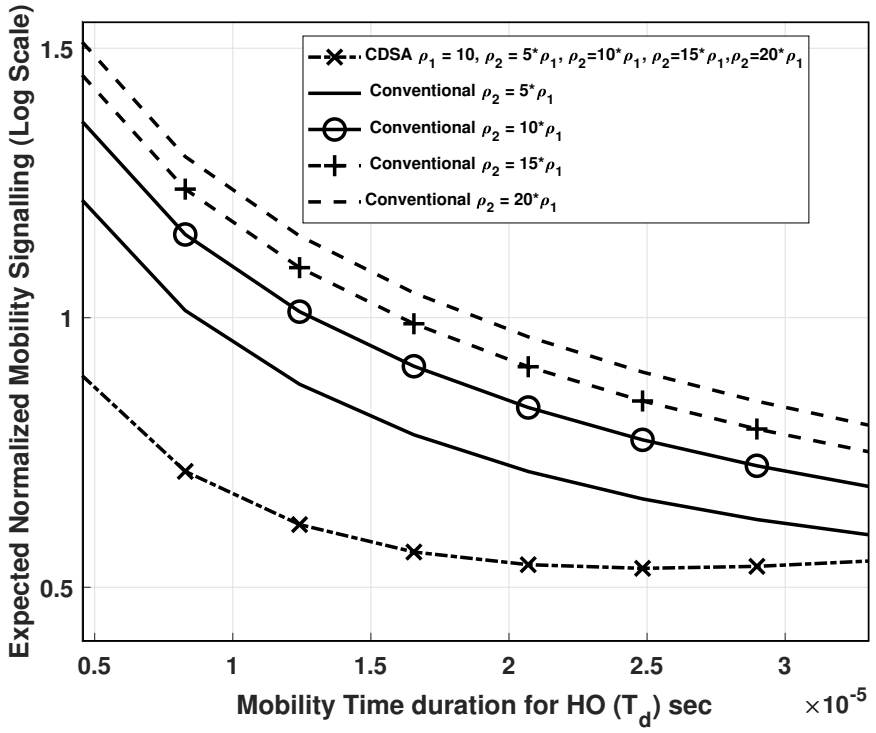

Fig. 7.4: Expected normalized mobility signalling vs mobility time duration T_d for HO, for velocity, $v = 60$ km/hr and session duration E[λ]= 5 mins. Expected mobility signalling is normalized with S_f and S_s

density. This result and the analysis done in this dissertation can be used to determine the optimal value of intercell coverage overlap fraction c for a given HO parameters, mobility statistics and cell density to minimize the signalling load or vice versa.

7.5 Summary

In this chapter, we extended the analytical model derived in previous chapters to investigate the question how much of mobility signalling reduction can be achieved by choosing a proper cell overlap fraction for a given velocity and cell density. A quantitative method to plan cell overlap for minimizing mobility signalling does not yet exist in commercial cellular network planning. The presented framework quantifies mobility signalling overhead as a function of

cell overlap fraction for a given velocity and cell density and thus offers the first crucial step towards this goal.

CHAPTER 8

Finite Handovers and Mobility Signalling

8.1 Analytical Model for Finite Handovers Signalling

In a realistic scenario for mobility handover, one of two cases can happen. Either the handover is successful or the handover is not successful. In case of a realistic mobility scenario, the number of handovers if finite. User(s) remains in the system even in case of handover failure. For the case of continuous mobility, user remains RRC connected while moving from one CBS to another. The user generates mobility signalling as a result of handover success and handover failures. It is assumed that the user remains RRC connected with the CBS in the system even in case of handover failure and gets connected back to DBS through RRC connection re-establishment. We assume the session is continuous, therefore the probability of no mobility signalling is zero in this case. Whereas, probability of failure and probability of success are complementary of each other. In previous chapters we analysed this scenario assuming infinte number of HOs. In this chapter the goal is to find out the expected amount of CN signalling which is generated as a result of a fixed given number of HO success and failures. This probability of expected CN signalling which is generated as a result of finite HOs is abbreviated as P($CS_{i,j}$). It can be solved by calculating Markov chain shown in Fig. 8.1. Since the amount of signalling generated by the users movement (HO failures and HO success) increase with time, a transition from CN signalling state $CS_{m,n}$ to $CS_{p,q}$ has zero probability $m,\ n > p,\ q$

Looking at the 2D Markov chain in Fig. 8.1. It can be inferred.

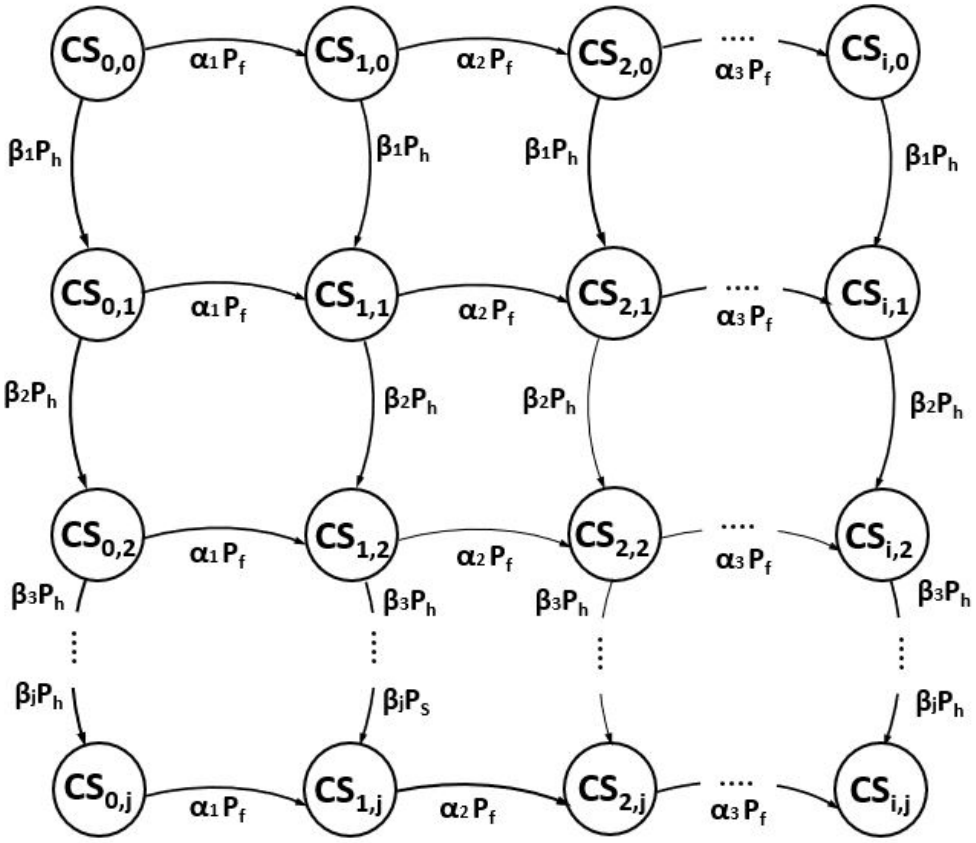

Fig. 8.1: Markov chain representing finite handover success and failure signalling scenarios

- P_f = Probability that CN signalling will be generated as a result of given handover failure in case of continuous mobility.

- P_h = Probability that CN signalling will be generated as a result of given handover success in case of continuous mobility

- $CS_{i,j}$ = Aggregate CN signaling load on account of i handover failures and j handover successes in case of continuous mobility

Based on this model, Prob. (CSp,q) can be formulated as:

$$Prob.(CS_{p,q}) = \begin{cases} \alpha_p P_f^p P(CS_{0,0}) & , for\ p > 0, q = 0 \\ \beta_q P_h^q P(CS_{0,0}) & , for\ p = 0, q > 0 \\ (p+q)\alpha_p P_f^p \beta_q P_h^q * P(CS_{0,0}) & , for\ p > 0, q > 0 \\ P(CS_{0,0}) & , for\ p = 0, q = 0 \end{cases} \tag{8.1}$$

Where α and β are HO coefficients in a cellular network and represent the difference in probabilities for HO from one cell to another cell. Using Lemma 1, we know for exponential cell residence time and session duration the values of α and β are equal to 1:

8.2 Expected Mobility Signalling Computation in Finite Handover Case

In order to compute an expression for CN signalling load generated as a result of given number of HOs scenario. The expected CN signalling is given as:

$$E[CS_{p,q}] = \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} CS_{p,q} * Prob.(CS_{p,q}) \tag{8.2}$$

where ,

$$CS_{0,0} = [0 * S_f] + [0 * S_s]$$

$$CS_{i,0} = [i * S_f] + [0 * S_s]$$

$$CS_{0,j} = [0 * S_f] + [j * S_s]$$

$$CS_{i,j} = [i * S_f] + [j * S_s]$$

Where S_f and S_s are normalized signalling load generated as a result of HO failure and success respectively. In order to compute the expected signalling in case of mobility, we need to compute the probability of state $CS_{0,0}$. Using Fig. 8.1 and we know that for a Markov chain:

$$\sum_{p=0}^{n}\sum_{q=0}^{m} Prob.(CS_{p,q}) = 1 \tag{8.3}$$

Expanding the equation above using Fig.8.1.

$$Prob.(CS_{0,0}) + \sum_{p=1}^{n} Prob.(CS_{p,0}) + \sum_{q=1}^{m} Prob.(CS_{0,q}) + \sum_{p=1}^{n}\sum_{q=1}^{m} Prob.(CS_{p,q}) = 1 \tag{8.4}$$

After inserting the values from equation (8.1) and solving for P($CS_{0,0}$)

$$P(CS_{0,0}) = \frac{1}{1 + \sum_{p=1}^{n} P_f^p) + \sum_{q=1}^{m} P_h^q + \sum_{p=1}^{n}\sum_{q=1}^{m}(p+q)P_f^p P_h^q} \tag{8.5}$$

In order to solve the equation above, we proceed as follows. The closed form solution for each of the summations in the denominator of the equation above is computed as follows:

$$
\begin{aligned}
A &= \sum_{p=1}^{n} P_f^p = \frac{P_f(P_f^n - 1)}{P_f - 1} \\
B &= \sum_{q=1}^{m} P_h^q = \frac{P_h(P_h^m - 1)}{P_h - 1} \\
C &= \sum_{p=1}^{n} pP_f^p \sum_{q=1}^{m} P_h^q = \frac{(nP_f - n - 1)(P_f^{n+1} + P_f)}{(1 - P_f)^2} * \frac{P_h(P_h^m - 1)}{(P_h - 1)} \\
D &= \sum_{p=1}^{n} P_f^p \sum_{q=1}^{m} qP_h^q = \frac{(mP_h - m - 1)(P_h^{m+1} + P_h)}{(1 - P_h)^2} * \frac{P_f(P_f^n - 1)}{P_f - 1}
\end{aligned} \tag{8.6}
$$

Simplifying equation (8.5) after inserting values from equation (8.6) respectively, the probability of state $CS_{0,0}$ is given as:

$$
P(CS_{0,0}) = \frac{1}{1 + A + B + C + D} \tag{8.7}
$$

To compute the expected mobility signalling in case of finite HOs mobility, we expand the equation (8.2) and insert the values from (8.6) and (8.7) as follows:

$$
\begin{aligned}
E[CS_{p,q}] = CS_{0,0} * P(CS_{0,0}) + \sum_{p=1}^{n} CS_{p,0} * P(CS_{p,0}) + \\
\sum_{q=1}^{m} CS_{0,q} * P(CS_{0,q}) + \sum_{p=1}^{n} \sum_{q=1}^{m} CS_{p,q} * P(CS_{p,q})
\end{aligned} \tag{8.8}
$$

Solving the equation above after including the values above.

$$
\begin{aligned}
E[CS_{p,q}] = \sum_{p=1}^{n} p * S_f * P_f^p * P(CS_{0,0}) + \sum_{q=1}^{m} q * S_s * P_h^q * P(CS_{0,0}) + \\
\sum_{p=1}^{n} \sum_{q=1}^{m} (pS_f + qS_s) * (p + q) * P_f^p * P_h^q * P(CS_{0,0})
\end{aligned} \tag{8.9}
$$

After algebraic manipulations of the above equation and inserting the values, the CN mobility signalling for mobile users in case of given HO successes and failures is as follows. Solving the equation (8.9) above after inserting values, it becomes:

$$E[CS_{p,q}] = \Big(S_1 \mathrm{S}_f + S_2 \mathrm{S}_s\Big) * \mathrm{P}(CS_{0,0}) \tag{8.10}$$

Where S_1 and S_2 are given as:

$$S_1 = \frac{P_f(nP_f^{n+1} - (n+1)P_f^n + 1)}{(P_f - 1)^2} + \frac{P_f(P_f+1) - P_f^{n+1}(n^2P_f^2 - (2n^2+2n-1)P_f + (n+1)^2}{(1-P_f)^3} * \frac{P_h(P_h^m - 1)}{(P_h - 1)} + \frac{P_f(nP_f^{n+1} - (n+1)P_f^n + 1)}{(P_f - 1)^2} * \frac{P_h(mP_h^{m+1} - (m+1)P_h^m + 1)}{(P_h - 1)^2}$$

$$S_2 = \frac{P_h(mP_h^{m+1} - (m+1)P_h^m + 1)}{(P_h - 1)^2} + \frac{P_h(P_h+1) - P_h^{m+1}(m^2P_h^2 - (2m^2+2m-1)P_h + (m+1)^2}{(1-P_h)^3} * \frac{P_f(P_f^n - 1)}{(P_f - 1)} + \frac{P_f(nP_f^{n+1} - (n+1)P_f^n + 1)}{(P_f - 1)^2} * \frac{P_h(mP_h^{m+1} - (m+1)P_h^m + 1)}{(P_h - 1)^2}$$

8.3 Numerical Results

In this section we verify the numerical results in the case of finite HOs mobility signalling in case of CDSA and compared it with conventional network deployment.

8.3.1 Finite Handovers and Mobility Signalling

In Fig 8.2 we show the expected normalized mobility signalling in case of 50 HO successes and 20 HO failures for an intercell coverage overlap factor of $c = 0.1$. Whereas, in Fig. 8.3 expected normalized mobility signalling is shown for similar number of HOs success and failure as in Fig 8.2 except for intercell coverage overlap value $c = 0.4$. It is evident from these two plots that the maximum value of signalling generated is the same at high velocity in

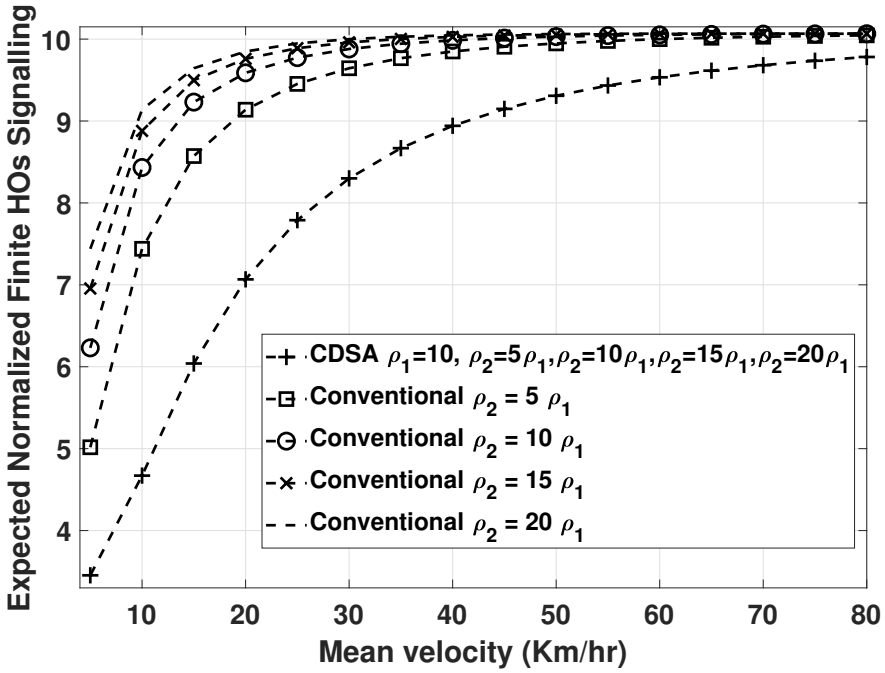

Fig. 8.2: Expected normalized finite HOs signalling vs. velocity for $c = 0.1$, HO successes m=50, HO failures n = 20 and Session Duration E[λ]=5 mins. Expected mobility signalling is normalized with S_f and S_s

both cases. However, during normal to medium mobility scenarios, signalling generated in case of large intercell coverage overlap value $c = 0.4$ is lower than smaller value $c = 0.1$ as visible in normal to medium velocity regions of Figs. 8.2 and 8.3 respectively. This insight tells us that for a fixed number of HOs allowed in a network, intercell coverage overlap factor decides the amount of mobility signalling generated in a network.

Additionally, it is also shown that CDSA results in least amount of mobility signalling compared to conventional networks in case of limited number of HOs allowed in a network. This informs the reader that in order to maintain quality of service for a given network when only limited number of HOs allowed CDSA ought be a clear choice over conventional network.

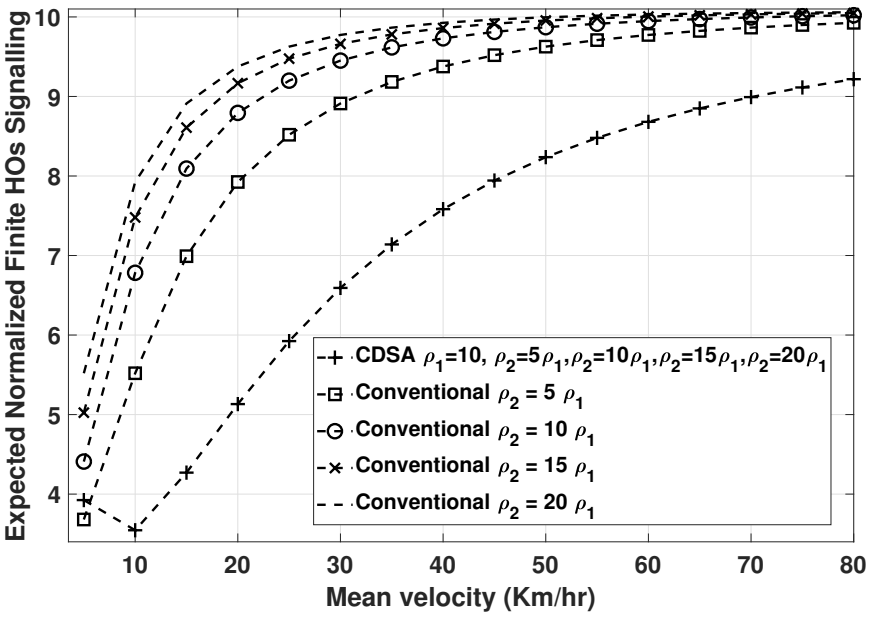

Fig. 8.3: Expected normalized finite HOs signalling vs. velocity for Intercell Coverage Overlap $c = 0.4$, HO successes m = 50, HO failures n = 20 and Session Duration E[λ]=5 mins. Expected mobility signalling is normalized with S_f and S_s

8.3.2 Intercell Coverage Overlap and Finite HOs Signalling

In Fig. 8.4, we have shown normalized expected mobility signalling in case of finite HOs vs. velocity for different values of intercell coverage overlap factor ranging from 0.1 to 0.5.

The curve trend shows that signalling increases with increase in velocity but this increase is the least for high values of coverage factor such as c=0.5. This insight is helpful for network planners and engineers from design perspective. While planning a network, high values of intercell coverage factor value can result in least amount of signalling. It also increases the probability of HO success.

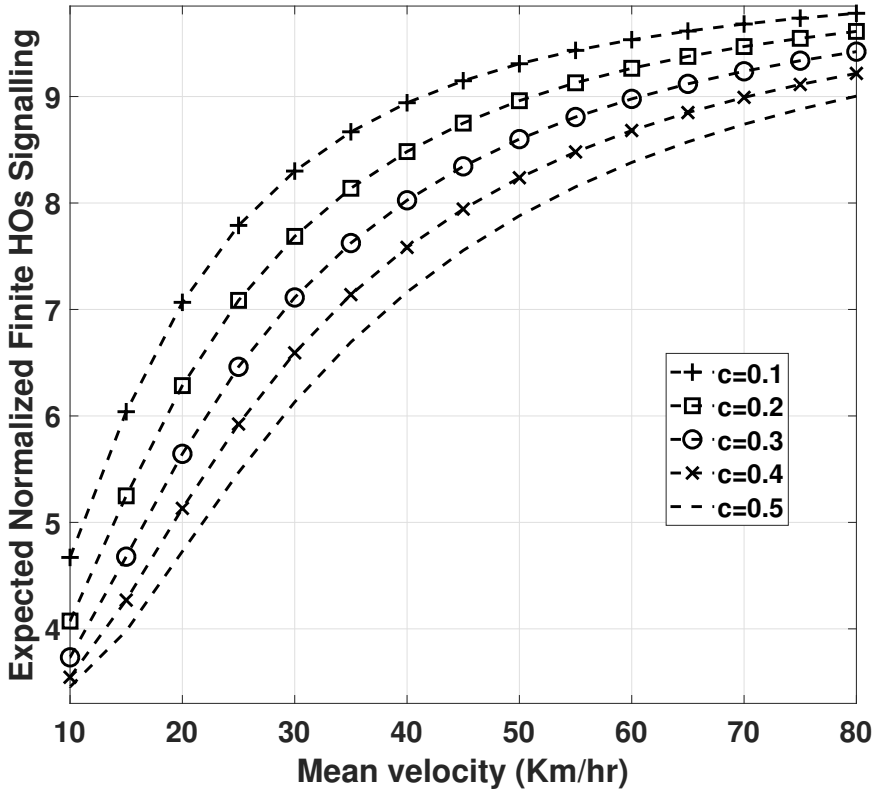

Fig. 8.4: Expected normalized finite HOs signalling vs. velocity for different values of c, $\mathrm{E}[\rho] = 10$ and $\mathrm{E}[\lambda] = 5$ mins. Expected mobility signalling is normalized with S_f and S_s

8.4 Summary

In this chapter, we developed a framework to analyze the effect of mobility signalling overhead for a given number of HO successes and failures respectively. A continuous mobility user is considered in this scenario, having handover success or handover failure. The Markov chain is used to evaluate the average mobility signalling in this case.

It is worth noting that analysis in earlier chapters considered infinte number of HOs thus provided insights into asymptotic behavior of amount of signalling generated in different scenarios. Whereas, the analytical framework derived in this chapter allows calculation of HO signalling for given number of HO failures and success. This framework thus can be readily used to calculate mobility signalling overhead in a real network for both CDSA and conventional HetNet deployments.

CHAPTER 9

Conclusion and Future work

9.1 Conclusions

This dissertation presented the handover model in active mode and the associated signalling generated in case of a handover. Analytical models were derived to investigate the probability of signalling in case of handover success, handover failure and no handover scenarios. Mobility signalling model(s) were presented to quantify how much of an overall signalling load is generated as a result of handover (success and failure) in active mode by a UE. In addition, a new parameter (intercell shared coverage factor c) was introduced which affects the overall signalling load both in case of handover success and failure. The performance of the proposed analytical model was evaluated using conventional architecture deployment versus CDSA. Numerical results show that CDSA-based network deployment can perform better overall while resulting in much lesser amount of signalling in handover success and handover failure scenarios when compared to conventional network deployment under various cell densities. The proposed dissertation gives insight and directions to the reader regarding what approach should be taken in deploying future ultra dense networks. This new parameter introduced is equally applicable, as it can save the magnitude of signalling load generated during handover by proper setting for a given cell density and velocity. These insights can help in better tuning of handover parameters, fewer call drops and session failures. Ultimately it will improve the operational costs of the network.

Additionally, we also developed an analytical model to quantify how much

of mobility signalling reduction can be achieved by choosing a proper cell overlap fraction for a given velocity and cell density. A quantitative method to plan cell overlaps for minimizing mobility signalling does not yet exist in commercial cellular network planning. The presented framework quantifies mobility signalling overhead as a function of cell overlap fraction for given velocity and cell densities, and thus is the first attempt towards this goal. The design of optimal cell overlap that minimizes the mobility signalling, as enabled by the presented framework has potential to become an integral part of cell planning in 5G and beyond.

The analytical insights provided by this dissertation quantify the advantages of CDSA over conventional network in terms of mobility management. In CDSA a continuous and reliable coverage layer is provided by CBS, where the large footprint ensures robust connectivity and mobility. Conversely, the data plane is supported by flexible, adaptive, high capacity and energy efficient DBSs, which provide data transmission along with the necessary signalling, as shown in Fig. 1.1 in Chapter 1. In a conventional network, the network remains on all the time and signalling and data connectivity operations are controlled by the eNodeB alone. Every single HO generates CN signalling which adds load on the network elements and increases delay. In case of CDSA, any HO between DBS to DBS is transparent to the CN and does not generate any CN signalling. This saves a lot of capacity and resources of the CN. The only time when CN mobility signalling is generated in CDSA is when a HO takes place between CBS to CBS.

From the results presented in this dissertation, it is clear that CDSA is a clear winner when it comes to ultra dense HetNet deployment.

9.2 Future work

The work presented in this dissertation assumes too-late handover as the reason for handover failure. Handover failure can take place due to various reasons other than too late handover, such as transport network reliability, i.e., S1 interface is down, poor RF conditions, radio link failure and partial handover. In future work we plan to compute handover signalling as a result of various handover failure reasons mentioned earlier.

The presented analytical framework in this dissertation is applicable for both intra-frequency and inter-frequency handover scenarios. In the CDSA network, the CBS and DBS are usually deployed in separate frequency bands to avoid inter-layer interference. Although this may complicate the UE radio frequency design, a separate frequency deployment is being considered in the new radio guidelines of the 3GPP [103]. In this direction, HOs within the footprint of the same CBS require changing the DBS only, hence they are considered intra-frequency HOs. On the other hand, inter-CBS HOs, i.e., HOs between two different CBSs, require changing both the CBS and DBS, i.e., a two-link HO. Such a scenario may involve intra-frequency and inter-frequency HOs. Consequently, a longer or shorter measurement gap may be required depending upon the cell deployment density. According to Mahbas et al [104], using smaller values of measurement gap, better system performance can be achieved in case of dense cell density and higher values of measurement gap in case of sparse cell density. This issue can be solved by limiting the number of CBSs / DBSs that are being monitored by the UE, e.g., the UE monitors the top n-cells per cell categorization to ensure that data transmission is balanced against the accurate measurement cycle. However, modeling this specific scenario is beyond the scope of this dissertation, and can be a goal pursued in

future work.

Bibliography

[1] C. Inc., "Cisco visual netnetwork index: Global mobile data traffic forecast update, 2016-2021," Tech. Rep., 2017.

[2] N. S. Networks, "2020: Beyond 4g, radio evolution for the gigabit experience," White Paper, aug 2011. [Online]. Available: http://nsn.com/file/15036/2020-beyond-4g-radio-evolution-for-the-gigabit-experience

[3] A. Imran, A. Zoha, and A. Abu-Dayya, "Challenges in 5g: how to empower son with big data for enabling 5g," *IEEE Network*, vol. 28, no. 6, pp. 27–33, Nov 2014.

[4] R. Yang, B. Lin, H. Yang, P. Si, E. Sun, and Y. Zhang, "The transmission optimization of the massive mimo hetnets," in *2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, Oct 2017, pp. 1–6.

[5] M. Agiwal, A. Roy, and N. Saxena, "Next generation 5g wireless networks: A comprehensive survey," *IEEE Communications Surveys Tutorials*, vol. 18, no. 3, pp. 1617–1655, thirdquarter 2016.

[6] J. G. Andrews, H. Claussen, M. Dohler, S. Rangan, and M. C. Reed, "Femtocells: Past, present, and future," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 3, pp. 497–508, April 2012.

[7] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5g be?" *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1065–1082, June 2014.

[8] N. Bhushan, J. Li, D. Malladi, R. Gilmore, D. Brenner, A. Damnjanovic, R. T. Sukhavasi, C. Patel, and S. Geirhofer, "Network densification: the dominant theme for wireless evolution into 5g," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 82–89, February 2014.

[9] Huawei, "Whitepaper: Five trends to small cells 2020," Barcelona.

[10] A. Mohamed, M. Imran, P. Xiao, and R. Tafazolli, "Memory-full context-aware predictive mobility management in dual connectivity 5g networks," *IEEE Access*, vol. PP, no. 99, pp. 1–1, 2018.

[11] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5g be?" *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1065–1082, June 2014.

[12] P. Banelli, S. Buzzi, G. Colavolpe, A. Modenini, F. Rusek, and A. Ugolini, "Modulation formats and waveforms for 5g networks: Who will be the heir of ofdm?: An overview of alternative modulation schemes for improved spectral efficiency," *IEEE Signal Processing Magazine*, vol. 31, no. 6, pp. 80–93, Nov 2014.

[13] D. Liu, W. Hong, T. S. Rappaport, C. Luxey, and W. Hong, "What will 5g antennas and propagation be?" *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 12, pp. 6205–6212, Dec 2017.

[14] N. Zhang, P. Yang, J. Ren, D. Chen, L. Yu, and X. Shen, "Synergy of big data and 5g wireless networks: Opportunities, approaches, and challenges," *IEEE Wireless Communications*, vol. 25, no. 1, pp. 12–18, February 2018.

[15] A. Gavras, S. Denazis, C. Tranoris, H. Hrasnica, and M. B. Weiss, "Requirements and design of 5g experimental environments for vertical industry innovations," in *2017 Global Wireless Summit (GWS)*, Oct 2017, pp. 165–169.

[16] S. N. Khan, L. Goratti, R. Riggio, and S. Hasan, "On active, fine-grained ran and spectrum sharing in multi-tenant 5g networks," in *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Oct 2017, pp. 1–5.

[17] W. Kiess, M. R. Sama, J. Varga, J. Prade, H. J. Morper, and K. Hoffmann, "5g via evolved packet core slices: Costs and technology of early deployments," in *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Oct 2017, pp. 1–7.

[18] J. Alcaraz-Calero, I. P. Belikaidis, C. J. B. Cano, P. Bisson, D. Bourse,

M. Bredel, D. Camps-Mur, T. Chen, X. Costa-Perez, P. Demestichas, M. Doll, S. E. Elayoubi, A. Georgakopoulos, A. MÃďmmelÃď, H. P. Mayer, M. Payaro, B. Sayadi, M. S. Siddiqui, M. Tercero, and Q. Wang, "Leading innovations towards 5g: Europe's perspective in 5g infrastructure public-private partnership (5g-ppp)," in *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Oct 2017, pp. 1–5.

[19] Y. Li, B. Cao, and C. Wang, "Handover schemes in heterogeneous lte networks: challenges and opportunities," *IEEE Wireless Communications*, vol. 23, no. 2, pp. 112–117, April 2016.

[20] B. B. SÃąnchez, Ã. SÃąnchez-Picot, and D. S. D. Rivera, "Using 5g technologies in the internet of things handovers, problems and challenges," in *2015 9th International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing*, July 2015, pp. 364–369.

[21] G. Hu, A. Huang, R. He, B. Ai, and Z. Chen, "Theory analysis of the handover challenge in express train access networks (etan)," *China Communications*, vol. 11, no. 7, pp. 92–98, July 2014.

[22] S. Park, Y. Lee, and P. Kim, "Challenges of seamless handover for merging wired and wireless infrastructures," in *2005 IEEE 61st Vehicular Technology Conference*, vol. 4, May 2005, pp. 2216–2219 Vol. 4.

[23] L. Taylor, R. Titmuss, and C. Lebre, "The challenges of seamless handover in future mobile multimedia networks," *IEEE Personal Communications*, vol. 6, no. 2, pp. 32–37, Apr 1999.

[24] A. S. Priyadharshini and P. T. V. Bhuvaneswari, "Investigation on handover performance analysis through control parameter configuration in lte-a hetnet," in *2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, March 2017, pp. 1883–1887.

[25] S. Meenakshi, J. Chaubey, and K. V. Babu, "Handover requirement analysis in device-to-device communication for next generation networks," in *2017 International Conference on Communication and Signal Processing (ICCSP)*, April 2017, pp. 0101–0105.

[26] M. B. Yassein, S. Aljawarneh, and W. Al-Sarayrah, "Mobility management of internet of things: Protocols, challenges and open issues," in *2017 International Conference on Engineering MIS (ICEMIS)*, May 2017, pp. 1–8.

[27] "Study on small cell enhancements for e-utra and e-utran: Higher layer aspects," *3GPP TR 36.842*, no. Release 12, 2013. [Online]. Available: http://www.3gpp.org/DynaReport/36842.htm

[28] "Mobility enhancements in heterogeneous networks," 3GPP TR 36.839, Tech. Rep., 2012.

[29] Y. K. H. Ishii and H. Takahashi, "A novel architecture for lte-b c-plane/u-plane split and phantom cell concept," *IEEE Globecom Workshops*, pp. 624 – 630, 2012.

[30] A. Mohamed, O. Onireti, M. A. Imran, A. Imran, and R. Tafazolli, "Control-data separation architecture for cellular radio access networks: A survey and outlook," *IEEE Communications Surveys Tutorials*, vol. 18, no. 1, pp. 446–465, Firstquarter 2016.

[31] A. Taufique, M. Jaber, A. Imran, Z. Dawy, and E. Yacoub, "Planning wireless cellular networks of future: Outlook, challenges and opportunities," *IEEE Access*, vol. 5, pp. 4821–4845, 2017.

[32] J. Kang, O. Simeone, J. Kang, and S. Shamai, "Control-data separation across edge and cloud for uplink communications in c-ran," in *2017 IEEE Wireless Communications and Networking Conference (WCNC)*, March 2017, pp. 1–6.

[33] J. Gang and V. Friderikos, "Control plane load balancing in wireless c/u split architectures," in *2016 IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Sept 2016, pp. 1–6.

[34] A. Zakrzewska, D. LÃşpez-PÃľrez, S. Kucera, and H. Claussen, "Dual connectivity in lte hetnets with split control- and user-plane," in *2013 IEEE Globecom Workshops (GC Wkshps)*, Dec 2013, pp. 391–396.

[35] X. Weiliang, W. Zhouyun, Y. Fengyi, and Y. Tao, "Analysis and field trial results on c-plane and u-plane split scheme in virtual sectorization system," *China Communications*, vol. 12, no. 11, pp. 1–9, November 2015.

[36] I. G. et al, "Earth project deliverable d3.3: Final report on green network technologies,," 2012. [Online]. Available: http://bscw.ict-earth.eu/pub/bscw.cgi/d70472/EARTHWP3D3.3.pdf

[37] X. Xu, G. He, S. Zhang, Y. Chen, and S. Xu, "On functionality separation for green mobile networks: concept study over lte," *IEEE Communications Magazine*, vol. 51, no. 5, pp. 82–90, May 2013.

[38] C. H. K. S. Liu, J. Wu and V. Lau, "A 25 gb/s(/km2) urban wireless network beyond imt-advanced," *IEEE Communication Magazine*, vol. 49, no. 2, pp. 122 – 129, February 2011.

[39] A. Capone, A. F. dos Santos, I. Filippini, and B. Gloss, "Looking beyond green cellular networks," in *2012 9th Annual Conference on Wireless On-Demand Network Systems and Services (WONS)*, Jan 2012, pp. 127–130.

[40] J. Zhang, J. Feng, C. Liu, X. Hong, X. Zhang, and W. Wang, "Mobility enhancement and performance evaluation for 5g ultra dense networks," in *2015 IEEE Wireless Communications and Networking Conference (WCNC)*, March 2015, pp. 1793–1798.

[41] G. T.-R. W. Meeting, "Mobility statistics for macro and small cell dual-connectivity cases," 3GPP Nokia Siemens Networks, Chicago, IL,USA, techreport, apr 2013. [Online]. Available: http://www.3gpp.org/DynaReport/TDocExMtgâĂŞR2-81bâĂŞ30048.htm

[42] "Feasible senarios and benefits of dual-connectivity in small cell deployment," 3GPP TSG-RAN WG2 Meeting 81, 3GPP Huawei, HiSilicon, St.Julian's , Malta, Tech. Rep., feb 2013. [Online]. Available: Available:http://www.3gpp.org/DynaReport/TDocExMtgâĂŞR2-81âĂŞ30047.htm

[43] A. Mohamed, O. Onireti, M. A. Imran, A. Imran, and R. Tafazolli, "Predictive and core-network efficient rrc signalling for active state handover

in rans with control/data separation," *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1423–1436, March 2017.

[44] M. I. Poulakis, A. G. Gotsis, and A. Alexiou, "Multi-cell device-to-device communications: A spectrum sharing and densification study," *IEEE Vehicular Technology Magazine*, vol. PP, no. 99, pp. 1–1, 2018.

[45] J. Liu, M. Sheng, L. Liu, and J. Li, "Network densification in 5g: From the short-range communications perspective," *IEEE Communications Magazine*, vol. 55, no. 12, pp. 96–102, DECEMBER 2017.

[46] F. Lagum, I. Bor-Yaliniz, and H. Yanikomeroglu, "Strategic densification with uav-bss in cellular networks," *IEEE Wireless Communications Letters*, vol. PP, no. 99, pp. 1–1, 2017.

[47] M. I. Poulakis, A. G. Gotsis, and A. Alexiou, "Multi-cell device-to-device communication benefits in the presence of densification," in *2017 IEEE International Conference on Communications Workshops (ICC Workshops)*, May 2017, pp. 1335–1340.

[48] M. Bembe, G. Sibiya, and Y. Han, "Spectral efficient cell selection for lte-advanced's network densification," in *2017 14th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*, June 2017, pp. 1–2.

[49] W. Feng, Y. Wang, D. Lin, N. Ge, J. Lu, and S. Li, "When mmwave communications meet network densification: A scalable interference coordination perspective," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 7, pp. 1459–1471, July 2017.

[50] V. M. Nguyen and M. Kountouris, "Performance limits of network densification," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 6, pp. 1294–1308, June 2017.

[51] M. Xu, J. H. Yan, J. Zhang, F. Lu, J. Wang, L. Cheng, D. Guidotti, and G. K. Chang, "Bidirectional fiber-wireless access technology for 5g mobile spectral aggregation and cell densification," *IEEE/OSA Journal of Optical Communications and Networking*, vol. 8, no. 12, pp. B104–B110, December 2016.

[52] F. E. Idachaba, "5g networks: Open network architecture and densification strategies for beyond 1000x network capacity increase," in *2016 Future Technologies Conference (FTC)*, Dec 2016, pp. 1265–1269.

[53] J. G. Andrews, X. Zhang, G. D. Durgin, and A. K. Gupta, "Are we approaching the fundamental limits of wireless network densification?" *IEEE Communications Magazine*, vol. 54, no. 10, pp. 184–190, October 2016.

[54] B. Romanous, N. Bitar, A. Imran, and H. Refai, "Network densification: Challenges and opportunities in enabling 5g," in *2015 IEEE 20th International Workshop on Computer Aided Modelling and Design of Communication Links and Networks (CAMAD)*, Sept 2015, pp. 129–134.

[55] X. Zhang, *LTE optimization engineering handbook*, 2018.

[56] "Evolved universal terrestrial radio access (e-utra); "mobility enhancements in heterogeneous networks,"," 3GPP TR 36.839, Sophia Antipolis Cedex, France, Tech. Rep., dec 2012.

[57] B. Wegmann, A. Awada, K. Kordybach, and I. Viering, "Inter-rat mro in 3gpp rel.11: What works, and what does not," in *2013 IEEE 77th Vehicular Technology Conference (VTC Spring)*, June 2013, pp. 1–5.

[58] A. Awada, B. Wegmann, I. Viering, and A. Klein, "A location-based self-optimizing algorithm for the inter-rat handover parameters," in *2013 IEEE International Conference on Communications (ICC)*, June 2013, pp. 6168–6173.

[59] F. E. Morabet, A. E. Moussaoui, and N. Aknin, "Adaptive handover time scheme for dynamic load balancing in lte systems," in *2012 International Conference on Multimedia Computing and Systems*, May 2012, pp. 578–582.

[60] T. H. Kim, Q. Yang, J. H. Lee, S. G. Park, and Y. S. Shin, "A mobility management technique with simple handover prediction for 3g lte systems," in *2007 IEEE 66th Vehicular Technology Conference*, Sept 2007, pp. 259–263.

[61] W. Zizhou, F. Chen, W. Yafeng, and Y. Dacheng, "A novel network architecture for 3g evolution," in *2006 IEEE 17th International Symposium on Personal, Indoor and Mobile Radio Communications*, Sept 2006, pp. 1–5.

[62] M. Wang, E. Ramos, Y. P. E. Wang, N. Lidian, S. Nammi, and M. Curran, "Evaluation of mobility performance and deployment scenarios in umts heterogeneous networks," in *2014 IEEE 79th Vehicular Technology Conference (VTC Spring)*, May 2014, pp. 1–5.

[63] L. M. Abdullah, M. D. Baba, and S. G. A. Ali, "Parameters optimization for handover between femtocell and macrocell in lte-based network," in *2014 IEEE International Conference on Control System, Computing and Engineering (ICCSCE 2014)*, Nov 2014, pp. 636–640.

[64] S. Ranjan, P. Singh, M. Mehta, P. Rathod, N. Akhtar, and A. Karandikar, "A self-configured vertical handover algorithm for lte and wlan interworking," in *2015 Twenty First National Conference on Communications (NCC)*, Feb 2015, pp. 1–6.

[65] A. S. Priyadharshini and P. T. V. Bhuvaneswari, "A study on handover parameter optimization in lte-a networks," in *2016 International Conference on Microelectronics, Computing and Communications (MicroCom)*, Jan 2016, pp. 1–5.

[66] E. Demarchou, C. Psomas, and I. Krikidis, "Mobility management in ultra-dense networks: Handover skipping techniques," *IEEE Access*, vol. PP, no. 99, pp. 1–1, 2018.

[67] F. Ali, S. Jangsher, and F. A. Bhatti, "Resource sharing for d2d communication in multi small cell networks," in *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Oct 2017, pp. 1–5.

[68] Y. Chen, Z. Yang, and H. Zhang, "Opportunistic-based dynamic interference coordination in dense small cells deployment," in *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Oct 2017, pp. 1–5.

[69] J. Guo, H. Zhang, L. Yang, H. Ji, and X. Li, "Decentralized computation

offloading in mobile edge computing empowered small-cell networks," in *2017 IEEE Globecom Workshops (GC Wkshps)*, Dec 2017, pp. 1–6.

[70] I. Allal, B. Mongazon-Cazavet, K. A. Agha, S. M. Senouci, and Y. Gourhant, "A green small cells deployment in 5g x2014; switch on/off via iot networks energy efficient mesh backhauling," in *2017 IFIP Networking Conference (IFIP Networking) and Workshops*, June 2017, pp. 1–2.

[71] A. Paatelma, D. H. Nguyen, H. Saarnisaari, N. Kandasamy, and K. R. Dandekar, "Reinforcement learning system to mitigate small-cell interference through directionality," in *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Oct 2017, pp. 1–7.

[72] L. Qiang, J. Li, and C. Touati, "A user centered multi-objective handoff scheme for hybrid 5g environments," *IEEE Transactions on Emerging Topics in Computing*, vol. 5, no. 3, pp. 380–390, July 2017.

[73] F. B. Tesema, A. Awada, I. Viering, M. Simsek, and G. P. Fettweis, "Evaluation of adaptive active set management for multi-connectivity in intra-frequency 5g networks," in *2016 IEEE Wireless Communications and Networking Conference*, April 2016, pp. 1–6.

[74] J. StaÅĎczak, "Mobility enhancements to reduce service interruption time for lte and 5g," in *2016 IEEE Conference on Standards for Communications and Networking (CSCN)*, Oct 2016, pp. 1–5.

[75] A.-L. S. B. Nokia, "'"make-before-break" ho in lte and its implications," 3GPP TSG RAN2 Meeting 94, 2016.

[76] P. Kela, X. Gelabert, J. Turkka, M. Costa, K. Heiska, K. LeppÃď nen, and C. Qvarfordt, "Supporting mobility in 5g: A comparison between massive mimo and continuous ultra dense networks," in *2016 IEEE International Conference on Communications (ICC)*, May 2016, pp. 1–6.

[77] J. S. Kim, W. J. Lee, and M. Y. Chung, "A multiple beam management scheme on 5g mobile communication systems for supporting high mobility," in *2016 International Conference on Information Networking (ICOIN)*, Jan 2016, pp. 260–264.

[78] A. S. Cacciapuoti, "Mobility-aware user association for 5g mmwave networks," *IEEE Access*, vol. 5, pp. 21 497–21 507, 2017.

[79] B. Lannoo, D. Colle, M. Pickavet, and P. Demeester, "Radio-over-fiber-based solution to provide broadband internet access to train passengers [topics in optical communications]," *IEEE Communications Magazine*, vol. 45, no. 2, pp. 56–62, Feb 2007.

[80] M. D. Cia, F. Mason, D. Peron, F. Chiariotti, M. Polese, T. Mahmoodi, M. Zorzi, and A. Zanella, "Using smart city data in 5g self-organizing networks," *IEEE Internet of Things Journal*, vol. PP, no. 99, pp. 1–1, 2017.

[81] J. Wu, J. Liu, Z. Huang, and S. Zheng, "Dynamic fuzzy q-learning for handover parameters optimization in 5g multi-tier networks," in *2015 International Conference on Wireless Communications Signal Processing (WCSP)*, Oct 2015, pp. 1–5.

[82] F. B. Tesema, A. Awada, I. Viering, M. Simsek, and G. Fettweis, "Evaluation of context-aware mobility robustness optimization and multi-connectivity in intra-frequency 5g ultra dense networks," *IEEE Wireless Communications Letters*, vol. 5, no. 6, pp. 608–611, Dec 2016.

[83] L. L. Vy, L. P. Tung, and B. S. P. Lin, "Big data and machine learning driven handover management and forecasting," in *2017 IEEE Conference on Standards for Communications and Networking (CSCN)*, Sept 2017, pp. 214–219.

[84] U. Karneyenka, K. Mohta, and M. Moh, "Location and mobility aware resource management for 5g cloud radio access networks," in *2017 International Conference on High Performance Computing Simulation (HPCS)*, July 2017, pp. 168–175.

[85] B. L. Dang, R. V. Prasad, I. Niemegeers, M. G. Larrode, and A. M. J. Koonen, "Toward a seamless communication architecture for in-building networks at the 60 ghz band," in *Proceedings. 2006 31st IEEE Conference on Local Computer Networks*, Nov 2006, pp. 300–307.

[86] N. Pleros, K. Tsagkaris, and N. D. Tselikas, "A moving extended cell concept for seamless communication in 60 ghz radio-over-fiber networks,"

IEEE Communications Letters, vol. 12, no. 11, pp. 852–854, November 2008.

[87] F. A. Hossain and A. M. Chowdhury, "User mobility prediction based handoff scheme for 60 ghz radio over fiber network," in *2014 IEEE 11th Consumer Communications and Networking Conference (CCNC)*, Jan 2014, pp. 557–562.

[88] X. Ge, B. Du, Q. Li, and D. S. Michalopoulos, "Energy efficiency of multiuser multiantenna random cellular networks with minimum distance constraints," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 2, pp. 1696–1708, Feb 2017.

[89] R. Balakrishnan and B. Canberk, "Traffic-aware qos provisioning and admission control in ofdma hybrid small cells," *IEEE Transactions on Vehicular Technology*, vol. 63, no. 2, pp. 802–810, Feb 2014.

[90] S. Bu, F. R. Yu, and H. Yanikomeroglu, "Interference-aware energy-efficient resource allocation for ofdma-based heterogeneous networks with incomplete channel state information," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 3, pp. 1036–1050, March 2015.

[91] F. Guidolin, I. Pappalardo, A. Zanella, and M. Zorzi, "Context-aware handover policies in hetnets," *IEEE Transactions on Wireless Communications*, vol. 15, no. 3, pp. 1895–1906, March 2016.

[92] H. Ibrahim, H. ElSawy, U. T. Nguyen, and M. S. Alouini, "Mobility-aware modeling and analysis of dense cellular networks with c -plane/ u -plane split architecture," *IEEE Transactions on Communications*, vol. 64, no. 11, pp. 4879–4894, Nov 2016.

[93] R. Arshad, H. ElSawy, S. Sorour, T. Y. Al-Naffouri, and M. S. Alouini, "Velocity-aware handover management in two-tier cellular networks," *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1851–1867, March 2017.

[94] "Signaling is growing 50 percent faster than data traffic," *Nokia Siemens Networks*, 2012.

[95] L. Wang, Y. Zhang, and Z. Wei, "Mobility management schemes at radio network layer for lte femtocells," in *VTC Spring 2009 - IEEE 69th Vehicular Technology Conference*, April 2009, pp. 1–5.

[96] H. Zhang, W. Ma, W. Li, W. Zheng, X. Wen, and C. Jiang, "Signalling cost evaluation of handover management schemes in lte-advanced femtocell," in *2011 IEEE 73rd Vehicular Technology Conference (VTC Spring)*, May 2011, pp. 1–5.

[97] H. Xie, S. K. Xie, and S. Kuek, "Priority handoff analysis," *IEEE Vehicular Technology Conference*, pp. 855–858, 1993.

[98] V. Luarini, "From symmetry breaking to poisson point process in 2d voronoi tessellations: the generic nature of hexagons," *Statistical Physics*, vol. 130, no. 6, pp. 1047 – 1062, 2008.

[99] K. Dimou, M. Wang, Y. Yang, M. Kazmi, A. Larmo, J. Pettersson, W. Muller, and Y. Timner, "Handover within 3gpp lte: Design principles and performance," in *2009 IEEE 70th Vehicular Technology Conference Fall*, Sept 2009, pp. 1–5.

[100] A.Helenius, "Performance of handover in long term evolution," Master's thesis, School of Electrical Engineering, Aalto University, 2011.

[101] 3GPP, "Feasibility study for evolved universal terrestrial radio access (utra) and universal terrestrial radio access network (utran)," 3GPP, techreport, 2017.

[102] Z. Li and W. Mick, "User plane and control plane separation framework for home base stations," *Fujitsu scientific and technical journal*, vol. 46, no. 1, pp. 79–86, 2010.

[103] "Evolved universal terrestrial radio access (e-utra) and nr; multi-connectivity; stage 2 (release 15)," 3GPP TS 37.340, Tech. Rep., 2017.

[104] A. Mahbas, H. Zhu, and J. Wang, "The role of inter-frequency measurement in offloading traffic to small cells," in *2017 IEEE 85th Vehicular Technology Conference (VTC Spring)*, June 2017, pp. 1–4.