Mobility management for the information centric future internet

Muhammad Shoaib Saleem

THESE DE DOCTORAT CONJOINT TELECOM SUDPARIS et L'UNIVERSITE PIERRE ET MARIE CURIE

Spécialité
Informatique et Télécommunications

Ecole doctorale Informatique, Télécommunications et Electronique (EDITE) de Paris

Présentée par

Muhammad Shoaib SALEEM

Pour obtenir le grade de
DOCTEUR DE TELECOM SUDPARIS

Gestion de la Mobilité pour l'Internet du Futur Centré autour de l'Information

Soutenance le 19 Novembre, 2012

Devant le jury composé de :

Jalel Ben-Othman	**Rapporteur**	**Professeur, Université Paris 13, France**
Sidi-Mohammed Senouci	**Rapporteur**	**Professeur, Université Bourgogne, France**
Selma Boumerdassi	**Examinateur**	**Maitre de conférences (HDR), CNAM, France**
André-Luc Beylot	**Examinateur**	**Professeur, INP-ENSEEIHT, France**
Mathieu Jaume	**Examinateur**	**Maitre de Conférences (HDR), LIP6, France**
Eric Renault	**Directeur**	**Maitre de Conférences (HDR),Télécom SudParis, France**

Thèse n° 2012TELE0035

Dedication

To my Loving Parents & Siblings

Acknowledgements

So finally I am able to end this thesis by writing few lines to acknowledge the support of those who encouraged and helped me during my stay at Télécom SudParis. Last three and a half years were like a roller coaster ride for me. There were critical moments during this period where some came to my rescue while there were moments of joy and excitement which I shared with few more. I am deeply thankful to all of them for their valuable company.

I would first like to express my gratitude and thanks to the jury members for being part of my thesis committee. I am especially grateful to my reviewers, Professor Jalel Ben Othman and Professor Sidi Mohammad Senouci, for taking their time out of their busy schedules to review my thesis report.

Professor Éric Renault surely is the most important person during my stay at Télécom SudParis. Thanks a lot for all of your support during all these years. Thanks for listening to me patiently and appreciating my ideas and opinions.

I would like to thank RS2M department for providing support for my research. I really appreciate the administrative help from Valérie Mateus (Executive Assistant of RS2M department). My colleagues and friends at work: Khalil, Raheel, Imran, Farhan Hanif, Newaz and Sidharta, with whom i have spent a very good time, thank you all. My friends: Ayaz, Mubashir and Yasir, thank you for all those memorable weekends. A very special thanks to my friends Mr.& Mrs. Farhan Mirani and Mr.& Mrs. Abdul Malik Khan for their generous support and help during my period of illness. Thank you Ramya for being such a good friend and for all those morale boosting advices.

How can I forget my loving parents and my siblings. Your prayers, love and affection have always been with me. I am grateful to you for all your encouragements. I dedicate all my accomplishments to my family.

Muhammad Shoaib Saleem

Abstract

The contemporary Internet ecosystem today has gone through series of evolutionary changes during the last forty or fifty years. Though it was designed as a network with fixed nodes, it has scaled well enough with the development of new technologies both in fixed and wireless networks. Initially, the communication model of the Internet was based on the telephone network (and can be considered as the 1st Generation Internet). Later, its transition as a client-server model made it a network where communication systems exchange data over dedicated links. This 2nd Generation Internet, over the years, has been challenged by many problems and issues such as network congestion, path failure, DOS attacks, mobility issues for wireless networks, etc. The Internet users always look for some information, irrespectively where it is located or stored. This approach is the basic building block for a network architecture where information is considered as the premier entity. Such networks, in general, are termed as *Information Centric Network* (ICN), where information takes centric position superseding the node centric approach like in the current Internet. The problems faced by the current Internet architecture, mentioned above, can be handled with a unifying approach by putting the information at the center of the network architecture. On a global scale, this network architecture design is termed as the Future Information Centric Internet.

Similarly, Mobile Internet usage has increased overwhelmingly in the last decade. There has been an estimated 1.2 billion mobile broad-band subscriptions for 2.4 billion Internet users in 2011. Because of the increased spectrum efficiency and ubiquitous availability of cellular connectivity, the seamless mobility and connectivity is now considered as daily life commodity. However, in the case of the Internet, IP based mobility solutions cannot catch up in performance with the fast evolution of cellular networks. Therefore, one of the primary goals for the Future Internet is the design of mobility management schemes that overcome the issues in wireless networks such as handover and location management, multihoming, security, etc.

In this thesis, we have proposed a mobility management solution in wireless networks in the context of ICN in general and in the context of *Network of Information* (NetInf) in particular. NetInf is an ICN-based Future Internet architecture. We propose a NetInf Mobile Node (NetInf MN) architecture which is backward compatible with the current Internet architecture as well. This cross architecture design for mobility support works closely with Central Control Unit (CCU) (network entity) for improved performance in case of handover management in wireless networks. The Virtual Node Layer (VNL) algorithm explains how different modules of NetInf MN and CCU units work together. The game theoretical and Reinforcement Learning (CODIPAS-RL) scheme based mathematical

model shows how handover management and data relaying in the wireless networks can increase the network coverage through cooperative diversity. Simulation results show that the proposed model achieves both Nash and Stackelberg equilibria where as the selected CODIPAS-RL scheme reaches global optimum.

Finally, as a use case example of NetInf architecture, we propose the NetInf Email service that does not require dedicated servers or dedicated port unlike the current email service. The use of asymmetric keys as user's ID is the unique feature proposed for this service. The NetInf email service architecture framework presented, explains how different architectural components work together. We discuss different challenges and requirements related to this service. The prototype developed for the Network of Information will be used for the implementation of this service.

Keywords

Information Centric Network (ICN), Network of Information (NetInf), Mobility Management, NetInf Mobile Node (NetInf MN), Game Theory, Nash Equilibrium, Stackelberg Equilibrium, Reinforcement Learning (RL), NetInf Email Service.

Résumé

L'Internet, basé sur TCP/IP, a traversé série de changements évolutionnaires dans les quarante ou cinquante dernières années. Il a été conçu pour un réseau avec des noeuds fixes. Néanmoins, il s'est adapté assez bien avec le développement de nouvelles technologies de réseaux fixes et sans fil. Au début, le modéle de communication de l'Internet a été basé sur le réseau téléphonique (considéré comme 1er Generation Internet). Plus tard, il a été mis á jour comme un modéle "client-serveur" oú la systéme de communication fait l'échange de données sur les ports dédiées. Cette 2éme génération Internet, au cours des années, a été contestée par de nombreux problèmes tels que la congestion du réseau, panne de chémin, les attaques par déni de service (DOS), gestion de la mobilité pour les réseaux sans fil, etc. Les utilisateurs d'Internet cherchent toujours des informations, indépendamment leur lieu de stockage (noeud ou serveur). Cette approche est la base d'une architecture où l'information est considerée comme l'unité primaire. Ces réseaux, en général, sont appelés en tant que Network of Information (NetInf), où l'information prend une position centrée, remplaçant l'approche oú, le réseau est centrée autour de noeud comme l'Internet aujourd'hui. Les problèmes rencontrés par l'Internet mentionné ci-dessus peuvent être traitées avec une approche unificatrice en mettant l'information au centre de l'architecture du réseau. À l'échelle mondiale, cette conception de l'architecture du réseau est nommé comme Future Information Centric Internet.

En paralléle, l'utilisation de l'Internet mobile a été augmentée durant la dernière décennie. Il a été environ 1,2 milliard abonnements de mobile broad band pour 2,4 milliards d'utilisateurs d'Internet en 2011. En raison d'augmentation de l'efficacité spectrale et ubiquitaire disponibilité de la connectivité cellulaire, la mobilité sans couture et la connectivité est désormais considérée commodité essentielle de la vie quotidienne. Néanmoins, en cas d'Internet, les solutions de mobilité basées sur des protocols de la suite TCP/IP ne peuvent pas rattraper son retard dans la performance avec l'évolution rapide des réseaux cellulaires. Par conséquent, l'un des principaux objectifs pour l'internet du futur est de concevoir des systémes de gestion de mobilité qui permettent de surmonter les problèmes dans les réseaux sans fil tels que handover et la gestion de la localisation, multihoming, sécurité, etc.

Dans cette thèse, nous avons proposé une solution de gestion de mobilité dans les réseaux sans fil dans le cadre du Information Centric Networking (ICN) en général et dans le contexte de Network of Information (NetInf) en particulier. NetInf est une architecture du Futur Internet basée sur le concept du ICN. Nous proposons un noeud mobile qui s'appelle NetInf Mobile Node (NetInf MN). L'architecture de ce noeud est compatible avec l'architecture d'Internet basée sur TCP/TP. Ce conception de l'architecture travaille en collaboration avec Central Control Unit (CCU) (une entité proposée pour les réseaux sans fil) pour améliorer les performances en cas de handover dans les réseaux sans fil. La

Virtual Node Layer (VNL) algorithme explique comment les différents modules de NetInf MN avec les composants du CCU travailler ensemble. La modéle mathématique basé sur Théorie de Jeu et Renforcement Learning (CODIPAS-RL) montre comment handover et data relaying sont géré dans les réseaux sans fil. Les résultats des simulations montrent que le modéle proposé réalise à la fois de Nash et de Stackelberg équilibres alors que la régime de CODIPAS-RL atteint un optimum global.

Enfin, comme un exemple de cas d'utilisation de l'architecture du NetInf, nous proposons le NetInf Email Service qui ne requiert pas des serveurs et ports dédiés contrairement au service e-mail existante. L'utilisation de clés asymétriques comme ID de l'utilisateur est la caractéristique unique proposée pour ce service. Le NetInf Email service architecture présenté, explique comment différents éléments architecturaux travail ensemble. Nous discutons des défis différents et des besoins relatifs à ce service. Le prototype développé pour NetInf sera utilisée pour la mise en oeuvre de ce service.

Mots-clés

Information Centric Network (ICN), Network of Information (NetInf), Gestion de la mobilité, NetInf Mobile Node (NetInf MN), Théorie ded Jeu, Équilibre de Nash, Équilibre de Stackelberg, Apprentissage par renforcement (RL), CODIPAS-RL, NetInf Email Service

Table of contents

Chapter 1

Introduction

Contents

The Internet today has evolved as a media entity where the creation and consumption of information (especially audio and visual) content is going to continuously grow in the coming years. This trend has resulted changes in the ways the content is published and distributed. The current Internet architecture was not designed to contest with this continuous growing amount of data. Peer-to-Peer (P2P) overlays and Content Distribution Networks (CDN) played an important role by sharing this load. The architecture that was designed on the basis of node centric paradigm has now transformed into content/information sharing platform. The end user seeks information (he/she is looking for on the web) whereas, the location of that information is irrelevant in this case. In fact, the majority of the Internet traffic today is the information dissemination as mentioned by Van Jacobson [1]. The Internet architecture has evolved for all these years and has somehow managed many challenges but the exponential growth of Internet users in recent years demands better service in terms of security, mobility, content distribution, etc. However, with such fast growth of content and users simultaneously, the incremental changes

or solutions for the current Internet architecture will hardly improve anything. Does this mean that the basic Internet architecture requires major overhauling? Or do we need a network where information should be independent from its locations and is ubiquitously available?

Many proposals have been given in recent years for a clean-slate design for the architecture of the Future Internet [1], [2], [3], [4]. The notion of *Information Centric Networking* (ICN) is the cornerstone of all these projects. The contemporary Internet architecture has IP addresses for the locations where information is stored. This approach has made every layer of TCP/IP protocol suite diverse with a wide range of protocols for different operations across the network. However, IP addressing has its limitation when it comes to ICN paradigm. In ICN, Information Objects (IOs) are location independent and information mobility does not change their ID unlike in the current Internet where end user's mobile device IP address changes whenever it changes its connection. This category of Future Internet architectures has addressed different challenges faced by current Internet architecture such as privacy and security, scalability, content naming and addressing, name resolution, routing and mobility. This thesis primarily addresses mobility issue in Information Centric Networks based on 4WARD [3] project's work package 6 called *Network of Information* (NetInf). The details of NetInf is provided in Chap.2. In the following section, we discuss the mobility management in ICN.

1.1 Mobility in Information Centric Networks

The current Internet architecture had to face mobility challenge when wireless networks integrated the Internet in their architectures. Though, the solution was provided in the form of Mobile IP [5], [6] and other location independent protocols like Host Identification Protocol (HIP) [7]. However these protocols have often complicated the working of underlying protocols. The ICN approach is different in the sense that it shifts it self away from location oriented dependencies. Though mobility issues has been addressed extensively for IP-based wireless networks, this area is still an open issue in the case of ICNs. This section discusses the mobility solution in some of the ICN designs along with a brief introduction to their architecture.

1.1.1 A Data Oriented (and beyond) Network Architecture (DONA)

In DONA [2], information name is of type P:L, where P is the cryptographic hash of the information publisher's public key and L is the label for the information identification. DONA uses Resolution Handlers (RH) that index information stored in the storage space of the architecture. DONA resembles as a DNS in a way that RHs are structured into a

tree topology. The request from the user or subscriber is first searched in the local RHs. If unsuccessful, the request is queried up in the tree structure until the information is found. In DONA, information publishers or creators follow the early-binding approach by registering the identifier-locator mappings that must be resolved before deliveries can be performed (Figure. 1.1).

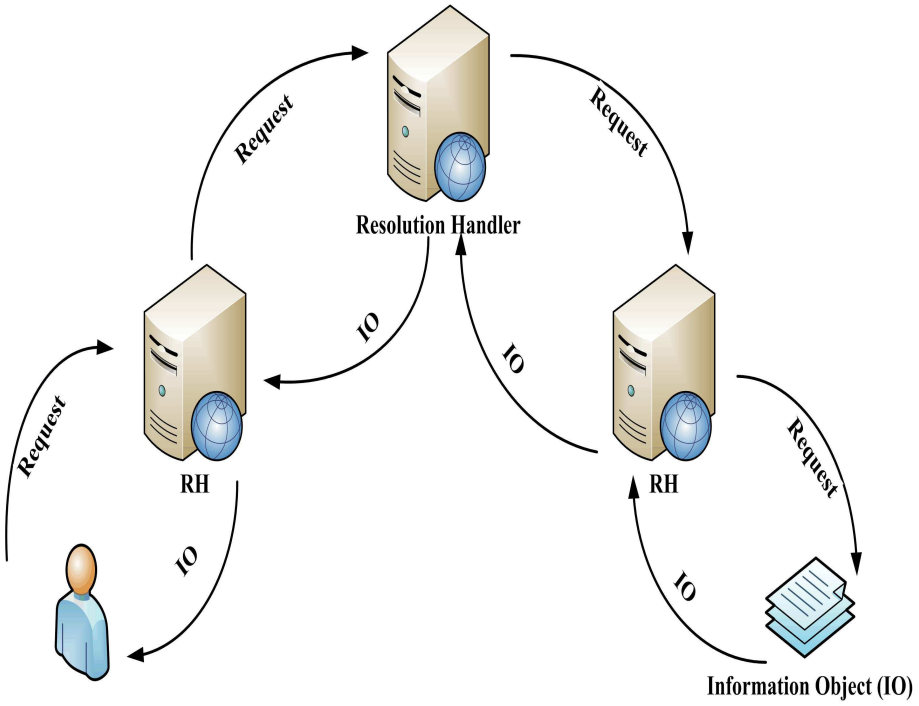


Figure 1.1: DONA

1.1.1.1 Mobility Management

In order to manage the end user's mobility, DONA simply changes the host's RH to that of the new network to which the end user is now connected to. A new request has to be made to this new RH. This complicates the situation as it requires a connection re-establishment. An open issue is how to deal with the delay and overhead that is involved in relocating and re-establishing the new connection. A possible solution consist in using cached copies of the content (if found on the route) when re-establishing the new link.

1.1.2 Network of Information (NetInf)

NetInf [3] is another example of early-binding in which IOs are published and registered in the Name Resolution Service (NRS) along with the locators. A NetInf user when retrieves an IO by sending a request, NRS resolves the request into a set of locators. The locators are used to retrieve the IO copy form the best available source. Mutilevel DHT (MDHT) [8], in NetInf, provide a global lookup system. Users' search is responded with the list of potential sources and the optimal source is chosen to retrieve the information using any supporting transport protocol (Figure. 1.2).

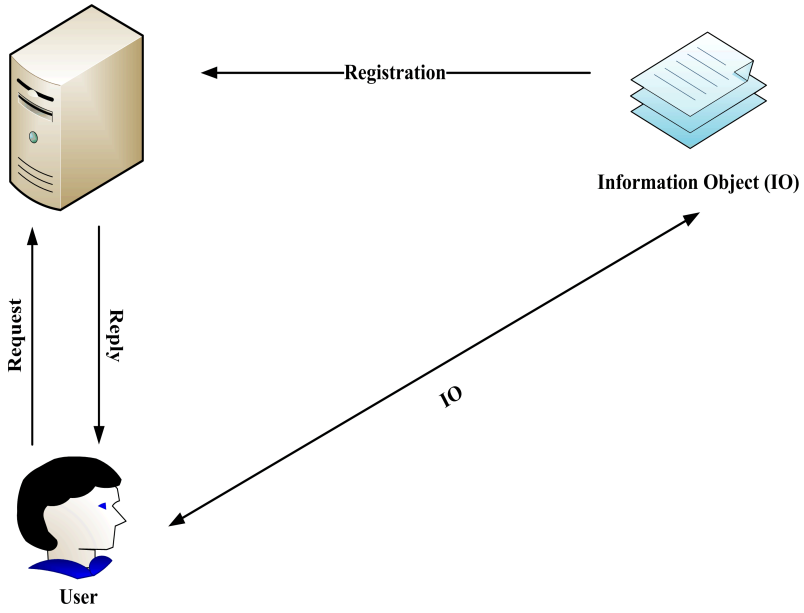


Figure 1.2: NetInf

1.1.2.1 Mobility Management

Mobility support in NetInf is provided by means of MDHT and Late Locator Contruction (LLC) [9] methods. In the case of MDHT, in user controlled mode, after a mobility event, the response to the user's request is provided with the list of locators of the potential information sources. The optimal source is selected among the received list. In the case of MDHT controlled mode, a single source (locator) is provided to the user. Under this condition, the high probability is that the user has to request the NRS again if the provided source is not an optimal choice . In LLC case, which is the extension of MDHT for edge networks, special registers called Attachment Registers (ARs) within the core network update the identifiers whenever a mobility event occurs in the edge network (details of the

procedure are given in Chap.2).

1.1.3 Content Centric Networking (CCN)

In CCN [1], the content request is sent by issuing an Interest (request) packet. The main idea in CCN is that this request for an IO is routed towards the location in the network where that IO was published. In other words, it is routed to the closest instant of the content. As the Interest packet traverses the route, the cache memories of intermediate nodes are looked up for the possible availability of the required IO. If available, the data packet is routed back to the requester through the same path. All the intermediate nodes cache the copy of the requested IO. Unlike DONA and NetInf, CCN follows late-binding in which the content request is resolved to a specific location (the node where the IO is discovered). As shown in Figure. 1.3, User 2's Interest packet is acknowledged at the first hop (as a result of caching of the same IO at the intermediate nodes when User 1 requested it for the first time).

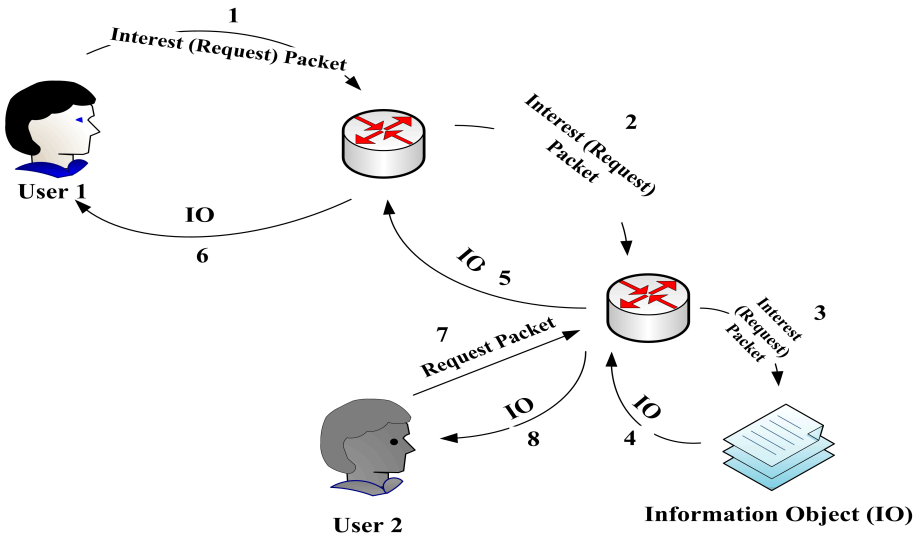


Figure 1.3: CCN

1.1.3.1 Mobility Management

User mobility in CCN is inherent and intrinsic. A node when switching its connection in the network can resend the interest packets which were not acknowledged by the data packets. This can happen seamlessly as no new registration is required in this process.

1.1.4 Publish Subscribe for Internet Routing Paradigm (PSIRP)

In PSIRP [4], the publisher publishes the content or IO in the network to which the user or subscriber subscribes to. To subscribe to an IO in the network, users use Application Identifiers (AI) which are human readable identifiers. AIs are mapped to Rendezvous Identifiers (RI) which are used by rendezvous systems to generate Forwarding Identifiers (FI) (locators) that are used for routing IO packets within the network as shown in Figure. 1.4.

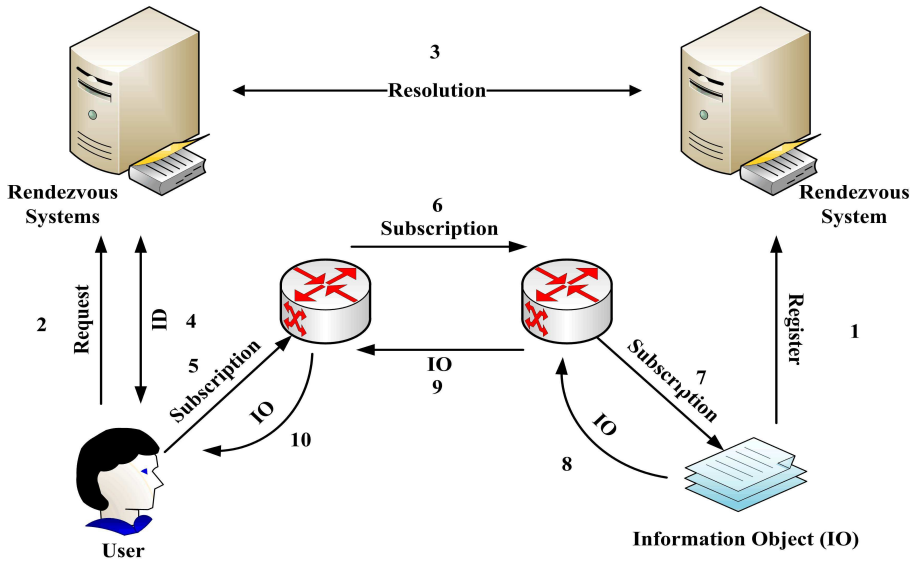


Figure 1.4: PSIRP

1.1.4.1 Mobility Management

The subscriber or user mobility in PSIRP is not a big issue and can be achieved in a straight forward way. A user when moving from one network attachment point to another, is simply re-subscribed to the content which it was accessing at its previous location. The whole process is repeated to generate the FI for that content. The mobility process is termed efficient if this re-subscription process or the generation of FI is completed at the earliest. In [10], PSIRP is claimed to be more efficient than Mobile IPv6.

1.2 Open Issues

The Future Internet architecture is a recent domain of research that has addressed multiple issues which the current Internet architecture is facing. The *Information Centric* approach is not the only key research issues based on which different projects were initiated to design a clean-slate Future Internet architecture. A list of different projects related to the Future Internet design is given in Table 1.1. Key research issues based on which these projects are initiated are listed. Though, the variety of these issues have touched various aspects for the design of an architecture that can replace the current Internet architecture. However, some important points need attentions and should be addressed. These points reflect our perspective towards ICN-based network architecture and highlight the issues addressed in this thesis.

Country	Future Internet Project	Key Research Issue
USA	FIA [11](NDN [12], Mobility-First [13], NEBULA [14], XIA [15], etc), FIND [16] (CABO, DAMS, MAESTRO, NetSerV, etc), GENI [17] (Over 100 different Projects)	Named Data Networking (NDN),Mobility Management, Cloud Computing Architectures, Test Beds
Europe	4WARD [3], PSIRP [4], FIRE [18], P2P-Next, MASTER, INSPIRE, etc	Future Architecture, Internet of Things, Information Centric Network, Network Virtualization
ASIA	NWGN,AKARI [19], JGN2 [20], JGN-X, CNGI, CERNET [21], CERNET2 [22], 3T-NET, etc.	New/Next Generation Internet, Next Generation Broadcasting, Test beds,etc.

Table 1.1: Some Future Internet Projects

(i) The first issue is the notion of clean-slate design for the Future Internet Architecture. The Internet architecture in use today has evolved with time. Designed initially for non-commercial use, it was not anticipated that this node-based network will undergo through revolutionary changes in the coming years. The idea of developing a clean-slate Future Internet architecture replacing the current Internet architecture requires some additional assumptions. The revolutionary ideas last long only if they can sustain

the evolutionary changes. This is one of the main reason behind the success and survival of the TCP/IP-based Internet architecture. Future Internet architectures can be considered as a new breed of networks in which there is a major shift (evolution) in the policies that are adopted for inter-networking and unlike the current Internet architecture, the approach for inter-networking is information centric and not node centric. However, while designing algorithms and protocols for different functions such as routing, mobility management and others, it should be made sure that they are sound and scalable to sustain both major as well as minor (but incremental) changes in the network architecture. Therefore, the new architecture for the Internet should have embedded feature of being flexible towards any changes.

(ii) The second issue of the utmost importance is the backward compatibility of these architectures with the current Internet architecture. It is believed that these architectures are proposed bearing in mind this reality. Nodes and infrastructures today should be able to communicate with the entities of the new architecture. This requires a common communication space available to the end user to access the services of both architectures. A cross architecture support is envisaged in such case. Many challenging issues may arise while adopting this approach such as mobility support for mobile nodes when re-locating itself between old and new architecture. How this process will be facilitated through cross architecture support? In [23], content-centric features of the current Internet architecture are discussed and explain how today's IP- protocol suite can be compatible with Content-Centric Networks.

(iii) The third issue is the deployment of a new architecture. The implementation should and will be started at a small scale. The success of a new technology involves a variety of factors ranging from end user's interest to its market value. The proposal of incentives for users adopting this new technology can be beneficial. This requires investment at initial stages from the business point of view. The success and acceptance of this new architecture can be judged by measuring the payoff. The increase payoff will result in large-scale deployment of this new architecture which will eventually reduce the overall infrastructure cost (less than the initial stages) followed by incremental removal of the old architecture.

1.3 Problem Statement

In this thesis, we have focused on the mobility management issue in *Information Centric Networks* in general and in *Network of Information* (NetInf) [3] in particular. This issue is addressed in the context of a cross-architecture support between the IP-based Internet and NetInf for wireless networks. As discussed above, for better performance and adaptation to a new Internet architecture, legacy protocols and nodes should be able to communicate over it. It means that new network entities should be designed in a way that makes them compatible with NetInf and non-NetInf (IP-supported) sites. Similarly, new algorithms and protocols should also be designed to support different functions. However, we have focused on mobility management in particular within wireless environment. This means that during mobility, the QoS of an ongoing communication session should not deteriorate. The handover situation, when a node switches its connection from one network to another (vertical handover), is the most critical moment and demands seamless mobility. Algorithms or protocols designed for seamless mobility during vertical handover should have the following characteristics:

- For ongoing sessions, the handover latency should be minimum. This means that the time taken for route diversion of the data traffic from the old access point (node's old point of attachment) towards the new access point (node's new attachment point) should be minimum.
- The data loss should be minimum during handover.
- The probability of handover failure should be zero or minimum.

The handover management can either be network controlled or mobile controlled. However, the performance of a handover algorithm can be increased if both players, i.e network and mobile mutually control this phenomena. This requires some modification or addition to the networks. For example, the Proxy Mobile IPv6 [24] protocol requires an additional infrastructure for its working, whereas Fast Mobile IPv6 (FMIPv6) [25] requires mobile node modifications. Partially controlled handover process means partial assistance from one entity in the wireless network to the other and vice versa. In the current perspective, modification in mobile nodes and network architectures can be helpful. The contribution in this thesis, discussed in the next section, is based on the issues discussed so far. Architectural modifications in network entities can be beneficial during mobile node handover. The major contribution in this thesis, suggests to adopt these modifications.

1.4 Thesis Contribution

The first contribution is reported in Chap.3. As discussed in Sec.1.3, modifications in the architecture of network entities can be useful in order to support mobility in wireless networks. This chapter proposes a *Network of Information Mobile Node (NetInf MN)* architecture that defines a cross-architecture design that is capable of working both in the current Internet and the proposed *Information Centric Network* called NetInf. The Central Control Unit (CCU) is the additional infrastructure introduced in the wireless network. In NetInf MN, the Virtual Node Layer (VNL) is introduced with its working modules that provide handover, data relaying and power management supports. The Inner Locator Construction Routing (ILCTR) and the Outer Locator Construction Tunnel Routing (OLCTR) support communication between NetInf and non-NetInf sites. The working units of CCU are Mobility Pattern Data Base Unit, Mobility Prediction Unit, Mobility Zone Allocation Unit and VNL Coordination Unit. Theses units provide support in different situations discussed in deeper details in Chap.3. The proposed VNL algorithm explains how different modules of NetInf Mobile Node and different units of CCU work together to support mobility in a wireless network. The handover scenario explains the working principle of NetInf MN implementing the VNL algorithm in a wireless network.

The cooperative diversity in the network among mobile nodes can maximize the overall network coverage. However, the cooperation among mobile nodes to facilitate each other during mobility events (especially during handover) is difficult to achieve because of their selfish behaviour. This issue is solved in our second contribution presented in Chap.4. The methodology adopted to solve this issue involves mathematical modelling. We have selected Game Theory [26] to follow this approach. The *Stackelberg Leadership Model* [27] is used to formulate the problem of our system model. Players in the game are selected from a wireless network scenario where the network (an access point) acts as a leader and advertise incentives for cooperating nodes, whereas mobile nodes in the network act as followers. Each node in the network wants to maximize its utility or payoff function but may or may not cooperate with other mobile nodes in the network. The cooperation results into rewards (incentive from the network leader). The leader of the game aims to maximize its revenue. The problem here is the selection of a strategy by mobile nodes in the network that can maximize its utility. To solve this problem, we have used the Reinforcement Learning (RL) scheme known as *Combined Distributed Payoff and Strategy Reinforcement Learning* (CODIPAS-RL) [28]. Using this scheme, mobile nodes experiment different strategies with different sets of actions and receive payoffs. The strategy returning high payoff values is repeated with higher probability. The numerical results of our proposed method show that the adopted strategy achieves the Nash Equilibrium [29]. Similarly, for

the 2-level Stackelberg Model, the Stackelberg Equilibrium is achieved at a point where the network maximizes its overall revenue.

The final contribution reported in Chap.5 is a NetInf-architecture use case example. We have proposed a NetInf Email application or service that is based on the NetInf architecture. Asymmetric key cryptography is used for user's email ID. Using the NetInf architecture platform, this service does not require dedicated servers or ports unlike the current email service. We have presented the NetInf email architecture framework that has three dedicated components, namely, the Storage space, the Index space and the NetInf API for end users to access this service. The message format describes how different parts of the message are considered as separate objects and are assigned with unique IDs when published in NetInf. The NetInf email sending and receiving procedures explain the working principle of this service. The qualitative evaluation of this service highlights different positive and advantageous characteristics of this application.

1.5 Thesis Organization

This document is structured into six chapters. The following chapter is dedicated to the state of the art explaining the NetInf architecture and its mobility management proposals. It examines different mobility protocols in IP protocol stack. Chap.3 presents the NetInf MN architecture in detail and the VNL algorithm for mobility management in wireless networks. Chap.4 proposes the game theoretical approach towards handover management and data relaying in wireless networks. The formulated mathematical model is evaluated in this chapter. The last contribution is related to the NetInf Email service presented in Chap.5. We conclude this thesis and give possible directions for future research in Chap.6.

Chapter 2

State of the Art

Contents

The contemporary Internet is a four decade old architecture, proliferating during all these years and has undergone through various changes along with the continuous development of wide range of applications over it. It has been noticed recently that, this same infrastructure is now primarily used for content distribution [1], [30]. The paradigm shift from being a node-centric Internet architecture to an information dissemination platform is the basic influencing factor for research community to think about the basic architectural redesign of this network. The response in this regard is the proposal of an *Information-centric Future Internet architecture*. This means that, unlike the existing host-to-host routing infrastructure, a content-based system should be put into place. Hence, applications that are involved in generating content requests, are routed using content identifiers. In other words, information have become more and more important in communication and networking. This is evident from the fact that most of the traffic today on the Internet is related to content distribution. This content distribution today does not involve end-to-end data exchange, but rather fragmented into smaller chunks that are named as information objects. They can be accessed in a variety of ways such as in peer-to-peer (P2P) networks where a node can act both as a client or a server. However, accessing named data objects at the center of the network is unique. This *Information Centric* approach has made things challenging for the current Internet architecture and has made *Information Centric Networking* an important research domain.

As discussed above, in *Information Centric Networking* (ICN), accessing information is different from client-server model based Internet. In the current Internet architecture, the (audio/visual) content demand is fulfilled using a P2P overlay network that shares the core network payload. Whereas in ICN, information are considered explicitly as a *first class entity*. ICN is a concept for the Future Internet architecture that has been addressed by a number of projects world wide like DONA [2], 4WARD [3], ,PSIRP [4], NDN [12], PURSUIT [31], SAIL [32], etc. These projects have proposed different architectures but the corner stone for each architecture is to have information as the centric network entity. The aim is to make content distribution fast and reliable in case of disruptions during communication. Instead of using P2P or Content Distribution Networks (CDN) explicitly for content distribution, the network architecture should natively supports such mechanisms. For example, a request made by a user, seeking some information, is delivered in the same manner as in P2P networks from the nearest possible location rather than from an end host. Caching plays an important role in terms of information dissemination in ICN. In case of mobile devices, mobility support becomes easy in ICN as a result of ubiquitous availability of information. In the present context, mobility is the physical displacement of a mobile device, relocating its point of attachment within the network. During an active communication, mobility can seriously effect QoS from the network's point of view and end

user's quality of experience.

In this chapter, we present an overview of the solutions proposed to manage mobility in ICNs. The Network of Information (NetInf), work package of 4WARD, is the focus of our discussion here. The discussion also includes comprehensive overview of mobility management solutions in general and handover solutions in particular across all the layers of the TCP/IP-protocol suite for mobile nodes in wireless networks.

2.1 Network of Information

The *Network of Information* (NetInf) is one of the work packages of the 4WARD project that deals with the development of the ICN-based Internet architecture. In NetInf, the device centric concept is replaced by the *Information Centric Concept* where Information is given the highest priority. Users, regardless of the information location, interact with NetInf and retrieve this information as shown in the Figure.2.1. The information is hierarchically divided into different levels. The highest level represents the information abstraction based on the publish/subscribe paradigm or in other words represents the semantic details of the information. The prime objective of NetInf is to design a new networking approach with new protocols and algorithms for routing, new name resolution schemes, security and privacy framework and innovative mobility management ideas.

However, before going into details, we first briefly explain the motivation that manifested the need of an information-centric Future-Internet architecture called *Network of Information* (NetInf). As said earlier, for efficient content distribution in the network, a new Internet architecture was needed. NetInf architecture is one of such efforts in the ICN domain to fill this void.

2.1.1 Efficient Content-Distribution Issue

The increasing demand for data distribution and replication resulted into a massive use of P2P overlays (e.g., Akamai [33], BitTorrent [34], etc). P2P overlays handle this overwhelming demand of data by sharing core network load and distributing it among dedicated peers (node dedicated for file sharing). These peers are self-organized and data is retrieved through these (available) peers in the form of fragments (chunks). Similarly, Content Distribution Network (CDN), more or less, works in a similar way, but is more transparent to the end users. The main objective of these two approaches represents a deviation towards a network which is content or information based. In mobile networks, the Internet data demand will be doubled by 2014. Most of this demand is for video content. Wireless network domains like 3GPP-LTE with increased bandwidth demand in the near future require less load on core network. The obvious solution is the use of P2P/CDN-like services. The

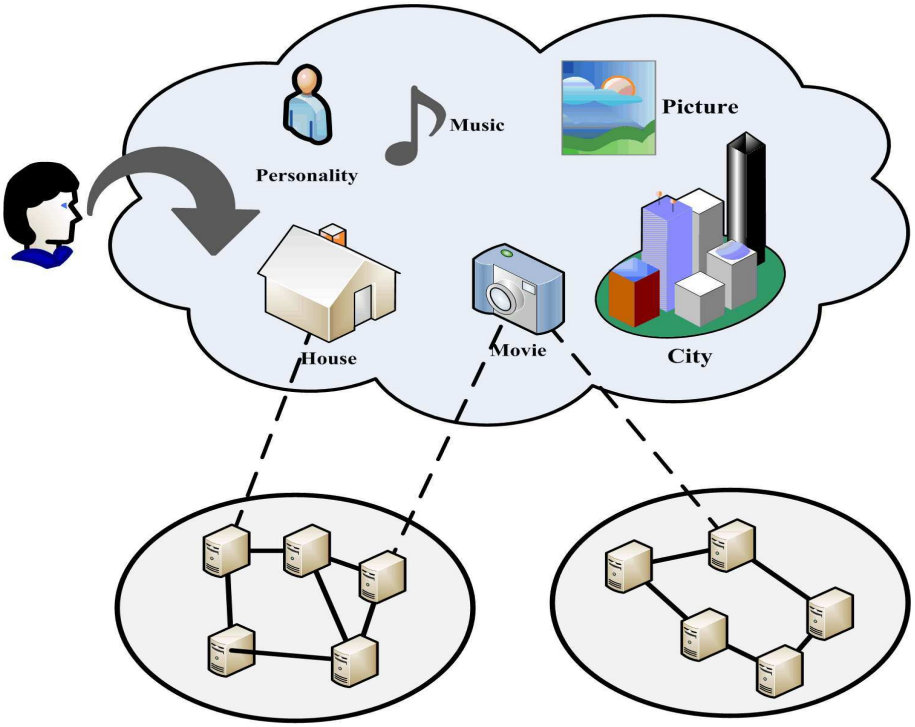


Figure 2.1: Network of Information.

question is: can there be a network architecture that distributes contents in a scalable and efficient way as per users demand anywhere and at any time?

2.1.2 Persistent Name and Unique Identifier Issue

Designing and implementing a naming scheme for an information-centred network is a major issue. Since, in ICN, accessing information by its name means that regardless its location anywhere in the network, the name or ID of the information remains the same. Applying this information centric concept in the current Internet architecture where URIs (Uniform Resource Identifiers) are used to access information is challenging because object's (information/content) URI also works as a locator. Any location change of the host hosting certain information or content (video/text/audio) changes its URI. A user accessing the same information next time experiences an error message such as "Host not Found". Even if replicas of the same information are cached at different locations in the same network, different URIs are required for different locations. This leads towards the issue where an information object cannot have a unique identifier. In general, the current Internet uses URIs as locators and identifiers at the same time which makes persistent naming of the data

impossible. This results in the disruption of an ongoing communication during mobility as relocating the information once again takes time whenever the connection is re-established. The question here is; how a content- centric network can solve the persistent name and unique identifier issue?

2.1.3 Content Privacy and Security Issue

The content security is another major issue for ICN. Today, the technologies used to secure the retrieval of information provide end-to-end connection authentication instead of information authentication. Popular security methods are Transport Layer Security (TLS) and Secure Socket Layer (SSL). Both are cryptographic protocols that provide security over the Internet. This means that the extracted information is authentic if the connection was secure. The object authenticity is not verified even if it is a basic concern. In the case of ICN, object authentication is more important than the end-to-end connection authentication. How this issue can be addressed in ICN paradigm?

2.1.4 Intermittent Connectivity

It is obvious that environments where all-time connectivity is not possible are challenging because of the frequent disruptions and disconnections. In such challenged environments, the end-to-end connection is difficult to establish. The challenging reasons are high-speed mobility of nodes, the sparse connectivity and disruptions because of environmental issues. If the possibility of having an end-to-end connection is minimum or almost zero, DTN [35] like store-and-forward approach can be adopted. A simple example can be considered of passengers traveling in a train and accessing videos or other information through their mobile devices. They experience frequent disruption whenever they move out of the range of their respective base station or access point. An IP-based network is not an efficient way to continue the communication whenever there are chances of connectivity. In case of ICN, how this problem can be solved?

2.1.5 Mobility and Multihoming

In ICN, the system tries to access information from the end-point that does not require a transport layer session which is bounded to a specific host. IP-based networks handle mobility by routing and indirection. Mobile IP is the famous and practical solution for all such networks. Multihoming is a technique to increase the reliability of the Internet connection at the user end through simultaneous attachments to multiple networks. In the current Internet architecture, having IP-address is problematic for named data object as during mobility, Mobile IP requires indirection and transport layer session re-initiation

to continue the ongoing communication. The required data object located at the nearest cached location is thus not accessible in such situation. In the case of an ICN, without any transport layer connection, the information can be retrieved from the nearest topological cache. Similarly for multihoming, TCP connections are bounded to locators which make it difficult to employ all possible access opportunities. How an ICN addresses mobility and multihoming issues?

The above issues can be considered as a set of problem statements that NetInf has identified and addressed. Details of the proposed solutions can be found in [3]. The following sections, briefly discusses them. The mobility management in NetInf is explained in deeper details as our work is related to this particular domain of ICN.

2.2 Network of Information Architecture

The following components, explained briefly one by one, constitute the NetInf architecture. It should be noted that, more or less, almost all the proposed ICN architectures have similar sets of architectural components. They work together and provide support to perform various tasks.

2.2.1 Information Model for NetInf

In *Network of Information*, *data* and *information* have different meanings. The classification is based on the *Information model* [36] proposed for the NetInf architecture. The concept of *Information* is the generic definition for any entity; real or virtual. Real-world objects include all physical objects such as a famous city (Paris) or a monument (Eiffel Tower), a building or a personality as shown in Figure. 2.1. Virtual information include text files, videos, pictures, etc. This *Information abstraction* is digitally represented as *Data* composed of streams of bits and bytes. In the NetInf Information Model, information are represented as an abstraction at the highest level showing semantic details of the object (physical/virtual), whereas the data abstraction is the set of bits representing the actual payload. Hence, in general, each entity or object, physical or virtual, is considered as an object and based on the above discussion, objects in the NetInf Information Model are classified as *Information Objects* (IO) and *Bit-level Objects* (BO).

Figure. 2.2 shows a typical hierarchical representation of an IO. The semantic description of an IO can have different meanings and is replicated and disseminated ubiquitously at various locations within the network. The semantic search helps to accessing the desired information, whereas the bit-level access is possible once the optimal location of BO becomes available.

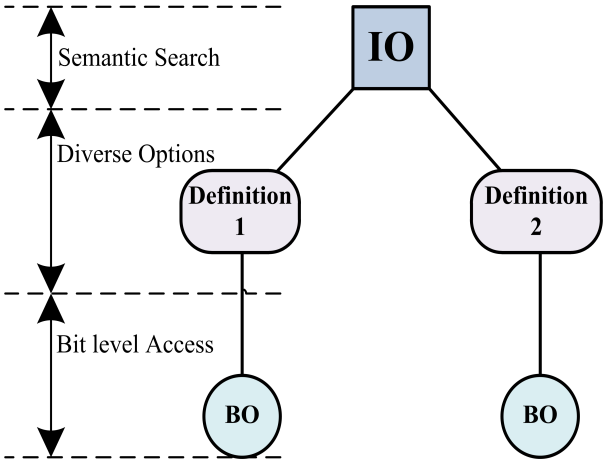


Figure 2.2: Information and Bit Level Objects Hierarchical Representation.

2.2.2 Object Naming Architecture

In the today's Internet architecture, URLs are used to define the location where information is stored and are resolved to identifiers through a name resolution process. In NetInf, IO are named having location-independent identifiers unlike URL. However, the question that first comes into our minds is: which entities should be named? Data objects (or BOs) (as mentioned earlier) representing virtual contents like files, images, etc. is one category of objects to be named. An IO, though, represents the semantic description of the data, can also be named. For example, in the case of naming real-world physical objects, their virtual representation (in the form of an image on the map with GPS coordinates) can be named by the NetInf naming scheme. The naming architecture sets some rules and goals for defining the object name in NetInf architecture. These rules are as follows:

(a) As mentioned in the Sec 2.1.3, data in ICN must be secured unlike in the current Internet architecture where end-to-end secure connections are taken into account. The information authentication in the today's Internet is not the priority. In NetInf, security is not bounded to the locations where the information is stored. The naming architecture provides embedded security within the IO structure.

(b) The object identifier in NetInf is unique and location independent and follow the locator/Identifier split concept.

(c) Objects have the self-certification property. It means that without relying upon any third party (infrastructure for verifying integrity), the object is capable of self verification and proving its integrity.

(d) The confidentiality of information in NetInf means that only eligible users are allowed to access. This is made sure by using cryptographic schemes.

(e) The change or modification of information does not affect the persistence of names in NetInf. Persistence means that the object's name or identifier remains valid and unique even in cases where it changes its location in the network.

(f) Once an information is published, it is available to all authorized entities in the network. The access to information to an authorized user group is controlled by defining access rights (e.g., read, write, delete) by the owner of the information.

2.2.3 Integrated Name Resolution and Routing Approach

Being a name-based network architecture, name resolution and routing service for NetInf IO are provided through an integrated approach. The NetInf architecture has adopted a name based routing scheme for routing which is integrated with the name resolution process. The two proposed schemes for this integrated approach are Multiple Distributed Hash Table (MDHT) [8] and Late Locator Construction (LLC) [9] for core and edge (access) networks respectively. MDHT handles name resolution and routing in the core network while LLC works in access networks at the edge of the core domain. A detailed discussion on these two methods is presented in Sec. 2.3.

2.2.4 Data Caching and Storage

In NetInf, data caching at multiple sites (caching at network edge and in-network caching) in the network optimizes the overall performance. Intermediate nodes, that relay data retrieval requests towards their source, cache most of the data objects requested on its way towards the destination. Repeating the same request for the same data object does not need to be routed again towards the same source location as it can be retrieved from the nearest site that has previously cached the same information. Information cache is a temporary information storage. On the other hand, persistent data storage for NetInf can become a part of the NetInf architecture. Two approaches that can be adopted for this purpose are: **(a)** either storage resources are installed within the infrastructure, or **(b)** network nodes can dedicate portions of their storage memory whenever connected to the service provider network.

2.2.5 NetInf Application Programming Interface (API)

The NetInf API is based on REpresentational State Transfer (REST) [125] style of software architecture for distributed systems such as the World Wide Web (WWW). The API

adopts publish/subscribe paradigm in which an information is published by its creator (owner/publisher) and the user (subscriber) can retrieve that information using different sets of functions. Using these functions, users request for IOs using their IDs. For example, function **Publish** is used to publish an IO whereas function **Revoke** is used for revoking an IO.

2.3 Mobility Management in the Network of Information

In the NetInf architecture, mobility management is embedded or in other words natively provided. A (Distributed Hash Table) DHT-based integrated name resolution and routing structure is proposed for this architecture, called *Dictionary* or *Dictionary Node* [8]. Dictionary Nodes are hierarchically arranged at multiple levels of the NetInf architecture. This overall arrangement forms a Multiple Distributed Hierarchical Table (MDHT) named resolution system. The hierarchical arrangement of DHT-based Dictionary nodes provide support for ubiquitous data dissemination, routing and locality resolution (both at local and global levels of the network). In other words, a Dictionary is like a DHT based DNS for name resolution of the information requested (IO in the case of NetInf). Together, the group of these dictionary nodes form a global system of routing and name resolution for the Network of Information architecture. The scope of the resolution can be local or global (depending upon the availability and location of the required IO).

2.3.1 MDHT

MDHT, as mentioned above, is the hierarchical combination of Dictionary nodes at multiple levels of a network. The structure of this name resolution system is shown in Figure. 2.3 [3].

The multiple DHT-based name resolution system can be of different sizes with respect to the size of a network. As shown, at the lowest level of the system are the end user equipments (nodes). Each node is equipped with a DHT-based Dictionary Node. In case of MANETS and VANETS (or even in case of structured networks), these nodes can become information storage sites as well. One of the prime goals of the NetInf is information dissemination ubiquitously and each node within the NetInf architecture can be enabled to cache and store information. The other levels more or less have the similar implementation of DHTs at their levels.

It should be noted that at the lowest level (The User Equipment Level), Dictionary Nodes are less in number. Distributed in nature, these DHT-based Dictionary nodes have low churn rates. Thus, efficient DHT algorithms can be used for routing, with less number of hop counts.

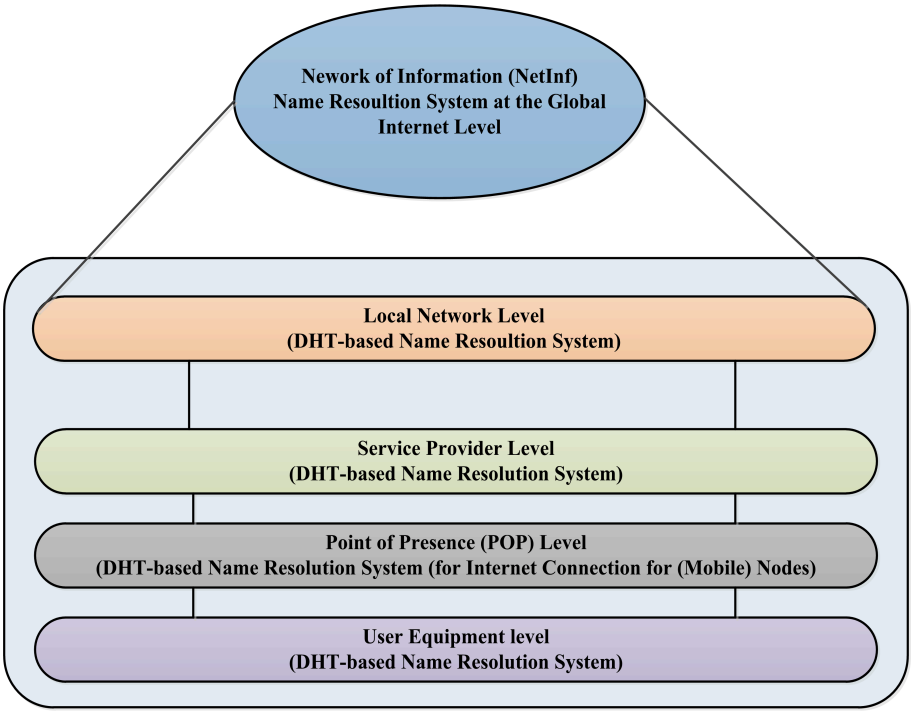


Figure 2.3: MDHT Name Resolution System for NetInf.

2.3.2 LLC

In order to provide mobility to edge and access networks and their entities (mobile nodes), an extension of MDHT is used. An edge network can be distinguished from the core network by frequent changes in its topology as a result of mobility events. The proposed solution designed for edge network topologies is known as the Late Locator Construction (LLC) [9] scheme. In LLC, the name resolution system resolves the name into a locator on demand whereas the resolved locator is the hierarchical representation, describing the path from core network to the host node. This locator is constructed at the latest possible stage in order to support the dynamic nature of an edge network. In other words, the constructed locator describes the most recent topology of an edge network.

The working mechanism of LLC is shown in Figure. 2.4. Attachment Registers (ARs) are assigned by the LLC mechanism for every object, network, router, etc. They are located in the core network and on the sites where Edge Routers (ERs) are deployed. Each AR stores its own ID as well as its neighbouring entities. An AR of a network entity also stores the locator of the neighbouring AR during the time of locator construction. The locator construction mechanism is shown in the Figure. 2.4 and is briefly explained by the following

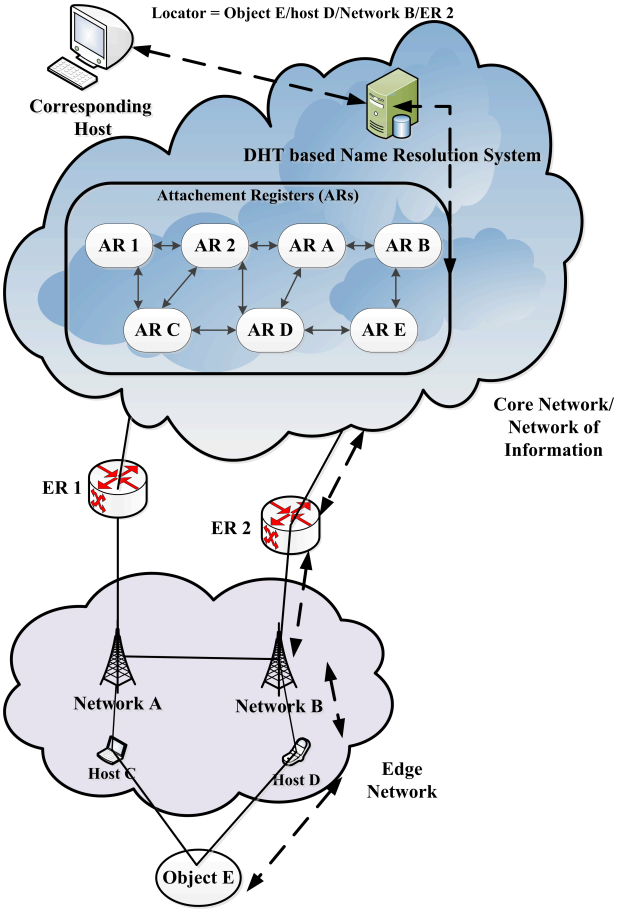


Figure 2.4: Late Locator Construction Scheme in NetInf.

steps:

(i) During the first step, a corresponding host (CH) looks for an object (E) and requests the name resolution system in the core networks for its locator.

(ii) The DHT-based name resolution system queries AR for object (E). Once found, the binding information available in the neighbouring ARs are used to build the locator of the object. The built locator for object E is: **Obj E loc** = {**Object E/Host D/Network B/ER 2**}.

(iii) The selected path is based on different factors. The most important factor is to find the shortest possible path. Though, Host C also has the same information but the path via ER 2, Network B and Host D is the shortest.

(iv) Hence, once CH receives locator {**Obj E loc**} for object E, it is used in the packet

header of a data packet before transmission.

LLC uses the LOC/ID split concept such as presented in [7] and [37]. In order to reduce the routing state signalling within the core network, entities in edge networks limit the information exchange of routing state to their neighbours. The core network routing state is taken cared by the LLC system. The performance evaluation of LLC in heterogeneous cooperative multi-access wireless metropolitan networks (WMAN) [38] shows the advantages of the ICN approach. They have shown high connectivity and less handover latency in a cooperative multi-access environment where WiMax and 3GPP wireless technologies have been used alternatively to provide video services to the end users.

To this end, we have described in detail the concept associated with NetInf architecture and provided a brief overview of its different components. For further information, the complete details of these components can be found in [3].

2.4 Mobility Management in All-IP-Based Wireless Networks

This section highlights some mobility management solutions for wireless networks. The focus is on those telecommunication and access networks that have evolved into an IP-based heterogeneous network in which different access technologies transport information (voice, data, text and all other media) by encapsulating them into packets. Such heterogeneous environment, with diverse network access technologies, demands cooperation among wireless networks especially during mobility. Before going into the details of these available solutions, we first define mobility management. Mobility management in a wireless network is classified as:

(a) **Location Management:**
This is the process by which a mobile node updates the network about its current location. The main purpose of this procedure is to make sure that the network should always be aware of the mobile node's presence for better coverage for any sort of communications. Similarly, a network has to update its database on a frequent basis and location update of mobile nodes is an essential procedure in this regard. The location update process is periodic and there are various reasons for mobile nodes to update the network about their current location. In cellular networks, one such case is when a mobile node is switched on and location update of this mobile node is mandatory for any future correspondence within the same or other networks.

(b) **Handover Management:**
In telecommunications, the term handover is a mobility process in which an ongoing communication is maintained during connection relocation of a mobile terminal from

one network to another. In other words, it is the detouring the routing path of the ongoing communication session as a result of a change in the point of attachment of the mobile node. For example, a user with a mobile device, moving randomly and changing its point of attachment within the network while continuing its ongoing session undergoes handover process. The handover management allows a network to assist mobile nodes to maintain their connectivity during relocation of their attachment point in the wireless network.

The handover process basically involves three steps:

(i) The first step is the handover triggering or initiation. There are multiple reasons for initiating handover such as RSS value, data rate requirement or traffic congestion in the network.

(ii) The second step consists in finding a new connection as early as possible to resume the ongoing communication (if there is a session in process during handover).

(iii) The final step is the switching of the data flow from the old network to the new one.

The handover process in mobility management can be intra-system or inter-system. The intra-system handover is a handover for a wireless environment of homogeneous networks while inter-system handover is for the heterogeneous wireless networks. This major difference classifies the handover process into two sub-categories. They are:

(a) **Horizontal Handover:** The horizontal handover is the type of handover where a mobile terminal experiences RSS degradation from its serving access point (or a base station) and switches to another access point (or a base station) using the same wireless technology (e.g., handover within 3G networks or within WLAN/IEEE 802.11). The horizontal handover normally occurs whenever the RSS value goes down beyond a certain threshold value.

(b) **Vertical Handover:** The vertical handover is a handover type where a mobile node changes its data link layer technology used to access the network. In simple words, a mobile node changes its connection type to access a particular network technology. As an example, consider a smart device, capable of accessing both WLAN and cellular networks. WLAN provides high data rate while cellular networks give ubiquitous connectivity to access the Internet. Thus, the smart device switches its data link layer technology to access a network whenever one is available. If both networks are available simultaneously, the terminal will choose the service that satisfies the end user's requirement.

It must be noted that besides above definitions, handover is also classified as *soft handover* and *hard handover*. The soft or hard handover situation is possible both in vertical and horizontal handover.

- The situation where a mobile node connects to a new access point while still roaming in the coverage area of its old access point is a situation termed as *soft handover* or *make-before-break*. The mobile node, while performing a handover can have simultaneous connections to many access points at one time. The mobile node may move back and forth, and can switch to any one of the access points it is connected with. This situation is termed as ping-pong.

- The situation where the previous connection is first cleared and terminated before connecting to a new network is termed as *hard handover* or *break-before-make*. Hard handover situations are not desirable as they can result in data loss as well as handover process failure.

2.4.1 Mobility Management in IP-Based Internet

The interconnection of different wireless network technologies demands for a common infrastructure. Since almost all technologies today provide Internet access to end users, it is evident that a global architecture for future wireless communication can be IP-based. One such approach is called Next-Generation All-IP-Based Wireless Networks (NGN) where all sort of information (voice, video, text messages, etc.) are transported in the form of packets. The mobility management solutions are provided in the IP-based Internet architecture across different layers of the TCP/IP protocol stack. Here, we give a brief overview of the available solutions.

2.4.1.1 Network Layer Mobility Solutions

Network layer, also known as the IP-layer, is the inter-networking layer among different networks. It is obvious that network-to-network level issues are solved over this layer. In the case of mobility management, various protocols have been proposed and adopted at Network layer. However, mobility in the Network Layer can be broadly classified into two categories which are:

(a) **Macro-Mobility:**
The mobility of mobile nodes between different wireless network domains (inter-domain) is considered as macro-mobility, e.g.the movement of mobile nodes from a 3G network domain to a WLAN (IEEE 802.11) network domain is considered as macro-mobility.

(b) **Micro-Mobility:**
In this type of mobility, the movement of mobile nodes is limited to the same wireless network domain. However, the mobility is intra-domain, i.e. from one subnet to another.

The network-layer mobility management protocol performance is bounded to the above classification. Protocols are divided as macro and micro mobility management protocols. The list of mobility management protocols for both categories is long. However, as a use case example, we are going to discuss a few of them.

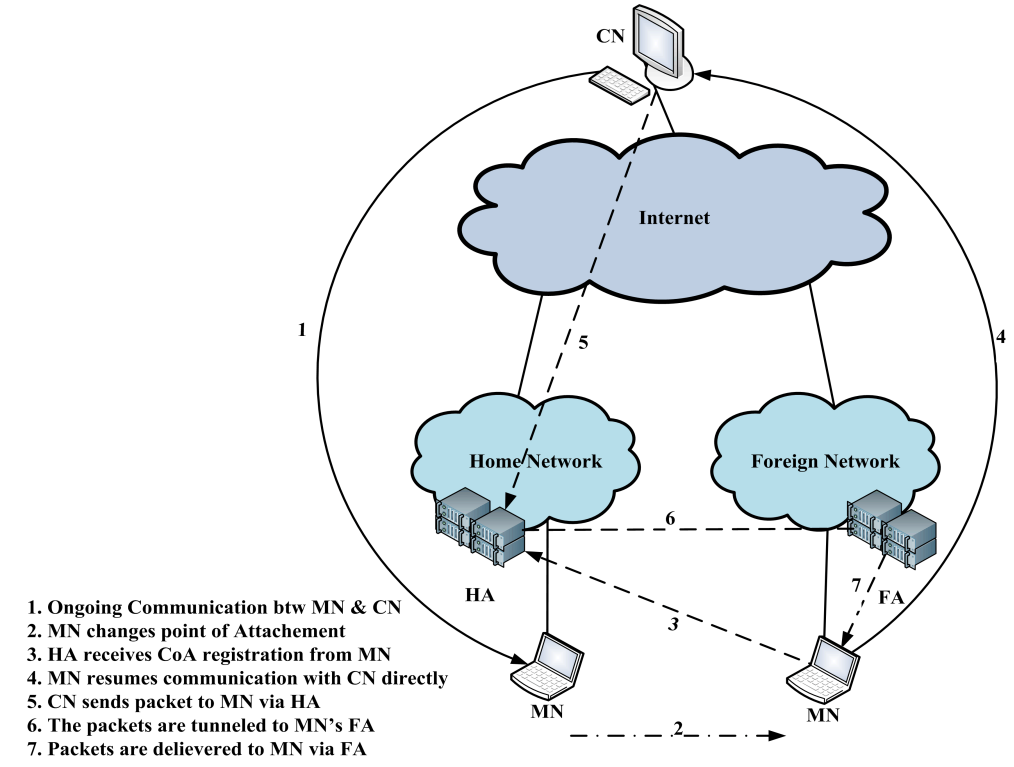


Figure 2.5: MIPv4 Working Principle.

(i) **MIPv4/MIPv6:** Mobile IP is a global mobility management protocol for the Internet architecture. MIPv4 [5] and its extension MIPv6 [6] provides an indirection support during mobility to sustain the end-to-end ongoing communication. In MIPv4, a Home Agent (HA) is an entity which maintains the local IP address of a mobile node whereas Foreign Agent (FA) is an entity that stores information related to visiting nodes. FA also provide Care of Address (CoA) used by mobile nodes visiting foreign domains. Figure. 2.5 shows the working of MIPv4. The problem with MIPv4 is that

once a mobile node is inside a foreign domain, the CN packets are routed via HA. This causes the triangular routing issue. It results into handover delay and can cause packet loss as long as the registration process is not completed at HA.

In MIPv6, as shown in Figure. 2.6, when a mobile node moves from its local network to another domain or subnet, it receives CoA through address configuration techniques as specified in [39] or [40]. The Mobile node registers its CoA at HA and at CN through Binding Update (BU) messages. HA and CN update their record of bindings. Messages can be directly communicated between the mobile node and CN. In other words, MIPv6 optimizes the triangular routing problem of MIPv4.

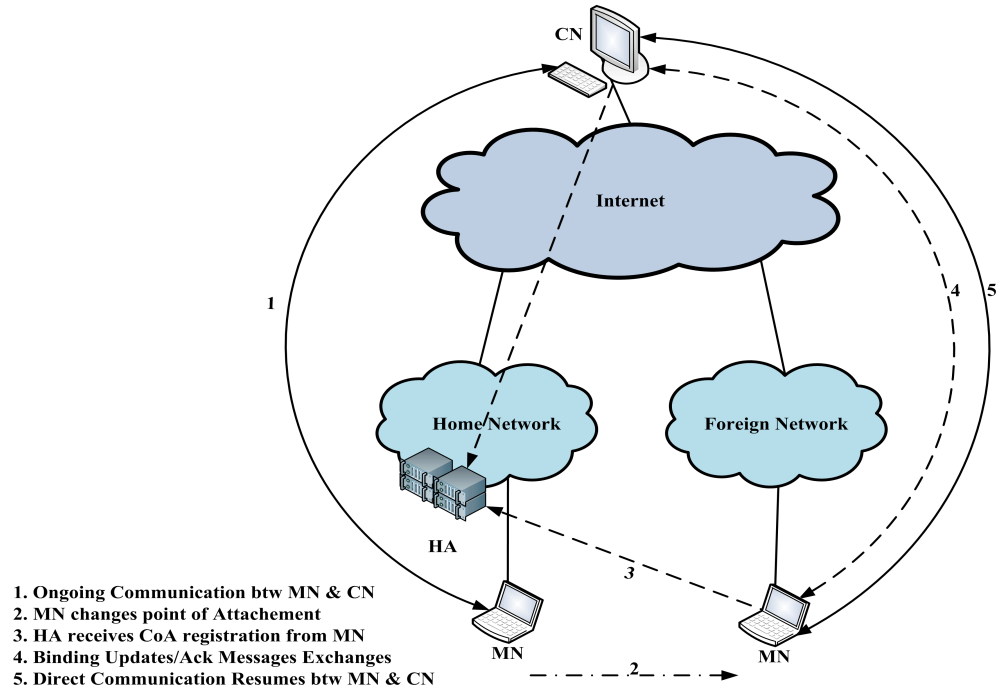


Figure 2.6: MIPv6 Working Principle.

(ii) **FMIPv6:**

The Fast Handover for MIPv6 [25] or FMIPv6 is designed for MIPv6 optimization. The main aim is to reduce the signaling during CoA configuration, binding updates, etc. This is done by acquiring the information beforehand that is required during handover and connecting via a new link. Cooperating access routers are used to get this job done which gathers information from other access routers that can be the possible candidates as the new point of attachment for the mobile node during handover.

So far the solutions discussed are designated for macro-mobility. In the case of MIPv4/MIPv6, the number of messages exchanged during handover increases within a wireless network with the increase in number of mobile users. This can result into signaling overhead putting extra load on the network as well as below average performance. In order to reduce signaling load and handover latency, micro-mobility solutions have been proposed.

The basic motivation behind these protocols is to provide mobility support during intra-domain handover, i.e. between subnets of the same domain. Many Network Layer micro-mobility solutions have been proposed such as Mobile IP regional registration (MIP-RR) [41], Hierarchical Mobile IPv6 (HMIPv6) [42], Proxy Mobile IPv6 (PMIPv6) [24], Intra-domain Mobility management Protocol (IDMP) [43], Cellular IP (CIP) [44] and Handover Aware Wireless Access Internet Infrastructure (HAWAII) [45].

2.4.1.2 Transport Layer Mobility Solutions

The Transport layer in the TCP/IP protocol suite is one of the most important layer as the end-to-end communication is possible because of the famous Transmission Control Protocol (TCP). The limitation for TCP in fixed networks is the network congestion that results into packet loss. However, in wireless networks this loss is increased because of the frequent connection disruptions making TCP an unfavourable choice for wireless communications. In case of mobility, to maintain the TCP connection, transport layer mobility solutions have been proposed such as TCP-Redirection (TCP-R) [46], TCP-Migrate [47], MSOCKS [48], Mobile UDP [49] and Mobile Stream Control Transmission Protocol (MSCTP) [50].

- **MSCTP:**
  As shown in Figure. 2.7, a mobile node commences a SCTP session with a corresponding node. The mobile node exchanges a list of IP addresses with CN. Among all these IP addresses, one is chosen as the primary path on which the current communication is being held while the others are specified as active IP addresses. After sometime, the mobile node starts to move and gets close to a new network and gets a new IP address. It updates CN with an Address Configuration Change message with a new IP address attached. CN adds this new IP address in the list and sends acknowledgement to the mobile node. The primary path of the mobile node may change while moving to the new IP address [51]. The SCTP session is therefore continued without interruption between the mobile node and CN. When the mobile node is in the new network, it informs CN to delete the IP address of the old network.

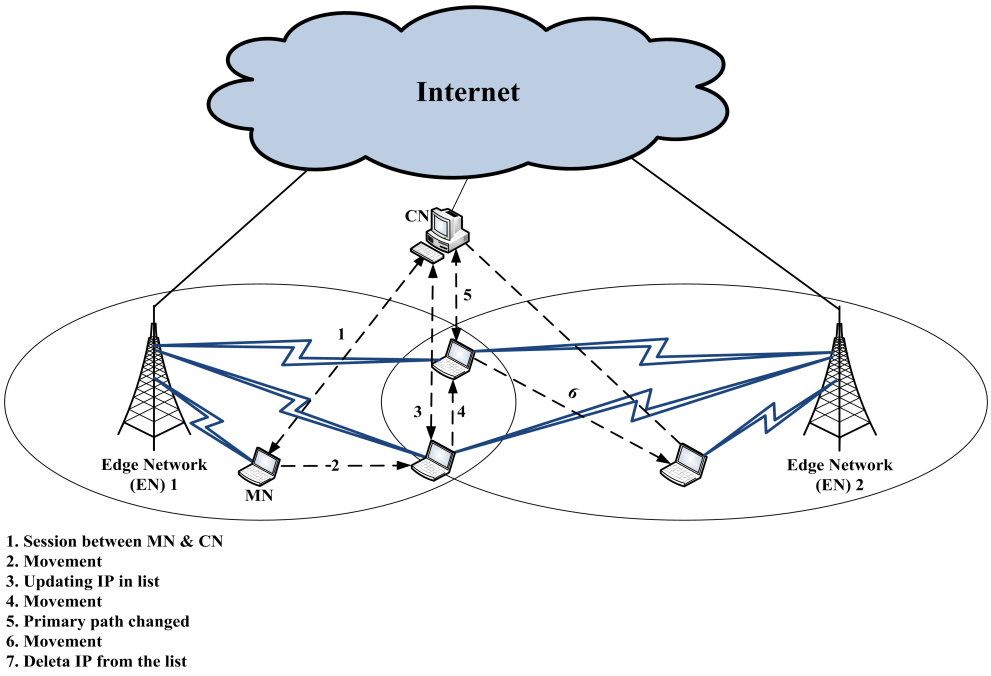


Figure 2.7: MSCTP Working Principle.

2.4.1.3 Application Layer Mobility Solutions

Besides its function that provides methods to make use of transport layer protocols for an end-to-end communication, the application layer can also be used for mobility support in wireless networks. Protocols such as Session Initiation Protocol (SIP) [52], Dynamic DNS (DDNS) [53], IKEv2 Mobility and Multihoming Protocol (MOBIKE) [54] are the examples of application layer protocols designed to support mobility.

- **SIP:** SIP was initially designed for multimedia applications by the IETF as an application layer multimedia signaling protocol. However, it has the potential for mobility management for the Internet. The main entities in SIP are the user agent, the redirect server and the proxy server. A SIP server works both as a redirect and proxy server. A SIP user agent (UA) is a logical network end-point used to create or receive SIP messages and thereby manage a SIP session. The redirect server allows proxy servers to direct SIP session invitations to external domains. A proxy server primarily plays the role of routing, which means that its job is to ensure that a request is sent to another entity closer to the targeted user. SIP is a text-based protocol with a syntax similar to that of HTTP. SIP messages have two different types: requests and

responses. SIP requests are codes for communication and SIP responses complement them. Some of the SIP messages are INVITE, REGISTER, ACK, CANCEL, BYE, OPTIONS, etc. Figure. 2.8 shows the SIP working principle.

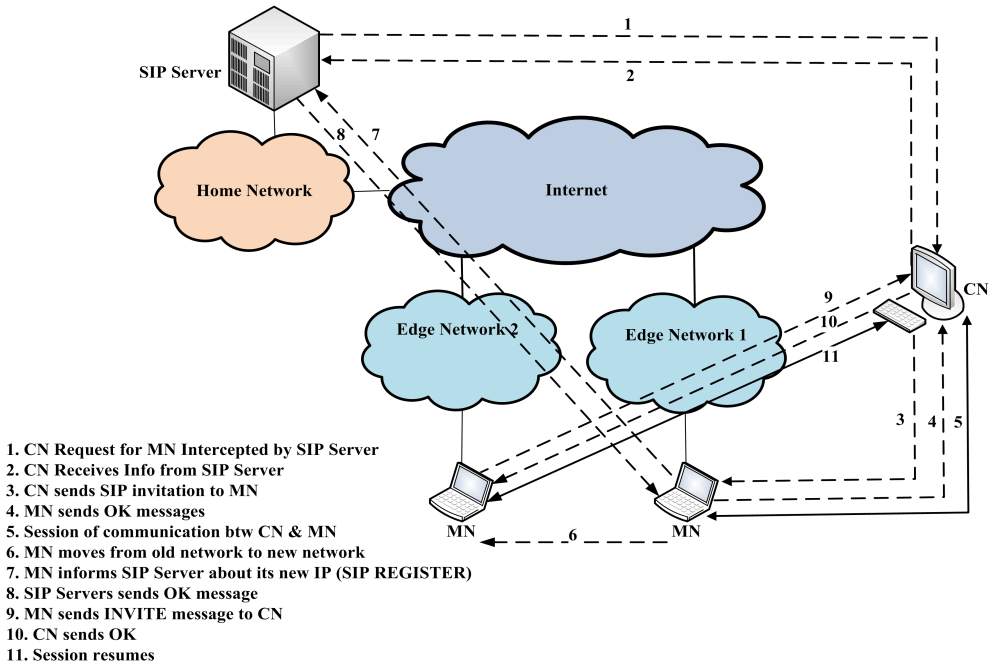


Figure 2.8: SIP Working Principle.

2.4.2 Analysis of Network, Transport and Application Layers Mobility Solutions

At the network layer, MIP based protocols are widely used. For the global Internet mobility management, MIPV6 is the best choice especially when a mobile node moves from one network domain to another. However, for intra-domain movement, from one subnet to another, macro-mobility solutions are inefficient. Micro-mobility issues are slightly different from macro-mobility solutions. Micro-mobility management solution in network layer have been compared on the basis of different criteria [45], [55], [56]. However, more or less, the objective of these solutions is to minimize the increased signaling that is normal in macro-mobility solution. In order to agree to a protocol that provides global as well as local mobility solutions, the idea of combining host-based and network based mobility management, such as in [57] and [58], is a good option.

Transport layer solutions for mobility management, like MSCTP, provide support for

seamless handover and improve services like connection oriented data stream, reliability and flow control. However, there are still some open issues that need attention in MSCTP. For example, by which criteria the primary path is changed and at which stage new IP addresses should be added or deleted in the IP addresses list during handover. Location management is another issue in MSCTP and [59] proposes to use Mobile IP for location management in MSCTP.

SIP, an application layer protocol, provides Internet mobility. However, its performance degrades adversely for real-time applications because of long handover delays and signaling overhead. This is the result of some procedures followed in the SIP-like acquisition of new IP addresses for every new connection or location update followed by the flow of SIP messages. In the case of a TCP connection, the IP encapsulation also causes the overload.

2.4.3 Locator-Identifier Separation Schemes for Mobility Management

The current Internet architecture includes a Domain Name Service (DNS) which is a mapping system resolving human readable URLs into IP addresses for the purpose of locating nodes or computers hosting different web services and information. In the contemporary Internet architecture, an IP address works both as a locator and an identifier for an end node. This dual nature of IP addresses causes problems especially during mobility. The change in location or attachment point of a mobile node results into the change of its IP address. This means a change in mobile node's identifier and as a consequence the desirable property of having a unique host identifier becomes unachievable. A naming scheme for the current Internet architecture can have a persistent naming if the identifier of the node is separated from its locator. This separation of identifier from locator is an important design feature introduced in the recent years. Such separation guarantees a seamless mobility along with many other benefits. Many solution have been proposed such as HIP [7], LISP [37], MAST [60], LIN6 [61], etc. We are going to briefly discuss Locator/Identifier Separation Protocol (LISP) and its extension LISP-Mobile Node (LISP-MN) [62].

2.4.3.1 LISP

The LISP protocol, currently developed by the IETF LISP Working Group, is based on an architecture through which the one number space IP address, working as a routing locator (where the node is attached) and the identifier (nodes ID/name), is separated. It means that the current IP address is actually split into End-point Identifiers (EID) and Routing Locators (RLOC). The nodes use EID as their identifier whereas RLOCs are IPv4/IPv6 addresses used for routing in the network. In order to reach an end node for any information retrieval using its EID, the first step is to look for the current RLOC of that node. In order to solve this issue, the LISP Mapping System [63] is used which is

designed to map EID-to-RLOC. Once RLOC is found, packets with EID information in their header are encapsulated within another header with the RLOC information and are routed to the destination. At the receiving end, the LISP header is first removed before packet is delivered to the destination. LISP has special tunnel routers at the edge of sites designated for LISP encapsulation and decapsulation. Figure. 2.9 shows the basic LISP architecture along with LISP-MN.

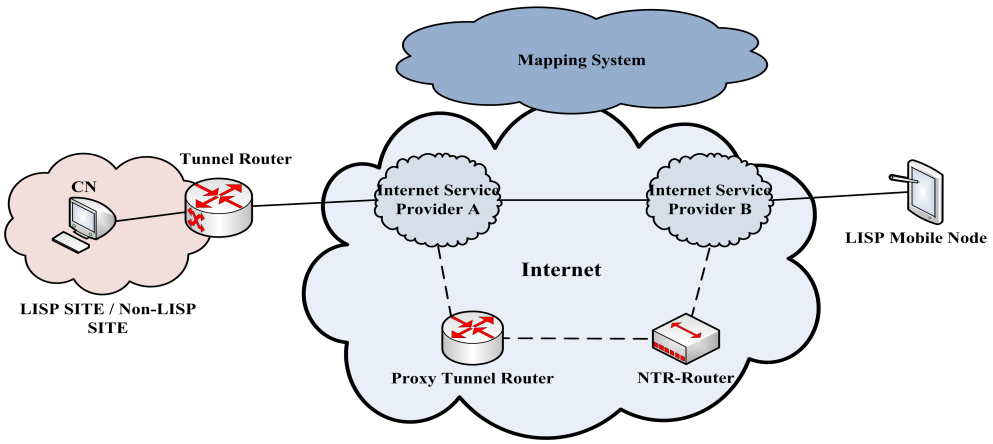


Figure 2.9: LISP-MN and LISP architecture.

2.4.3.2 LISP-MN

The Identifier/Locator split enables a seamless mobility by binding a permanent EID to a host. The RLOCs of the host are not bounded to be permanent and can be changed multiple times during an ongoing communication. LISP-MN, an extesion of LISP, takes advantage of this approach. It is assigned with a permanent EID which is used for every communication any time anywhere in the network. LISP-MN inherits the features the LISP architecture exhibits. Each LISP-MN acts as a LISP site. Packets originating from a LISP-MN are LISP encapsulated and are routed based on RLOCs to the destination host. In a handover situation, the LISP-MN receives a new RLOC and updates its EID-to-RLOC mapping in the associated mapping system to maintain its reachability at its new location. Though LISP-MN acts as a LISP site, still it depends on some of the LISP architecture components. They are briefly discussed below :

(i) **A Mapping System:**
The LISP Mapping System is the component that publishes location information associated with EIDs (EID-to-RLOC mappings). For any handover event, each LISP-MN updates its respective Mapping System with its new EID-RLOC mapping.

(ii) **LISP Inter-networking Components:**
These are the proxies that facilitate the communication between LISP and non-LISP sites or mobile nodes. They decapsulate packets sent by LISP sites and LISP-MNs to the non-LISP sites in the Internet and vice versa:

(a) **Ingress Tunnel Router (ITR):**
ITR is used for communication between LISP sites. When an ITR receives an IP packet at one end, it sends or forwards a LISP encapsulated IP packet from its the other end.

(b) **Egress Tunnel Router (ETR):**
ETR is also used for communication between LISP sites. When an ETR receives a LISP-encapsulated IP packet on one side, it sends or forwards a decapsulated IP packet on the other side.

(c) **Proxy Ingress Tunnel Router (PITR):**
It works in a similar way as ITR. However, the basic difference is that it acts as a gateway between LISP and non-LISP sites. In other words, PITR provides connectivity between the legacy Internet and LISP enabled networks sites.

(d) **Proxy Egress Tunnel Router (ITR):**
It acts as a gateway and decapsulates the packets sent from LISP sites or LISP Mobile Nodes. Similar to PITR, PETR also provides connectivity between the legacy Internet and LISP enabled networks sites.

(iii) **LISP NAT-Traversal Router (NTR):**
In order to facilitate LISP control and data messages (UDP encapsulated) to pass through NAT-boxes, NAT traversal Routers are used. They act as proxies between NAT boxes and MN's incoming packets.

2.4.4 Analysis of Locator/Identifier Separation Schemes for Mobility Management

The performance of LISP-MN compared to the LISP architecture are different. The deficiencies observed with LISP-MN are related to the processes of encapsulation and decapsulation by tunnel routers. This includes both normal ingress/egress tunnel routers (for

communication within LISP domains) as well as proxy ingress/egress routers (for communication between LISP and non-LISP sites). LISP-MN requires a double encapsulation when receiving traffic from non-LISP sites and as result two mapping lookups are needed in the Mapping System. Whereas in the LISP architecture a single mapping lookup is enough. The extra lookup increases the time until the buffered packets can be forwarded to their destinations. This raises the buffer overflow in tunnel routers and causes delay as well. Similarly, packets towards LISP-MN (within the LISP domain) carry two encapsulations until they reache the Egress Tunnel Router (for decapsulation) for the destined LISP-MN domain. This causes problems in the case of Maximum Transfer Units (MTUs) (which is the size of the largest data unit (packet) a layer can pass on), as large packets are problematic during communication errors. Improvements can be made by addressing these issues along with unnecessary path stretch during routing through tunnel routers when the communication between LISP and non-LISP sites takes place. The concept of LISP is new and is still an active working group of the IETF [62] and we hope that these short comings will be alleviated.

2.5 Conclusion

Mobility management in wireless networks is the research domain that has been widely addressed under different contexts. Solutions proposed by the research community with diverse approaches are the reasons for this research area to be still active. Mobility management has been studied with different aspects that include complete mobility management architectures as well as tools and techniques supporting mobility. For example, there are mobility solutions that are based on Internet layers as well as cross layer mobility management solutions such as in [64], [65], [66], [67]. In heterogeneous network environments, the mobility management becomes more interesting and challenging. There are mobility solutions that address the problem for two different wireless networking technologies working together as in [68], [69], [70], [71]. Solutions for specific network architectures have also been proposed as in [72], [73], [74], [75], [76]. IEEE 802.21 Media Independent Handover (MIH) [77] based on mobility architectures [78], [79], [80], [81] also contribute to solve this issue. Overall, the mobility issue has been addressed by the research community by taking into account different criteria, methodology, parameters and wireless technologies.

For our work, we chose to address mobility issues that are related to the current Internet as almost all wireless networks, directly or indirectly, are connected to the core Internet architecture. The short comings of mobility solutions in the TCP/IP stack has been addressed by the Information Centric approach. Built-in mobility solutions rather than add-on solutions (in the current Internet) in the Internet architecture seem more

promising. The contribution of this thesis is one such effort where we proposed mobility solutions that ensure mutual cooperation between network and mobile nodes. In a heterogeneous wireless network environment, this cooperation can be improved if the decision process (for handover) is made efficient. This can be done through individual learning (about the surrounding environment) of the entities of the network (mobile nodes in this case) and work for their individual interests (game theory approach) in a non-cooperative manner.

Chapter 3

Network of Information Mobile Node Architecture

Contents

In this chapter, we present a mobile node architecture framework for the Network of Information named NetInf Mobile Node. An extension of the basic node architecture proposed in NetInf, it is also backward compatible with the existing TCP/IP based networks. The Virtual Node Layer (VNL) and its modules introduced in the proposed framework provide the following support :

(a) Seamless handover with minimum latency.

(b) Data relaying among neighbouring nodes in the network to avoid data loss during mobility.

(c) Power management to avoid loss of mobile node's battery life.

Inner/Outer Locator Construction Routing (ILCTR/OLCTR) are two routing functions introduced in NetInf mobile nodes for NetInf and non-NetInf sites interaction. The basic purpose of NetInf Mobile Node is to maintain the QoS of an ongoing session during mobility events. Handover events are the critical situations during mobility where chances of QoS degradation of an ongoing session are higher. In this work, the proposed VNL algorithm is presented with the help of a scenario where an ongoing session is sustained during handover events between wireless networks. It is expected that the desired QoS during such events can be achieved.

In order to make Information-Centric Networking competitive enough to work along with the legacy TCP/IP protocol suite, one should provide solutions for problems existing in All-IP (Next Generation) networks. The new architecture must be backward compatible with the existing Internet architecture. The range of issues that contemporary IP-networks face are many folds and mobility management is one of them. The current Internet was not designed to cater every problem. Today most of the solutions are impermanent. With the advancement in mobile telephony, new protocols have been developed to handle mobility. The performance of these protocols remained steady in early years. However, during the last decade, the overwhelming development of new applications accessible on smart devices through wireless networks has urged to have more efficient algorithms and protocols. There are two possible solutions: **(i)** either, as usual, provide patch up solutions or **(ii)** design a clean slate architecture which provides built in solutions to all the problems and issues faced by TCP/IP-based networks.

There is a history of extensive work done for mobility management in wireless networks. This domain is still alive and very active as new emerging technologies in the wireless domain always bring new challenges. In a heterogeneous wireless network environment, the IP-layer provides mobility management (especially, vertical handover management) solutions using different protocols and approaches. Moreover, rapid advancements in the mobile

communication domain encourages the development of new ideas and frameworks. Ubiquitous QoS support in wireless network environment is a big challenge. In urban areas, data traffic congestion, channel fading and interference result into intolerable disconnectivity, poor coverage and lack of required QoS. In order to address these issues, the proposed framework is introduced with:

(i) A virtual node layer (VNL) in the NetInf MN. This VNL is a programming abstraction. The concept has been used before in [82], [83], [84], but in the context of Information Centric Networking, the idea is novel.

(ii) We introduce a central entity known as Central Control Unit (CCU) on the network sites supporting VNL coordination between various wireless network technologies especially during handover scenarios. CCU also records and updates the mobility pattern of mobile nodes, predicting mobile node motion and allocation of mobility zones for virtual nodes (explained in later sections).

(iii) Our proposal emphasizes collaboration of network and the end user mobile terminal. The mobility events discussed are neither network controlled nor mobile node controlled. In fact the control is partially divided among them. The VNL with its modules, explained in later sections, along with a cross-layer cooperation, provide seamless handovers with a minimum probability of failure.

3.1 Mobility Management and Quality of Service (QoS)

QoS is closely related to mobility management in wireless networks. The QoS of an application during an ongoing session is mostly disrupted during handover/handoff scenarios. Today, heterogeneous wireless network environment demands seamless handover between various network access technologies. A lot of work has been done so far in this domain. Today, factors that trigger handover are not just limited to the measured value of the Received Signal Strength (RSS). The list involves data rate requirement for a particular application, end-user demand for high/low data rate, packet loss and more. There is a wide range of protocols and algorithms across all the layers in the TCP/IP protocol stack for supporting mobility. For example, Mobile IPv4/IPv6 and LIN6 at the Network layer. The Stream Control Transmission Protocol (SCTP) and the Datagram Congestion Control Protocol (DCCP) provide mobility support at the transport layer level. Session Initiation Protocol (SIP) , Dynamic DNS (DDNS) and MOBIKE are few examples of application layer protocols. The performance of these protocols are limited due to the lack of cross-layer cooperation between lower and higher layers. The dual nature of an IP address, acting as a locator as well as an identifier, degrades the efficiency of these protocols during

mobility events. The separation of locator from its identifier aims at countering this issue. Host Identity protocol (HIP), Locator Identifier Separation Protocol (LISP) are examples of such proposals. HIP works in collaboration of transport and network layers. LISP is a network-based approach and focusses on limiting the size of routing tables and improving scalability and routing in the system. LISP Mobile Node, which is the extension of the LISP architecture, has multiple design goals including wide range of communication possibilities for different mobility cases.

The use of devices capable of accessing different radio interfaces urges to have an architecture that facilitates the seamless handover of sessions among different wireless access technologies. The IEEE 802.21-Media Independent Handover (MIH) standard supports the collaboration between various link layer access technologies. MIH consists in a framework providing services to users. The framework includes various entities for transmitting and receiving messages, to share information about various access network's capacity. In other words, this framework defines a protocol stack, implemented on each mobile device for seamless handover. The cross-layer cooperation also has been studied extensively [85], [86], [87], [88]. The information exchange between different layers of the stack improves the overall handover process by avoiding false alarm signals and minimizing the latency during handover to reconnect or update.

3.2 Mobility Management in NetInf

To handle mobility issues in NetInf, integrated name resolution and routing schemes are proposed. The proposed solutions include a Multiple Distributed Hash Table (MDHT) [8] approach for the core network, the Late Locator construction (LLC) [9] scheme, which is an extension of MDHT for access networks, and an autonomous local resolution using multicast (providing access to all local contents even when the network is not connected to the core network). In MDHT, as the name indicates, DHTs are arranged in a hierarchical manner. As an example of an Internet Service Provider (ISP), there are four levels:

- Access node level
- Point of Presence
- Autonomous System level
- Global Internet level

However, the structure can be changed depending upon the size of the network. There are name resolution platforms/nodes at each level of MDHT. These special platforms/nodes

are addressed as Dictionary Nodes (DN). They are actually a network-based implementation of a DHT system. They perform a name resolution and location look up service both at local and global scope of the network. The object (information) is published and duplicated at all levels. As far as LLC is concerned, it separates the core network routing from the edge network routing. It uses path-based locator. In LLC, packets are forwarded in a connectionless manner. A path-based locator for a object consists in a core edge router prefix appended with a sequence of identifiers that describes a path across a sequence of edge networks towards the host, hosting destination object. In general, LLC employs an object identifier/locator split mechanism. The common ground between MDHT and LLC schemes is that any one can be switched in depending upon the requirements. It means what kind of mobility case NetInf is dealing with. In case of the mobile terminal mobility, LLC is taken into account and the MDHT scheme is approached for network mobility.

The evaluation of both schemes, namely MDHTs and LLC, encourages the need to have more optimization in terms of reducing the handover latency during the mobility event. Further, in case of intermittent connectivity the delays are intolerable. One research challenge also indicates to have an interaction between mobility and caching to reduce the overall look up delay. This latency issue is mostly concerned with the mobile access networks that are not part of the actual core netwrok. These mobile access networks include all radio networks ranging from mobile telephony (3G, 4G, GSM, CDMA, etc.) to WiFi and WiMax.

The issues discussed above are very crucial in multi-access environment. In a large metropolitan environment, multi-technology enabled mobile devices always demand the best connection for better QoS. For real time streaming, smooth shift or handover between different access technologies is anticipated. Factors mentioned earlier like congestion, channel fading and unprecedented mobility of users are the challenges hindering good service.

3.3 NetInf MN Architecture

In NetInf, the node architecture defines the general framework of a typical information centric network node. Based on this architecture, we developed the Network of Information Mobile Node (NetInf MN) [89] framework for wireless networks. Along with the new features introduced, a NetInf Mobile Node is compatible with any environment or wireless access technology. There are some similar characteristic between NetInf MN and Locator Identifier Separation Protocol Mobile Node (LISP MN) [62]. However, LISP Mobile Node has a totally different working framework. It includes servers that facilitate the mapping from end point identifiers to locators. NetInf has a totally different infrastructure for name resolution. Still there are some common features that NetInf MN shares with LISP MN such as both nodes encourage to:

- Keep ongoing sessions alive during mobility even if all nodes are moving simultaneously.
- Have the possibility of simultaneous connections.
- Make mobile nodes to act as servers.

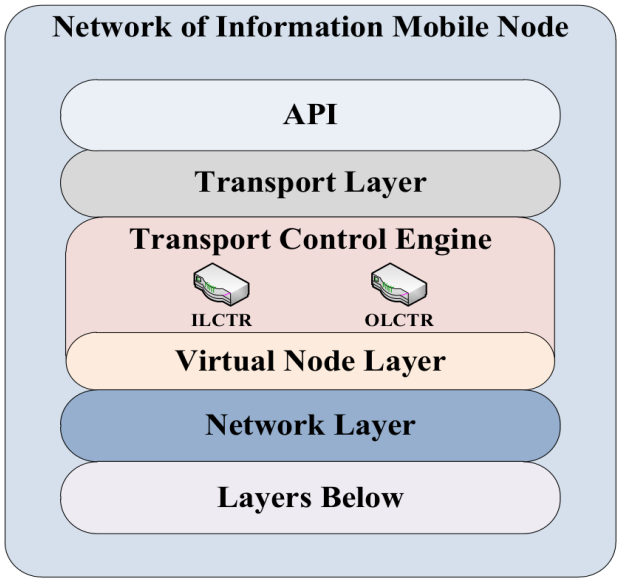


Figure 3.1: Network of Information Mobile Node.

In this work, the proposed solution for mobility management in the network is presented in the form of a mobile node architecture. Figure. 3.1 presents the NetInf MN architecture framework. We have an application layer which provides services to users. This API gets services from the transport layer below it. The Virtual Node Layer (VNL) includes a Transport Control Engine (TCE). Generally in the NetInf node architecture, TCE is responsible for the coordination of different protocols used for accessing NetInf objects. In the NetInf MN design, Inner Locator construction tunnel Routing (ILCTR) and Outer Locator Construction Tunnel Routing (OLCTR) functionalities are additionally included in TCE. Virtual Node Layer (VNL), ILCTR and OLCTR within VNL are responsible for mobility management of NetInf MN.

3.3.1 Virtual Node Layer (VNL) (Mobile Agent Generalization)

The main problem faced in wireless networks is the unpredictable motion of mobile nodes. This feature though facilitates end-user's mobility when compared with fixed (wired) telephone systems but results into surprise visits of mobile nodes in and out of a network. The availability of nodes leaving or joining a network thus cannot be predicted. This feature makes designing of algorithms difficult for wireless networks. Cases where the movement pattern of mobile nodes is known makes thing simple. For example, in the case of Wireless Sensor Networks (WSN) where fixed mobile nodes are placed in an area and a mobile node frequently visits that area to collect the information. In such a case, the mobile node movement is programmable and can even perform efficiently. Unfortunately, this cannot be done for users of mobile devices. It is, in fact, not practical for such devices to follow programmed instructions. Thus, the objective here is to ensure that, (a) the protocols or algorithms remain effective in terms of their performance and (b) they do not oblige mobile nodes to follow specified instructions.

VNL is a programming abstraction in NetInf-MN architecture. It can be considered as an autonomous program like a Mobile Agent (MA) that can migrate from one node to another in a network. In other words, the program running at a host can suspend its execution and move to another node and resume its working from the point of its suspension. In order to follow the objectives mentioned above, we propose that the proposed network algorithm should be implemented on VNL in the NetInf Mobile Node architecture. Executing an algorithm on the virtual layer can help in pre-determining the position of mobile nodes. This is done with the help of the Centralized Control Unit (CCU) (see Figure.3.5) in the network. We propose to have this unit installed in network sites to assist mobile nodes. A mobile node visiting a network site synchronises its VNL with the available CCU. CCU along with ILCTR and OLCTR, has different functions that are used to assist mobile nodes in different situations.

3.3.1.1 Transport Control Engine (TCE)

The Transport Control Engine, at the Transport layer, coordinates which protocol should be used for accessing the data. For NetInf defined objects, different protocols are used. In case of accessing IP-based information, TCE can switch NetInf Mobile Node to work as a normal TCP/IP-based mobile node. TCE also coordinates the use of ILCTR and OLCTR functions defined in VNL. In case of data relaying, TCE coordinates with Data Relaying module of VNL for buffering or storing the data temporarily during the disconnected periods. Similarly, in handover situations, TCE coordinates with the Handover module of VNL. In general, TCE chooses which mode NetInf mobile node should be in, as per the network

requirements.

The Inner and Outer Locator Construction Tunnel Routing (ILCTR/OLCTR) functionalities within VNL are mutually used by Handover and Data relaying modules. It is assumed that the network has two different mobile nodes: (*a*) a NetInf Mobile Node (NMN) that looks like a NetInf site originating packets with NMN encapsulated headers and (*b*) a Non-NetInf Mobile Node (NNMN) that originates packets with IP headers.

3.3.1.2 Inner Locator Construction Tunnel Routing (ILCTR)

The ILCTR function, as shown in Figure. 3.2, makes NetInf MN to work as a router which, when accepting a packet with an IP header only from NNMN, encapsulates it as an NMN packet. In other words all packets originating from the NMN and all packets being forwarded by NMN are NMN encapsulated.

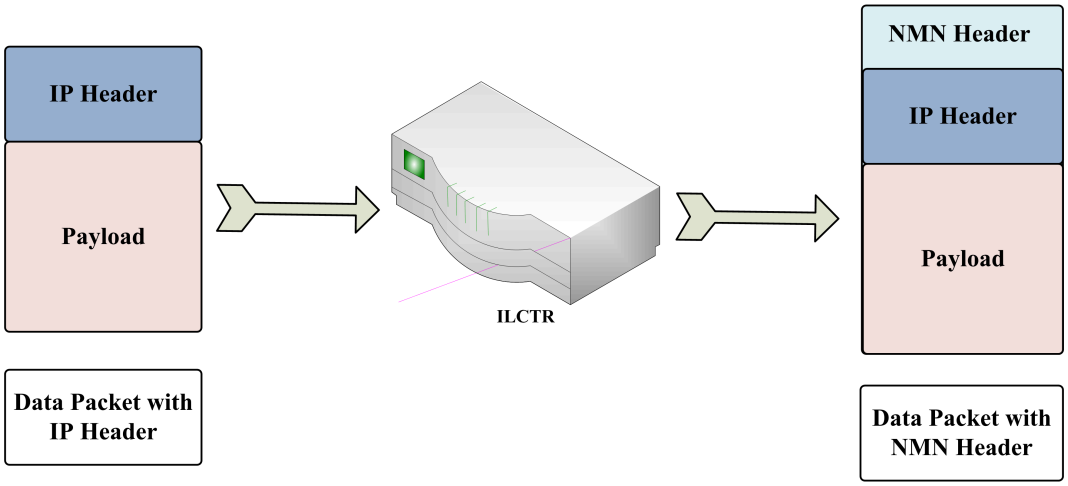


Figure 3.2: Inner Locator Construction Tunnel Routing

3.3.1.3 Outer Locator Construction Tunnel Routing (OLCTR)

OLCTR, see Figure. 3.3, also makes NMN to work as a router. However, it works in a different way. It decapsulates the received packets that have NMN headers. The outer header is taken off and the packet is forwarded if addressed to some other node.

Virtual Node Layer further classifies its functions into three different modules as shown in Figure. 3.4. Each module can work independently or mutually depending upon the task assigned to it. The details of each VNL module is given below.

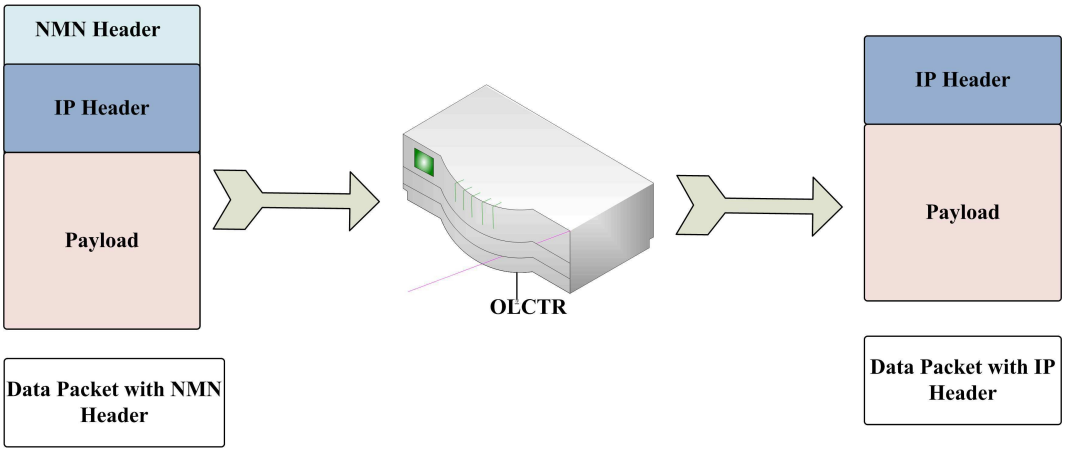


Figure 3.3: Outer Locator Construction Tunnel Routing

3.3.1.4 Handover Module

This module works closely with the Data Relay Module. The main characteristics of this module is to make sure that the handover should be seamless with minimum latency during ongoing communications, i.e. there should be a minimum or almost no disruption. Similarly packet loss should also be minimized during handover and the probability of handover failure should be limited. This last characteristic is difficult to achieve as it is directly related to the user's speed with which it moves. However, the scenario discussed in this chapter, considers a population where most of the users are pedestrians.

3.3.1.5 Data Relay Module

The Data Relay Module in VNL, together with the Handover Module, supports NMN during handover situations. Similarly in challenged environments, where long delays are expected due to frequent disruption and disconnections, this module provides DTN like service [35]. Challenged networks are characterised by high-latency, low-data rates, frequent disruptions, long queuing duration and path instability. Like in the DTN architecture, ILCTR and ILCTR makes NMN to work as storer and forwarder. It means, if there is a disconnection at some point during an ongoing communication in a multi-hop environment, NMN can act as a storage device until the connection is re-established.

3.3.1.6 Power Management Module

Handover and Data Relaying modules, with varying channel conditions, try to maximize or maintain the QoS during mobility events by respecting mobile battery power constraints. Still, in order to economize the use of mobile battery power, a mobile node should follow the optimal transmission scheme.

The power management module works in the conditions where the mobile node is inactive or idle or left with low battery power. In such cases, the regular query to update the mobile node location must be suspended. This is very useful in terms of power management of mobile nodes and reducing the signalling overhead of the overall network.

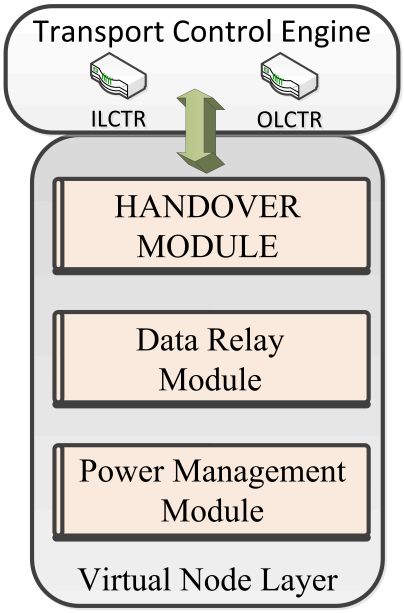


Figure 3.4: Virtual Node Layer Modules.

3.3.2 Central Control Unit (CCU)

Though, CCU is not part of NetInf Mobile Node architecture framework, without its coordination, mobile nodes (NMN or NNMN) cannot perform some tasks. CCU, as shown in Figure. 3.5, has different working units that provide assistance to mobile nodes. Each unit has different functions and are discussed in the following subsections.

3.3.2.1 Mobility Pattern Data Base Unit

It is generally considered that mobile nodes movement is random in a network. However, users of a particular region can have identical mobility patterns over a period of time. For

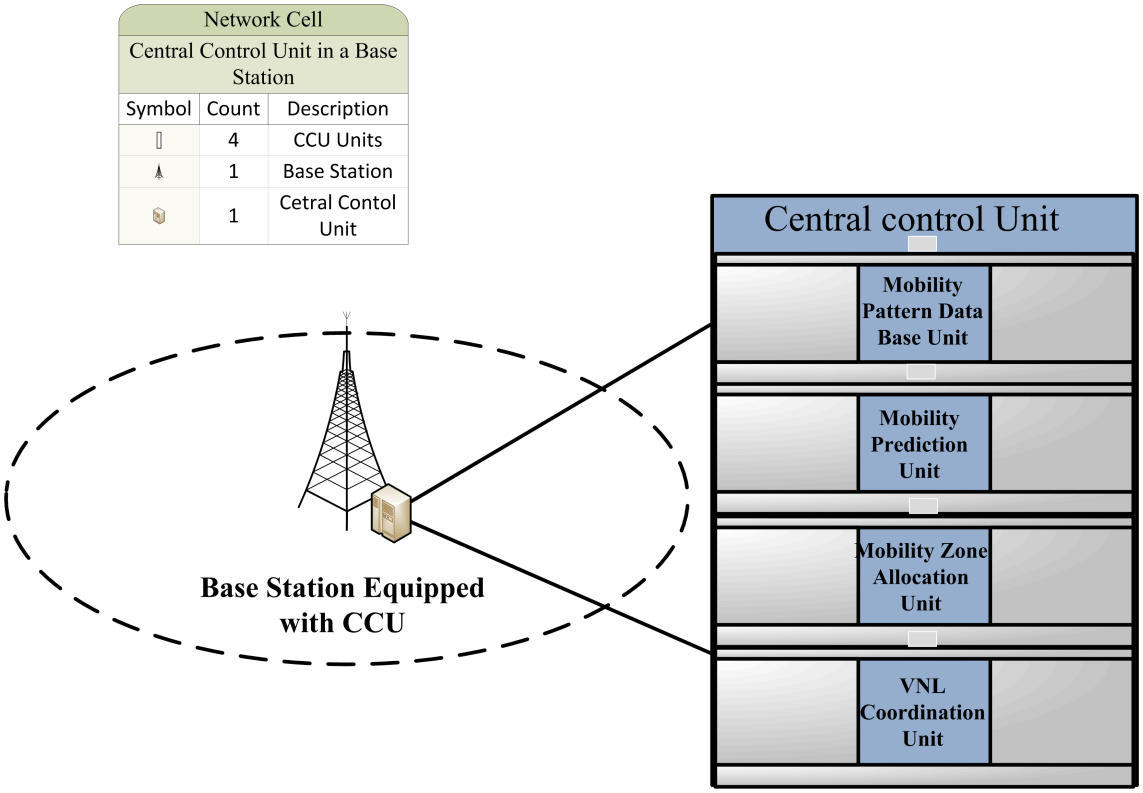


Figure 3.5: Central Control Unit.

example, a student living near to his university has a the same mobility pattern during weekdays (most of the time). Similarly, some mobile node users can have unpredictable random motion all the time. For example, a marketing salesman, visiting new areas on daily basis, promoting his/her company's new product. Such classification of cellular network subscribers is done by CCU's mobility pattern data base unit. The mobility pattern of mobile nodes provide useful information like their position, speed, acceleration, geographical information about the area, etc. All these information can be preserved as the mobile node's history of activities or mobility pattern and can be used for predicting mobile nodes mobility in the future. In CCU, Mobility Pattern Control Unit works closely with the Mobility Prediction Unit.

3.3.2.2 Mobility Prediction Unit

The Mobility Prediction Unit basically targets estimating mobile node's future position. Through prediction, mobile nodes location and handover process can be managed. Maintaining up to date information of mobile nodes location in the network helps in delivering data packets in Virtual Mobility Zones, to data relaying nodes (NMN) and to continue the relaying process during handover events without interruption. Similarly, during handover, predicting the next network cell to which a mobile node is going to migrate helps in initiating the handover process in advance. As a consequence, the probability of handover failure can be minimized.

Together with the Mobility Pattern Data Base Unit, this unit predicts the next move of a mobile node with no mobility pattern record available. There are various methods by which mobility prediction can be estimated. For example, the use of Kalman filters [90] for tracking the trajectories of mobile nodes in a cellular network. By predicting mobile nodes location, the quality of services can be improved extensively as reported in [91].

3.3.2.3 Mobility Zone Allocation Unit

The Mobility Zone Allocation Unit of CCU allocates specific coverage zones in a network area. These zones are known as Virtual Mobility Zones as shown in Figure. 3.6 and are allocated based on the mobile nodes activity inside a network cell. Since they are virtual, they can be removed and redefined based on network requirements. A network cell area that is not frequently visited by mobile nodes is not preferable to have a VMZ. Similarly, if an area (already allotted with such a zone) is not being visited frequently after sometime, it is also subject to be revoked from the list of Virtual Mobility Zones.

These zones can be considered as *locus of data relay and handover support*. It means, the more frequent a network area is visited by mobile nodes, the higher the probability this area has to be allotted with a VMZ. NetInf Mobile Nodes in a wireless network provide data relaying and handover support. Since all mobile terminals, including NetInf and Non-NetInf mobile nodes, move unpredictably in random directions, relaying data and support during handover seems to be difficult. This problem is solved by these allocated VMZs within the network cell. In case of a highly dynamic environment where mobile nodes are in continuous random motion, NMN nodes (providing data relay and handover support) moving out of VMZs transfer the relaying data and/or handover support links to the new NMN nodes entering these VMZs. This way, the connectivity is ensured all the time. This unit together with VNL Coordination Unit support mobile nodes during handover situations. These zones are allocated close to the cell boundaries and are used when mobile nodes have some probability of migrating to neighbouring cells of the same or different

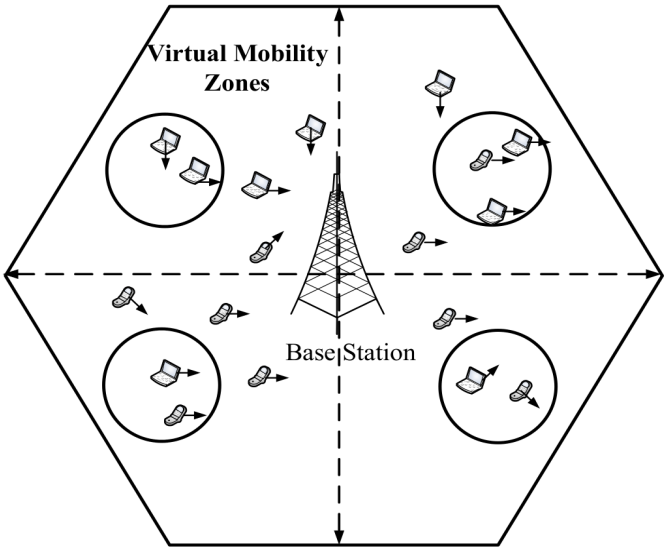


Figure 3.6: Virtual Mobility Zones.

wireless networking technology.

3.3.2.4 Virtual Node Layer Coordination Unit

A mobile node visiting VMZ is assisted by a VNL Coordination Unit along with the Mobility Zone Allocation Unit. A mobile node continuously monitors its RSS value. The VNL is activated when the RSS value starts deceasing, indicating that the mobile terminal is moving away from the base station or the access point. Along with Mobility Zone Allocation Unit, VNL modules participate in different operations. For example, in the case of a weak connecting link or during handover events, Data Relay and Handover modules are used. For power management, power management module is used.

3.3.2.5 Cross Layer Support for VNL

For Data Relay and Handover modules, a cross-layer support is necessary as shown in Figure. 3.7. Link layer measures the Received Signal Strength (RSS) and Network layer measures the Handover Delay Signal. The RSS value is affected by the mobile terminal speed and its position relatively to the Base Station or the Access Point. The handover delay signal is measured when a mobile node is in the close proximity of a neighbouring Base Station or Access Point. A simple way to measure this value is sending of a Network layer mobility management protocol's (e.g. MIPv6) registration request and compare the time difference between the transmission time and the reception time of the reply as suggested

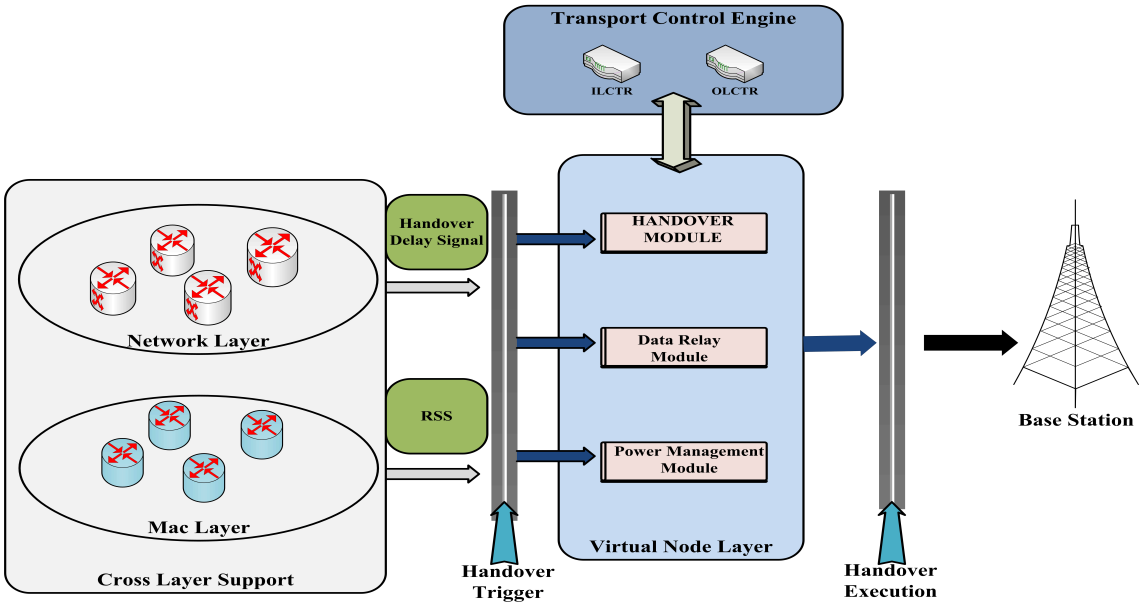


Figure 3.7: Virtual Node Layer with Cross Layer Support

in [87]. This measurement is important during handover as it can avoid the ping-pong effect during handover from one network to another. However, this method increases the signaling overhead. To avoid this situation, CCU uses a Mobility Zone Allocation Unit that allocates VMZ near the network cell boundary for mobile nodes assistance during the handover situations. For overlapping network cells, the former method can be adopted by limiting the number of queries, while the latter is useful for both overlapping and non-overlapping network cells.

3.4 VNL and CCU Support for Data Relaying and Handover

The basic purpose of VNL and CCU coordination is to support mobile nodes during mobility in general and during handover in particular. Another aspect of VNL is to relay data on behalf other neighboring nodes in the network. The main purpose of these services is

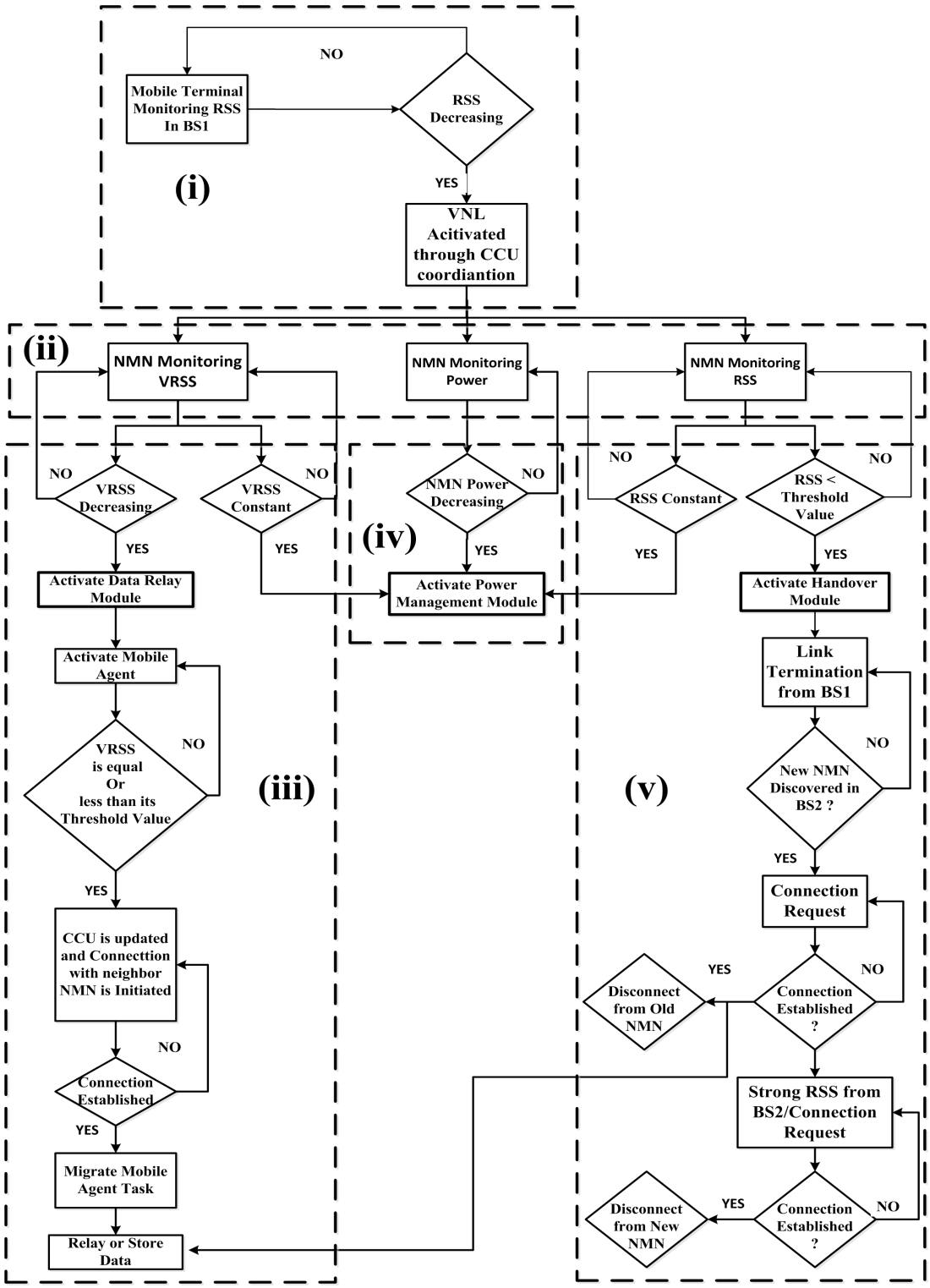


Figure 3.8: VNL Algorithm

to maintain QoS of an ongoing session during mobility events. The end user in a wireless network always wants ubiquitous connectivity and to experience the best service. However, in a highly dynamic heterogeneous wireless network environment, mobile nodes with multiple interfaces for different wireless network technologies, switch their connections because of different facts. The main fact is to be always best connected in order to experience the best QoS. However, this connection switching among heterogeneous network technologies or homogeneous networks should be seamless. In other words, during handover, VNL and CCU coordination should make sure to have:

(a) Minimum handover latency

(b) Minimum Packet loss

(c) Least probability of handover failure

3.4.1 VNL Algorithm

The VNL algorithm is executed in parts, both at the CCU (of a base station or an access point of a wireless network) and at the VNL of a mobile terminal (NMN). As shown in the Figure. 3.8, all of the three modules of a VNL take part depending upon different scenarios. The following steps briefly explain the VNL algorithm execution procedure:

(i) The Mobile terminal (NMN/NNMN) monitors the RSS from base station BS1 to which it is currently connected. Sensing that RSS is decreasing as a result of its movement away from BS1, this results in the activation of VNL through the CCU1 coordination.

(ii) The algorithm has now three options to work on. If it is in the VMZ, it monitors VRSS (Received Signal Strength within VMZ) and also constantly measures its battery power. If not in VMZ, RSS monitoring is continued.

(iii) In VMZ, there are two possibilities while measuring VRSS. If VRSS is constant, the Power Management Module is activated to save mobile energy. This means that the mobile node is now idle and not communicating. The other case is when VRSS is decreasing: in such a situation, the data relay module is activated.

- Once the Data Relay Module is activated, the mobile terminal measures VRSS and monitors it until it becomes equal or less than the assigned minimum threshold value ($VRSS_{min}$). The Mobile Agent that has already been activated along with the Data Relay Module is ready to take an action whenever the $VRSS_{min}$ value is reached.

- In the mean time, the mobile terminal contacts the neighbouring nodes in VMZ for data relaying whenever VRSS $\leq$ VRSS$_{min}$. CCU is updated and a connection with the relaying node (NMN) is established. The mobile agent transfers the data relaying task of the mobile terminal to the relaying node (NMN). NMN relays or stores the data depending upon the requirement.

(iv) When monitoring its power, the mobile terminal activates the Power Management Module in situations where the battery life is at its critical level. This is done to economize power utility.

(v) RSS monitoring after VNL activation can be divided into two phases. The first phase includes the activation of the Power Management Module if the RSS value becomes constant over some duration. The second phase includes the activation of the Handover Module when the mobile terminal is exiting VMZ and moving towards the cell boundary.

- In the first phase, if the mobile terminal is inactive or immobile, making RSS constant, the Power Management Module is activated.
- In the second phase, if the RSS value is decreasing continuously, the mobile terminal activates the Handover Module whenever RSS $\leq$ RSS$_{min}$.
- As mentioned earlier, the mobile terminal is temporarily connected to NMN (relaying node) in VMZ, even after the link between BS1 and the mobile terminal is terminated, the communication is still continued through BS1 via NMN (relaying node).
- In the mean time, if another relaying node (new NMN) is discovered in the vicinity of BS2, a connection request is sent as the link quality of the connection between the old NMN and the mobile terminal starts becoming weak.
- If the connection with the new NMN (relay node) is established, the link with the old NMN is aborted and the data of the mobile terminal is now relayed by new NMN.
- The new connection is continued in the BS2 cell until a strong RSS value ensures the existence of a stronger connection between mobile terminal and BS2. Once a new connection is established between the mobile terminal and BS2, the link between the new NMN and BS2 is terminated.

3.4.2 VNL Working Principle

The VNL working principle is explained with the help of a scenario presented in this section. As explained earlier, VNL is a programming abstraction, implemented on NetInf-

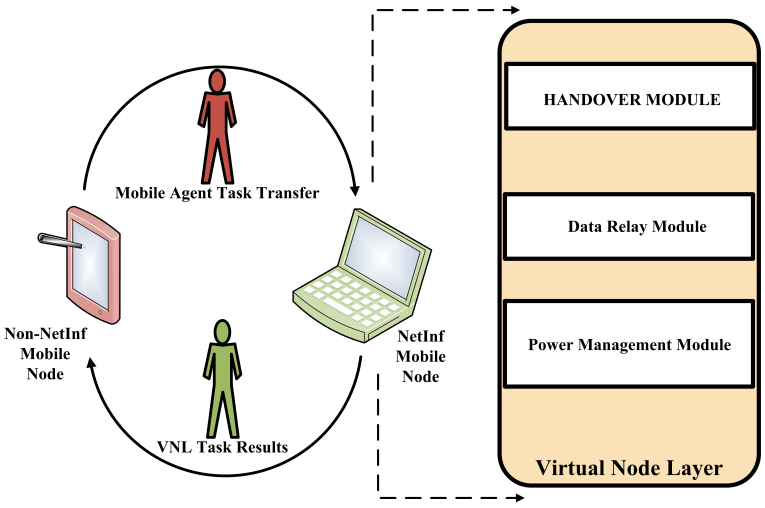


Figure 3.9: Mobile Agent and VNL Working.

MN (NMN). For Non-NetInf-MN (NNMN), a Mobile Agent (MA) is used which transfers its tasks to NMN (NetInf Mobile Node) as shown in Figure. 3.9. VNL, which is an extension of a Mobile Agent, completes the required task and submits it back to NNMN. The service provided by NMN can be different for different situations. A Mobile Agent is a computer program or a software that can migrate from one computer (suspending its execution) to another (resuming again from where it was suspended). The important features of a Mobile Agent are its autonomy for executing any task, its learning behaviour according to the environment and especially mobility. The last feature makes this technology favourable for the case presented in this work. As they are autonomous in nature, they can migrate to any another computer in the middle of their execution. For distributed applications, they are considered to be very powerful tools. They are easily portable and does not require specific system requirements.

3.4.3 Handover Scenario

The handover scenario in this section explains the working principle of VNL implemented in NMN. However, the following assumptions are considered for the setup shown in Figure. 3.10.

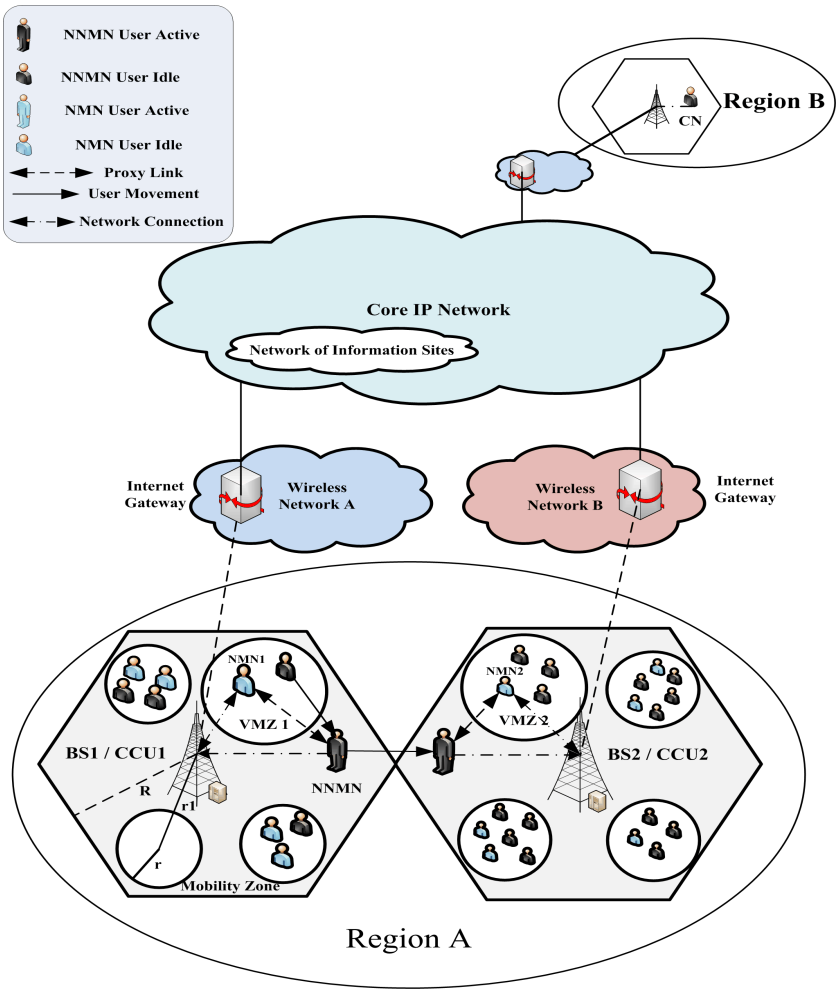


Figure 3.10: Handover Scenario

3.4.3.1 Assumptions

- The wireless network environment can either be heterogeneous or homogeneous, populated with NMN and NNMN nodes connected to their respective base stations (BS1/BS2). For a heterogeneous wireless network environment, mobile nodes having multiple interfaces can switch to different access technologies, e.g. Wi-Fi (IEEE 802.11 a/b/g/n), Cellular data service (3G/LTE) and WiMax (4G) technologies. The connection is changed either during a handover or whenever a user demands for high-data or bit rate per unit time. Sometimes, a mobile application also requires a network connection with a higher bit rate.

- Each network coverage area is assigned with VMZs. This is done through the CCU

of the network. A VMZ is considered as a *locus of mobile nodes* if it is frequently visited. An allocated VMZ is not always permanent. It can be changed or removed if mobile nodes activities in an allotted virtual zone is reduced. VMZs are assigned close to the cell boundary of the wireless network as shown in Figure .3.10.

- Though virtual, VMZs have geographical distance relative to their network's Access Point or Base Station. In the current case, the center to center distance between a VMZ and a Base Station is $r1$. The radius of a VMZ is r, whereas the radius of the network cell is R. $r1$ is measured by taking an average sum of all the $r1$ distances recorded by the CCU. The size (radius r) of a VMZ depends on the mobile nodes activity in that zone.

- The network cell and the VMZ, within the cell, consider two different RSS values. The general RSS value is considered all over the cell except in VMZ, whereas VRSS (Virtual Received Signal Strength) is measured for both NMNs and NNMNs instead of RSS in VMZ.

- Both RSS and VRSS have minimum threshold values, RSS_{min} and VRSS_{min} respectively. These limitations are useful for initiating handover and data relaying processes.

- All mobile nodes joining a new network register their presence as a first priority. The CCU of each cell (Base Station in the considered case as shown) maintains a history of mobile nodes activities. It holds useful information related to mobile nodes mobility trajectories and their behaviour. If the same nodes rejoin the network, it is easy to manage their requirements.

- We assume that one of the mobile nodes in the BS1 cell of Region A is in communication with corresponding node CN in Region B as shown.

3.4.3.2 Sequence of Events

The scenario in Figure. 3.10 are explained by the sequence of events shown in Figure. 3.11. The step wise details are provided below:

(i) In BS1, a NNMN is in session with a Corresponding Node CN as shown in Figure. 3.10. NNMN starts moving away from BS1. The RSS value starts decreasing. In VMZ1, NNMN measures VRSS.

(ii) While moving across VMZ1, moving away from BS1, VRSS decreases. When $\text{VRSS} \leq \text{VRSS}_{min}$, NNMN broadcasts a request for connection from NMNs in VMZ1. NMN1

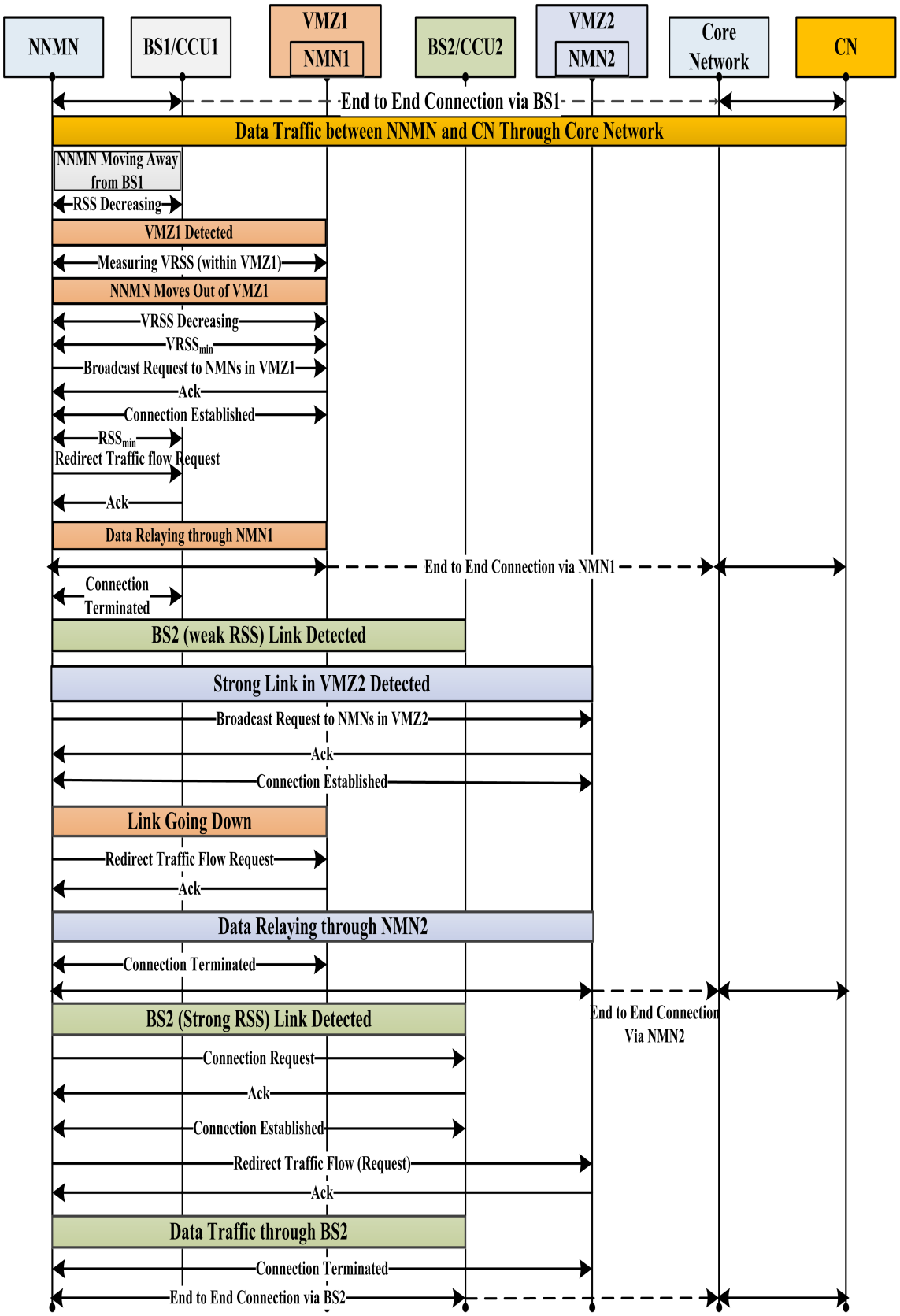


Figure 3.11: Sequence of Events in Handover Scenario.

accepts and acknowledges the request. The connection is established between NNMN and NMN1 in VMZ1.

(iii) Moving towards the BS1 cell boundary, the received signal strength becomes minimum (RSS_{min}). NNMN requests BS1 to redirect the data traffic flow via NMN1. BS1 acknowledges this request. The data is now being relayed through NMN1. The end to end connection between NNMN and CN is supported via NMN1. The connection between NNMN and BS1 terminates as the link between them goes down.

(iv) Moving towards the BS2, as shown in Figure. 3.10, NNMN detects a weak link from BS2. However, within BS2, NNMN detects strong links from VMZ2. Broadcasting connection requests to NMNs in VMZ2, NMN2 accepts and acknowledges the request. The connection between NNMN and NMN2 is established.

(v) Meanwhile, the link between NNMN and NMN1 weakens, NNMN requests NMN1 to redirect the data traffic. NMN1 acknowledges the request. The data is now being relayed through NMN2. The link between NNMN and NMN1 is terminated. The end to end communication between NNMN and CN is supported via NMN2.

(vi) As soon as NNMN detects strong RSS signals from BS2, a connection request is forwarded. BS2, after verifying the NNMN authenticity, accepts the request. Once the connection is established, the data traffic redirection request is sent to NMN2. Acknowledging the NNMN request, NMN2 forwards this request to BS2. The data traffic is now being relayed by BS2. The connection between NNMN and NMN2 is terminated. The end to end connection between NNMN and CN is continued via BS2.

3.5 Conclusion

This chapter presented a brief overview of the proposed mobility solution for the Network of Information followed by an introduction to our proposed NetInf mobile node architecture framework. The details discussed in the later subsections highlighted the different features exhibited by the NetInf mobile node. The Virtual Node Layer (VNL) abstraction is the generalization of Mobile Agents with additional features represented by Handover, Power and Data relaying modules. The ILCTR and OLCTR functions in the Transport Control Engine (TCE) facilitates the data transfer between NetInf and Non-NetInf sites. A NetInf mobile node works closely with the Central Control Unit (CCU) to perform its functions. During a handover and data relaying, mobile nodes are mutually assisted by the network and NetInf Mobile Nodes. The CCU unit operates to perform diverse tasks, e.g. updating

the data base of mobile nodes mobility patterns and prediction of mobile nodes movement. However, the basic role CCU plays is the allocation of Virtual Mobility Zones (VMZs) and assisting NetInf Mobile Nodes to activate VNL. This is made possible by proposing the VNL algorithm, explained through the handover scenario presented.

In the following chapter, we are going to present the Reinforcement Learning (RL) based on the Stackelberg model to show that the concept of data relaying during handover in wireless networks (presented in this chapter) can sustain the QoS of an ongoing session between mobile nodes. This is guaranteed by adopting strategies that maximize individual as well as overall utilities of mobile nodes and the network respectively. This game theoretic-model explains how in a non cooperative wireless network environment, mobile nodes learn through experimenting different strategies. The simulation results shows efficiency and stability of the adopted method and provide proof for the concept presented in this chapter.

Chapter 4

Reinforement Learning based Game-Theoretic Model for Data Relaying and Handover Management

Contents

This chapter presents a Price-Reward learning scheme to encourage mutual coordination between mobile nodes and their wireless networks. In order to maximize the overall network coverage through cooperative diversity, a Nash-Stackelberg Multiplicative Weighted Imitative CODIPAS-RL scheme is proposed. The wireless network implements a 2-level Stackelberg game by introducing Price-Reward parameters (λ,μ) whereas the Reinforcement Learning (RL) scheme paves the way for mobile nodes to reach a Nash-Equilibrium state. The performance evaluation of the learning scheme for the presented scenario proves a fast convergence towards the optimal solution by adopting different sets of actions for the selected strategies. This ensures QoS sustainability during handover situations by data relaying and avoids collisions among mobile nodes while accessing network resources.

The use of the Internet in wireless networks always demands for better QoS during mobility. In [92], we proposed a handover and data relaying algorithm which showed how cooperative diversity in wireless networks increases network coverage and revenue and ensures connection reliability during handover situations. In a wireless network where mobile nodes are highly dynamic and are moving in random directions, a mobile node will avoid cooperating with the other nodes because of (**a**) the loss of unnecessary battery life and (**b**) the disruption or delay of personal data. Hence, it is necessary to device a strategy which should encourage mobile nodes to cooperate in a wireless network.

Every mobile node in a wireless network competes to maximize its throughput. A Price-Reward strategy can be adopted that tempts nodes to cooperate. In this chapter, we have extended our effort to show how the network and nodes can work in a cooperative manner to maximize their individual utilities. This is achieved by following strategies that eventually lead towards Nash-Stackelberg equilibrium conditions [29], [27]. We developed a scenario to model our problem where a network acts as a leader and the mobile nodes as followers. In this 2-level Stackelberg game, the leader (WLAN Access Point in our case) proposes Price-Reward parameters (λ,μ) in the network. Mobile nodes avoid following their selfish non-cooperative behavior knowing that the network encourages data relaying.

In a dynamic environment where a mobile node follows random paths without knowing the transition probabilities from one state to another requires a Reinforcement Learning (RL) scheme [93]. In RL schemes, the learners interact with their environment and use their experience to choose or avoid certain actions based on their consequences. Actions that led to high payoffs in a certain situation (or state) tend to be repeated whenever the same situation (state) recurs, whereas choices that led to comparatively lower payoffs tend to be avoided.

Based on the above discussion, the proposed Nash-Stackelberg Multiplicative Weighted Imitative CODIPAS-RL scheme is designed for fast convergence to reach the Nash Equilibrium condition followed by the Stackelberg Equilibrium state.

The contributions reported in this chapter are :

- Data relaying and handover management algorithm to ensure seamless mobility.
- Proposal of a Nash-Stackelberg Multiplicative Weighted Imitative CODIPAS-RL scheme with fast convergence rate towards desirable solutions for all the players in the game.
- Enhancement of the overall network revenue and coverage through cooperation between mobile nodes and the wireless network.

4.1 Game Theory

Game Theory is a mathematical tool, used extensively for modelling situations with varied numbers of players. John Nash [29] contributions in this field made it an active discipline in recent decades. Though, Game Theory's major application is considered to be in the field of economics, it has also been widely used in other fields like Biology, Sociology, Political Science, Telecommunications and Wireless Networks. In Telecommunications and Wireless Networks, most of the addressed problems are related to resource allocation and routing.

The basic notion of using Game Theory is to solve any situation in an environment where there exist a competition among different actors or players for accessing available resources. The situation becomes more interesting when resources are limited and players have conflicting interests. This analogy is applicable in wireless networks where mobile nodes are the players and competing for maximizing their payoffs or utilities. In wireless networks, any situation in any game or strategy demands cooperation between players (wireless nodes). It means that each node has to share some of its resources in a distributed way. This rule gives each player the liberty of making independent decisions and is the basis of *non-cooperative games*. Whereas, in *cooperative games*, groups or coalitions of players enforce cooperation. Our contribution here deals with the theory of *non-cooperative games*.

Considering a wireless network, a non-cooperative game is defined by the system model (wireless network model) consisting of players (wireless nodes) or decision makers with personal interests. This selfish behaviour leads to possible conflicting situations. However, it is assumed that each player makes rational decision in order to maximize its payoff or utility. Where payoff/utility is defined here as the measure of satisfaction in terms of Quality of Service (QoS) experienced by a player. The decisions made by players are termed as strategies where a strategy in a game is an analogy to an action in wireless networks related to some functionality. For example, in a wireless ad-hoc network, the decision of node to send or not send a packet.

In order to represent a game and its components (discussed above), there are two forms: the Normal-form game and the Extensive-form game. For the sake of simplicity, the

Normal-form game is discussed here which, unlike Extensive-form, is not graphical but a matrix representation of the game. A game using the Normal-form structure is represented as :

$$G = \langle N, A, U \rangle \tag{4.1}$$

where $N = \{1, 2, 3, ..., m\}$ is the set of player numbers, $\mathcal{A} = \{a_1, a_2, a_3, ...a_m\}$ is an action or strategy set for each player and $U = \{u_1, u_2, u_3, ..., u_m\}$ is the utility or payoff set. Each player m, maximizes its utility u_m by taking available action a_m. Actions chosen by the other players are denoted as $\mathbf{a_{-m}}$. Together, the two actions $(a_m, \mathbf{a_{-m}})$ form an action tuple $\mathbf{a}$ which defines a strategy profile for a user m for different actions. Using different strategy profiles by player m, a steady state condition can be attain for game model G. This state is termed as the *Nash Equilibrium*. A Nash Equilibrium (NE) condition can be defined in terms of the best action $\bar{a}$ [94], player m takes in response to $\mathbf{a_{-m}}$, to maximize its utility u_m. This is given as:

$$\bar{a} \, \epsilon \, \{argmax \; u_m(a_i, \mathbf{a_{-m}})\} \tag{4.2}$$

In other words, a NE state for a player is defined by a strategy profile which takes into account all the players in the game and produces the best response subject to the condition that a unilateral deviation by any player will not be beneficial. Consider an example of a Multiple Access Game [95] between two players $p1$ and $p2$ accessing shared channel C. Each player has two possible actions Wait (W) and Access (A) as shown in the Table. 4.1. The simultaneous access to the channel result into possible collisions as shown and there is no payoff in return. The wait and access is the best possible strategy for both players in this case. The best possible return value for player $p1$ is $(1 - C)$ when player $p2$ is not accessing. Similarly the same strategy holds for player $p2$ as shown. A NE condition in this game for any player is expressed as,

$$u_m(a_m^*, a_{-m}^*) \geqslant u_m(a_m, a_{-m}^*), \; \forall a_m \epsilon \mathcal{A} \tag{4.3}$$

p1 \ p2	W	A
W	(0,0)	(0,1-C)
A	(1-C,0)	(-C,-C)

Table 4.1: Multiple Access Game in Normal-Game Form.

It should be noted that once the point of equilibrium is identified, it is not guaranteed that every player has achieved its desired or maximum outcome. This means that a game

can have multiple Nash Equilibria. The method for identifying the desired Nash Equilibrium in a game is called $Pareto - Optimality$. In this method different strategy profiles are compared. Formally this criteria is defined as:

$$u_m(a_m, a_{-m}) \geqslant u_m(a'_m, a'_{-m}) \tag{4.4}$$

which means that a strategy profile a_m is *Pareto-Optimal* if there does not exist any other strategy profile a'_m that increases player m utility without decreasing the utility of at least one player. There is a possibility to have several *Pareto-Optimal* strategies and the set of theses strategies is called the *Pareto Frontier*. The strategy profiles (A, W) and (W, A) in Table. 4.1 are Nash Equilibria as well as Pareto-Optimal. It should be noted that a *Pareto-Optimal* strategy may not necessarily be a Nash Equilibrium.

Game Theory framework is not limited to decision makers that play the game at the same level. A decision maker can be a policy maker representing an authority, e.g. in Telecommunications, a network operator or a service provider are good examples. In such cases, the decision maker has an objective, most likely overall profit, which he wants to optimize or maximize. There could be other common objectives among the network and its subscribers in terms of maximizing the provisioned QoS or minimizing the packet loss and useful resources (bandwidth). In Game Theory, such hierarchical relationship between a network operator and subscribers is modelled as Stackelberg-Leader-Follower problem [96].

From [97], if R is the revenue of the network that depends on some parameter a set by the network operator or manager and network subscriber's utility function is $u(a)$, in response to a, then the network objective function is defined as:

$$R(u(a), a) \tag{4.5}$$

The definition for NE in (4.3) can be redefined in the case of the Stackelberg problem. We consider that for some parameter a, player m has utility function $u^*_m(a)$, if there exists an equilibrium. The network manager, in this case, determines the value of a^* which maximizes R in (4.5). Thus, the Stackelberg Equilibrium (SE) state is defines as:

$$R(u^*_m(a^*), a^*) = \max_{a \epsilon \mathcal{A}} R(u^*_m(a), a) \tag{4.6}$$

This two level problem is useful for formulating optimization models as far as the decision maker (Leader) point of view is concerned. However, a lot of complexity is involved in terms of solving the users (Followers) problem to achieve NE. As mentioned earlier, it is not necessary that the achieved $u^*_m(a) \epsilon U^*(a)$ is optimal or unique (where $U^*(a)$ is the set of all achievable NE). To counter this problem the definition in (4.6) is redefined subject to the decision maker's objective. For different $u^*_m(a)$, the objective function R in (4.5) is

expressed depending on the network manager's interest. For example, in (4.7), minimizing $u_m^*(a)$ maximizes $R(u_m^*(a), a)$, while in (4.8), it is otherwise.

$$R(u_m^*(a^*), a^*) = \max_a \min_{u_m^*(a) \epsilon U^*(a)} R(u_m^*(a), a) \tag{4.7}$$

$$R(u_m^*(a^*), a^*) = \max_a \max_{u_m^*(a) \epsilon U^*(a)} R(u_m^*(a), a) \tag{4.8}$$

4.2 Reinforcement Learning

Reinforcement Learning (RL) is a research domain of machine learning which has been widely studied in the field of Computer Science. However, due to its wide scope of application in different domains, it has been applied and studied in various disciplines such as Game theory, Control theory, Genetics and Information theory. Consider an example of an agent in an environment. Any action taken by the agent results into some return value (a reward) from the environment. As a consequence, the agent learns by interacting with its environment and based on its experience (past rewards) it chooses or avoids some actions. This trial and error method is not only repeated to achieve a predefined value but to find some strategies or policies that maximize the reward.

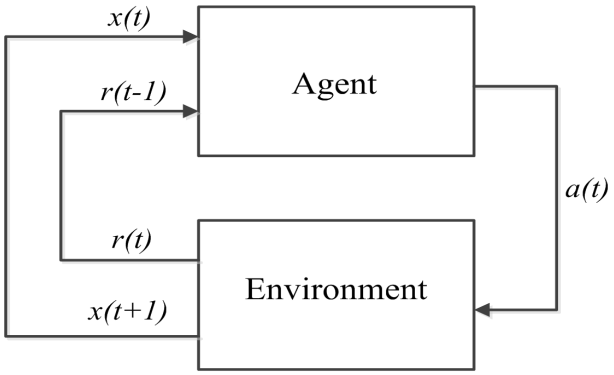


Figure 4.1: Reinforcement Learning Model.

A general RL model is shown in Figure. 4.1. The state of the environment is $x(t)$ at time t. It executes action $a(t)$ and receives reinforcement value $r(t)$. The state of the agent moves to $x(t+1)$. $r(t-1)$ defines all the past reinforcement measurements. The process continues till a terminal state is reached. The basic Reinforcement learning model based on the above discussion consists of:

1. A set S of environment states ,

2. An action set A,

3. Policies/Strategies to move from state x to state $(x+1)$,

4. Policies/Strategies to maximize future reward $r_a(t, t+1)$.

The mathematical framework for RL schemes is provided by Markov Decision Processes (MDPs) by modeling the scenarios where the output of a system is partly random and partly under the control of the agent. More precisely, MDPs are discrete time stochastic processes. At time t in some state $x(t)$, an agent chooses an action a from the action set A. As a result the system, the process or the environment move to a new state $x(t+1)$ while returning some reward $r_a(t, t+1)$.

As shown in Figure. 4.1, RL is a feed back process, the probability of a process to move to a new state depends upon chosen action a which is selected on the basis of the received reward r at $t-1$. This is known as the state transition probability $P_a(x(t), x(t+1))$ or simply $P_a(x, x')$. Mathematically it is represented as:

$$P_a(x, x') = Pr(x(t+1) = x' | x(t) = x, a_t = a) \tag{4.9}$$

The core objective of MDP is to find an optimal policy to maximize the return value. There exist methods that provide solutions for MDPs. Exploration methods with random selection of actions can result into poor performances. Even disregarding exploration mechanisms, the question remains: how to estimate which actions are good depending on past experiences? MDP problem is solved by taking into account some policy. Let π be the policy an agent follows to select action a_t in state x_t at time t:

$$a_t = \pi(x_t) \tag{4.10}$$

Algorithms that are used to select an optimal policy, values policy π with the expectation of the future reinforcement (return value or reward) with the learning rate (also known as discounted factor):

$$V^{\pi}(x) = E\{\sum_{t=0}^{\infty} \gamma^t r(x_t, \pi(x_t)) | x_0 = x\}, \forall x \epsilon \mathcal{X}, \tag{4.11}$$

where $\mathcal{X} = \{x_1, x_2, ...x_K\}$ is the set of finite number of states. (4.11) represents the expected reinforcement $E(r(x_t, \pi(x_t)))$ when the policy in (4.10) is applied. For simplicity, let $R(x, a) = E\{r(x, a)\}$. The above equation becomes:

$$V^{\pi}(x) = R(x, \pi(x)) + \gamma \sum_{x' \epsilon \mathcal{X}} P_{xx'}(\pi(x)) V^{\pi}(x') \tag{4.12}$$

where x' can be any state in set $\mathcal{X}$ other than x. $P_{xx'}$ is the transition probability to move from state x to state x'. The values are calculated by adopting different policies until a maximum value is achieved given as:

$$V^*(x) = \sup_{\pi} V^{\pi}(x) \tag{4.13}$$

and the policy which achieves these optimal values is also optimal (π^*) given as:

$$\pi^*(x) = arg \max_{a \epsilon A} \{R(x, \pi(x)) + \gamma \sum_{x' \epsilon \mathcal{X}} P_{xx'}(\pi(x)) V^*(x')\} \tag{4.14}$$

The value function calculation based on different policies can be achieved either by the *Policy Iteration Method* [98] or the *Value Iteration Method* [99].

4.2.1 Q-Learning

Based on the *Temporal Difference* (TD) [93] and the *Value Iteration* methods, *Q-learning* [100] has been developed as a RL scheme where transition probabilities and the reward values are unknown. In such a case, it is necessary to define the function Q that updates an array of information for the experience learned in a new state x' happened by applying an action a in state x. As a result of applying action a at time t, function Q updates its evaluation by taking into account: (a) the immediate reinforcement r_t and (b) the estimated value of the new state $V_t(x_t + 1)$ given as:

$$V_t(x_t + 1) = \max_{a' \epsilon A} Q(x_{t+1}, a') \tag{4.15}$$

The Q funtion is updated as:

$$Q(x_t, a_t) \longleftarrow Q(x_t, a_t) + \alpha r_t + \gamma V_t(x_{t+1}) - Q(x_t, a_t) \tag{4.16}$$

where α is a learning rate such that $0 \leqslant \alpha \leqslant 1$. The value of α determines the rate by which the newly acquired value can take over the old one. γ is the discount factor with values $0 \leqslant \gamma \leqslant 1$. Taking a value close to 0 for γ makes the agent consider the current rewards while a value close to 1 leads towards long-term high rewards. It should be noted that if an agent keeps on taking the action with the highest Q function value for a given state, it will probably end up to the local maximum. This is because it never tries all the available actions in the action set for all the states. This is termed as exploitation. However, such situation can be avoided by adopting an exploration strategy. In other words, in order to have the best possible expected reward, one should explore sufficiently all the possible actions in every state. This is time consuming and demands for trade-off [101], [102], [103], [104].

4.2.2 CODIPAS-RL

So far, the objective described for the RL scheme is to find an optimal strategy to maximize the expected return value. The learning algorithms based on partial information about the environment as well as past actions are defined as partially distributed learning algorithms. Cases where such information is not available demands for fully distributed learning algorithms which means how an optimal strategy can be learned if the only information available is agent's own return value (payoff,utility)? **C**ombined fully **D**istributed **PA**off and **S**trategy **R**einforcement **L**earning **(CODIPAS-RL)** [28], [105], [106], [107] is a learning scheme that updates the return value as well as the strategy function with the minimum information available.

Different from the learning schemes discussed before, CODIPAS-RL is flexible as it does not require past information about the agents in order to update their strategy profiles as well as the estimated payoff values. This distributed learning scheme can be represented in the following form:

$$\begin{aligned} x_{n_i,t+1} &= x_{n_i,t}(a_{n_i}) + r_{n_i,t}U_{n_i,t}(\mathbb{I}_{a_{n_i,t}=a_{n_i}} - x_{n_i,t}(a_{n_i})) \\ \hat{u}_{n_i,t+1}(a_{n_i}) &= \hat{u}_{n_i,t}(a_{n_i}) + q_{n_i,t}\mathbb{I}_{a_{n_i,t}=a_{n_i}}(U_{n_i,t} - \hat{u_{n_i,t}}) \end{aligned} \tag{4.17}$$

where the first equation in (4.17) is for the strategy profile update for agent n_i and the second equation defines the estimated future value of the expected return value (utility,payoff) for the action a taken by agent n_i at time t. $U_{n_i,t}$ is the returned utility value and $\mathbb{I}_{a_{n_i,t}}$ is an indicator function, indicating the use of action $a_{n_i,t}$ (being 1) by agent n_i at time t from all the available actions in action set A.

The use of learning schemes is to reach a state of equilibrium. In Game theory, they are used to find Nash Equilibrium. Players in a game, having incomplete information, can use CODIPAS-RL schemes both as heterogeneous learning or homogeneous learning. In heterogeneous learning, different learning schemes are used based on the requirement of the players while in homogeneous learning, same learning scheme is deployed. Based on the model in (4.17), different versions of the CODIPAS-RL methods have been proposed and studied [28], [105]. One of the reasons to device a distributed learning method like CODIPAS-RL is the dynamic nature of today's networks. In Wireless networks, where mobile nodes are moving randomly in different directions encourage to adopt CODIPAS-RL as it looks more real to the dynamic scenarios. A static network model with fixed demands and supply cannot involve the complexity of contemporary heterogeneous wireless networks. The ever changing environment of heterogeneous wireless networks require the study of dynamic behaviour of the nodes, their demands, system parameters and many other factors. CODIPAS-RL looks promising in the light of the above discussion. In a dynamic environment, a learning scheme that does not require any past measurements is

an optimal choice.

4.3 Problem Statement

A mobile node experiences varied channel conditions in wireless networks because of the factors like its geographical location, velocity, network delay, available bandwidth and SINR values. These factors greatly affect mobile node's throughput. A wireless network, connecting mobile nodes to the core network, holds overall load information. In case of congestion or route failure, network assistance for mobile nodes can avoid bad QoS experiences during mobility and handover situations. Thus, through mutual cooperation between network and mobile nodes, it is possible to design a mechanism which can ensure seamless handover with a minimum QoS degradation.

Based on the above discussion, we developed a scenario that represents the handover situation between two wireless access points. The area covered by an Access Point *(AP)* is divided into three power levels as defined in [108]. Each level represents a certain *RSS (Received Signal Strength)* threshold value. In Figure. 4.2, the three levels are (*(RX_Tresh * pr_lim)*, *(RX_Tresh)* and *(CS_Tresh)*) represented with their minimum and maximum threshold values.

The system has two wireless Access Points AP_1 and AP_2, two mobile nodes n_1 and n_2 and a Corresponding Node *CN* as shown in Figure. 4.3. We assume that n_1 is in session with *CN* and the coverage areas of both access points are not overlapping. The game involves three players (n_1, n_2 and AP_1). Though AP_2 is part of the scenario, its role is to provide assistance to n_1 when the handover reaches its final moments. Each player, individually, tries to maximize its utility. Here utility means the capacity of a service to satisfy a player's interests. In the current situation, n_1 and n_2 interests are to maximize the QoS they receive from the AP_1. In other words, their goal is to maximize their individual throughput. Network AP_1 interest lies in maximizing its revenue.

4.3.1 Algorithm

The algorithm proposed in [92] is explained here. It is composed of three phases that are classified into two processes called:

(a) Data Relaying Process

(b) Handover Process

This division of phases is necessary to avoid the ping-pong effect at the edge of the network during the handover. The algorithm is initialized when (λ,μ) values are advertised by the Access Point (AP_1).

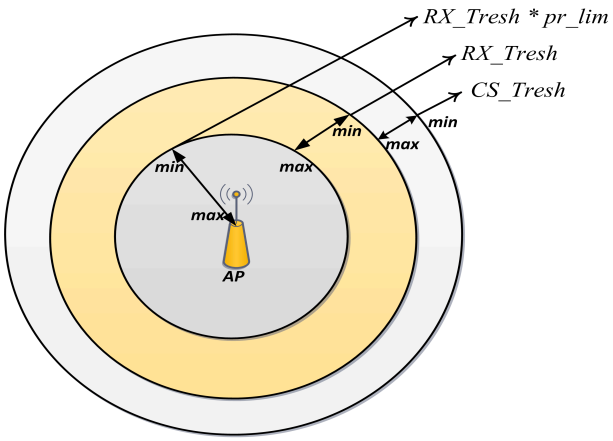


Figure 4.2: Power Levels in a Wireless Network.

4.3.1.1 Data Relaying Process

Phase I involves the *(RX_Tresh*pr_lim)* level where after time t, n_1 starts moving away from AP_1 as shown in Figure. 4.2. As a result, the *RSS* value decreases in *(RX_Tresh*pr_lim)* level where ($0 \leq pr\_lim \leq 1$) defines the minimum and maximum *RSS* values. This triggers the data relaying process. Node n_1 allocates **m** bits of its data stream to be relayed by a neighbouring node. It broadcasts data forwarding requests in the network. Mobile Node n_2, which has already updated itself with (λ,μ), acknowledges n_1 request by proposing **g** fraction of its channel to relay n_1*'s* **m** bits of data stream. n_1 upon receiving the reply, informs AP_1 and directs **m** bits of data stream towards n_2. The direct link between n_1 and AP_1 carries **(1-m)** bits of data and n_2 uses **(1-g)** fraction of its channel to stream its own data as shown in Figure. 4.3.

Phase II starts when the *RSS* becomes minimum in *(RX_Tresh)* level and the probability of packet loss between n_1 and AP_1 link increases. In such circumstances, n_1 finds it hard to traffic its data via direct link with AP_1. At this moment, the network announces new (λ,μ) values with increased incentive (μ) for relaying nodes (n_2 in our case). Following that, n_1 diverts most of its data stream, say $m_1 \mid m_1 > m$ to the link between n_1 and n_2. n_2 updates (λ,μ) information and allocates fraction of its channel, $g_1 | g_1 > k$, for n_1*'s* m_1 data stream. Once n_1 enters *(CS_Tresh)* level, n_2 starts relaying most of the n_1 data stream.

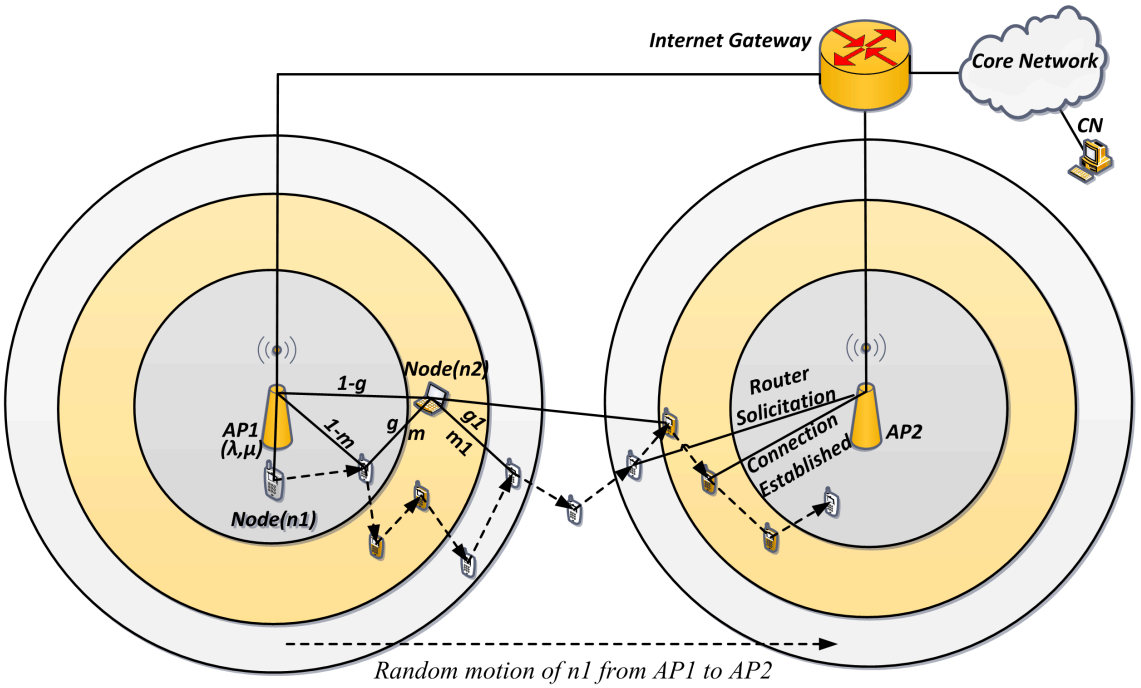


Figure 4.3: Data Relay and Handover.

4.3.1.2 Handover Process

Phase III starts when the *RSS* value in *(CS_Tresh)* level gets equal to zero. The link between n_1 and AP_1 break-offs and n_1 is completely relying on the connection it has with n_2. This critical phase lasts when eventually n_1 detects *(CS_Tresh)* level of AP_2. In $(RX\_Tresh)_{min}$ of AP_2, n_1 sends a router solicitation message to AP_2. AP_2 checks n_1 authenticity with the help of the Internet Gateway and sends back router advertisement. While accepting the request of stronger connection, n_1 requests AP_1 via n_2 to inform Internet Gateway to detour the *CN* traffic towards AP_2. Once the connection between n_1 and AP_2 has been established, n_1 terminates its link with n_2. The acknowledgement of route diversion is received by n_1 via AP_2.

4.3.2 System Model

Let $\mathcal{N}$ be the set of finite mobile nodes number in a wireless network. Each mobile node has a set of finite actions. For example, let consider a scenario with two mobile nodes (n_1,n_2) having $\mathcal{A}_1$ and $\mathcal{A}_2$ be the sets of finite action number respectively. $\mathcal{B}^n(t) \subseteq \mathcal{N}$ is the set of active mobile nodes. The active and non active states of a mobile node $n_i \mid i \in$

$\mathcal{N}$ are expressed as:

$$n_i = \begin{cases} 1 & \text{for} \quad i \epsilon \mathcal{B}^n(t) \\ 0 & \text{for} \quad Otherwise \end{cases} \tag{4.18}$$

At time t, set $\mathcal{S}_t$ defines the finite number of states of all mobile nodes in the network, e.g. $s_{n_i,j,t}$ is a state for any node n_i in the network, whereas $j = (1, 2, ...M)$ is one of the finite possible state at time t. $s_{n_i,j,t}$ of n_i represents its history composed of Strategies (x), Utilities (U), Actions (a), Bandwidth (b), Network delay (d), Velocity (v) and Location (l). At time t, $s_{n_i,j,t}$ is:

$$s_{n_i,j,t} = \Bigg((x_{n_i,0}, u_{n_i,0}, a_{n_i,0}, b_{n_i,0}, d_{n_i,0}, v_{n_i,0}, l_{n_i,0}), ..., \\ (x_{n_i,t-1}, u_{n_i,t-1}, a_{n_i,t-1}, b_{n_i,t-1}, d_{n_i,t-1}, v_{n_i,t-1}, l_{n_i,t-1}) \Bigg) \tag{4.19}$$

and

$$\mathcal{S}_t = (\{0, 1\} \times \prod_i s_{n_i,j})^t \tag{4.20}$$

where $\{0, 1\}$ is defined in (4.18) and $\prod_i s_{n_i,j}$ is the Cartesian product of states of all mobile nodes in the network at time t.

4.3.3 Utility Functions

As mentioned earlier, the utility is defined as the measure of satisfaction towards some service. In the present scenario, this definition is applied to the QoS experienced by users in terms of the service provided by their network. Various definitions of utility functions are proposed in the literature [109], [110], [111], [112]. The chosen utility function for our model is taken from [109]. This utility function is power dependent and is defined as the throughput achieved by a node for spending its own power, where throughput is the data in *bits* and power is measured in *joules*. Formally,

$$U_{n_i}(p_{n_i}) = \frac{T_{n_i}(p_{n_i})}{p_{n_i}} \; bits/joule \tag{4.21}$$

Throughput $T_{n_i}(p_{n_i})$ for node n_i is affected by the Signal-to-Interference and Noise-Ratio (SINR) $\gamma_{n_i} = \frac{p_{n_i} h}{W(I_{n_i} + N_o)}$, where h, W, I_{n_i} and N_o are path gain, bandwidth, interference and noise spectral density respectively. It is assumed that before each transmission, data bits are packed into M bit frames containing $L < M$ amount of information bits per frame. $M - L$ amount of bits are used for error detection at the receiver end. A frame M is retransmitted in case of any error in reception. From [109], the chosen throughput is expressed as:

$$T_{n_i}(p_{n_i}) = (L/M)Wf(\gamma_{n_i}(p_{n_i})) \tag{4.22}$$

where the efficiency function $f(\gamma)$ given by:

$$f(\gamma_{n_i}(p_{n_i})) = [1 - 2BER(\gamma_{n_i})]^M \tag{4.23}$$

is the modification of the frame success rate function:

$$f(\gamma_{n_i}(p_{n_i})) = [1 - BER(\gamma_{n_i})]^M \tag{4.24}$$

The inclusion of factor 2 in (4.23) is the approximation of (4.24) for the cases when $p_{n_i} = 0$, making $U_{n_i}(p_{n_i}) = \infty$. Utility functions for both nodes (n_1,n_2) are based on (4.22). The **Data Relaying Phase** is considered for the formulation of these functions. However, it should be noted that utility functions of both nodes (n_1 and $_2$) are depending on the (λ, μ) value pair generated by the network, more specifically by AP_1 in the considered case. Figure. 4.4 shows the picture of the game with all the players involved.

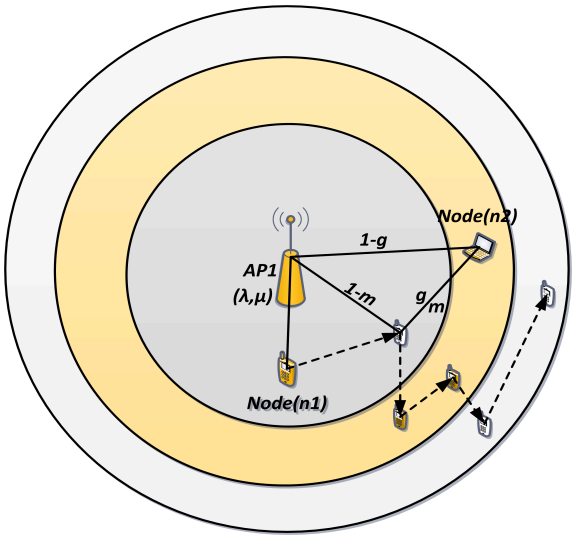


Figure 4.4: Utility Functions in 2-Level Stackelberg Game.

4.3.3.1 Access Point (AP_1)

The Access Point (AP_1), for any value (λ, μ), calculates its revenue. The charge for the service provided by AP_1 is the product of price λ and the service used by a node. Here, throughput T_{n_i} is considered as the service provided or used. It should not be confused with utility function U_{n_i} which is the satisfaction level of node n_i for the service provided. Similarly, two nodes in the same network, with the same amount of throughput and at

the same time, cannot have the same value for U_{n_i}. The radio resources provided by the network (AP_1) varies with time depending upon the channel conditions and the factors mentioned in (4.19).

The prime objective for the network is to maximize its revenue. All mobile nodes using network resources, pay some price (λ) for the provided service. For the case presented in Figure. 4.4, the total revenue calculation by a network also includes reimbursement factor μ. In this case, the reimbursement is defined as the product of reimbursement μ times throughput T_{n_r}, where n_r is a relay node. Net-revenue R_{net} is given as:

$$R_{net} = \sum_{all-nodes} \lambda(T_{n_i}) - \sum_{relaying-nodes} \mu(T_{n_r}) \tag{4.25}$$

By defining the revenue calculations in (4.25), net utility $U_{n_i}^{net}$ of node n_i is given by $U_{n_i} - \lambda T_{n_i}$. In the case of relay-node n_j, the net utility is $U_{n_j}^{net} - \lambda T_{n_j} + \mu T_{n_j}$. λT_{n_i} and μT_{n_j} are the charge and the reimbursement respectively. The reimbursement here should not be confused with any financial reimbursement. It is an incentive proposed by the network operator to the relay-nodes in terms of allocating them additional network resources (e.g, extra bandwidth or giving the priority to forward their data traffic, etc.) as a reward for relaying.

4.3.3.2 Node1 (n_1)

The **Data Relaying Process** explained in the Sec. 4.3.1 is taken into account to formulate the utility function of n_1. As discussed earlier, n_1 has allocated **m** fraction of its data stream for **g** fraction of n_2's bandwidth. Based on this division, the throughput for n_1 is the sum of the throughputs of two paths. The first path is the direct link between n_1 and AP_1 and the throughput of this link is denoted as $T_{n_1AP_1}$, where $T_{n_1n_2}$ is the throughput of the second link between n_1 and n_2. From the throughput definition in (4.22), $T_{n_1AP_1}$ and $T_{n_1n_2}$ are $T_{n_1AP_1}(p_{n_1,W_{n_1AP_1},h}) = W_{n_1AP_1} f(\gamma_{n_1AP_1}(p_{n_1}))$ and $T_{n_1n_2}(p_{n_1,W_{n_1n_2},h}) = W_{n_1n_2} f(\gamma_{n_1n_2}(p_{n_1}))$. Using these notations, n_1 pays the price which is the sum of the two links ($\lambda T_{n_1}AP_1 + \lambda T_{n_1}n_2$). The throughput of n_1 is the function of p_{n1}, $W_{n_1AP_1}$ and h, where h is the vector sum of all the path gains in the network. In order to maximize its utility, n_1 throughput $T_{n_1n_2}$ is dependent on n_2's throughput $T_{n_2AP_1}$. It should be noted that in a multi-hop network topology, the net throughput between two non-neighbouring nodes is the minimum of all links throughputs along the routes. In our case, the path between n_1 and AP_1 via n_2 has two links. The net throughput will be the minimum throughput value $min\{T_{n_1n_2}, T_{n_2AP_1}\}$ of the two links. In a non-cooperative game situation, the n_1 interest lies in maximizing its own utility and its optimization problem can be given as:

$$Node1(\lambda): maxU_{n_1}(p_{n_1}, m, h) \triangleq max \frac{T_{n_1AP_1}(p_{n_1,W_{n_1AP_1},h}) + min\{T_{n_1n_2}, T_{n_2AP_1}\}}{p_{n_1}} - \lambda[T_{n_1AP_1}(p_{n_1}, W_{n_1AP_1}) + T_{n_1n_2}(p_{n_1,W_{n_1n_2},h})] \quad (4.26)$$
$$subject\ to \begin{cases} 0 \leq p_{n1} \leq p^{max}, \\ W_{n_1AP_1} + W_{n_1n_2} = W \end{cases}$$

The utility of n_1 can be maximized under the constraints mentioned in (4.26). Since, the utility is power constrained, the problem imposed by $min\{T_{n_1n_2}, T_{n_2AP_1}\}$ can be solved by always maintaining the inequality between $T_{n_1n_2}$ and $T_{n_2AP_1}$ in such a way that $T_{n_1n_2} \leqslant T_{n_2AP_1}$. This can be done by decreasing W_{n1n2} and p_{n_1} slightly whenever $T_{n_1n_2} > T_{n_2AP_1}$. In this way, n_1 utility can be maximized by paying a less amount of service charges.

Based on the scenario presented in Figure.4.3, W is replaced by m and $1-m$. The total throughput of n_1 is the sum of throughputs between (n_1,AP_1) and (n_1, n_2), represented by $(1-m)f(\gamma_{n_1AP_1}(p_{n_1}))$ and $mf(\gamma_{n_1n_2}(p_{n1}))$ respectively. The objective function in (4.26) is subject to the two constraints, (a) the power constraint that defines an upper bound for n_1 power limit to avoid interference at AP_1 terminal and (b) the throughput constraint that limits $mf(\gamma_{n_1n_2}(p_{n_2}))$ to exceed $gf(\gamma_{n_2AP_1}(p_{n_2}))$. This limitation prohibits n_1 to pay the extra price (λ) for the throughput between n_1 and n_2.

$$Node1(\lambda): maxU_{n1}(p_{n1}, m, h) \triangleq max\big(\frac{1}{p_{n1}} - \lambda\big)\big[(1-m)f(\gamma_{n1a1}(p_{n1})) + mf(\gamma_{n1n2}(p_{n1}))$$
$$subject\ to \begin{cases} 0 \leq p_{n1} \leq p^{max}, 0 \leq m \leq 1\,, \\ mf(\gamma_{n1n2}(p_{n1})) \leq gf(\gamma_{n2a1}(p_{n2}))\,. \end{cases} \quad (4.27)$$

4.3.3.3 Node2 (n_2)

Node n_2 is the forwarding or relay-node. The throughput of the direct link between n_2 and AP_1 using (4.22) is given as $T_{n_2AP_1}(p_{n_2,W,h}) = Wf(\gamma_{n_2AP_1}(p_{n_2}))$. The service charges applied to n_2 by the network manager are $\lambda T_{n_2AP_1}$. The reimbursement that should be paid to n_2 by AP_1 for relaying n_1's data stream is given as $\mu T_{n_2AP_1}$. However, the reimbursement involves the $T_{n_1n_2}$ factor because of the path carrying the n_2 data is also shared by n_1's data. The reimbursement function is thus modified by the addition of this factor and becomes $\mu \min\{T_{n_2AP_1}, T_{n_1n_2}\}$. The optimization problem for n_2 is, therefore:

$$
\begin{aligned}
Node2(\lambda,\mu) : maxU_{n_2}(p_{n_2},g,h) \triangleq max & \frac{WT_{n_2AP_1}(p_{n_2})}{p_{n_2}} \\
& -\lambda T_{n_2AP_1}(p_{n_2}) + \mu min\{T_{n_2AP_1}, T_{n_1n_2}\} \\
& subject\ to \begin{cases} 0 \le p_{n_2} \le p^{max} \\ W \end{cases}
\end{aligned} \tag{4.28}
$$

For n_2 to maximize its utility, $min\{T_{n_2AP_1}, T_{n_1n_2}\}$ should be resolved. Factor $min\{T_{n_2AP_1}, T_{n_1n_2}\}$ describes the situation where the $T_{n_2AP_1} > T_{n_1n_2}$ condition decreases n_2 utility because of relaying more than required and eventually paying more service charges. Reducing W can increase n_2 utility whenever such situation arises.

Taking Figure. 4.3 and (4.28) into account, the objective function in (4.29) is subject to the power and throughput constraints like n_1. The power constraint inhibits the interference while the throughput constraint make sure that n_2 should not relay more than what n_1 has requested for. The fact is that n_2 will not be rewarded (μ) for any extra relaying. Hence, relaying more than requested will result in n_2 to pay an extra price (λ). The $T_{n_2AP_1} > T_{n_1n_2}$ situation is countered by tuning g for different value.

$$
\begin{aligned}
Node2(\lambda,\mu) : maxU_{n2}(p_{n2},g,h) \triangleq max & \Big[(1-g)\big(\frac{1}{p_{n2}} - \lambda\big) + g(\mu-\lambda)\Big] f(\gamma_{n2a1}(p_{n2})) \\
& subject\ to \begin{cases} 0 \le p_{n2} \le p^{max}, 0 \le g \le 1\,, \\ mf(\gamma_{n1n2}(p_{n1})) \ge gf(\gamma_{n2a1}(p_{n2}))\,. \end{cases}
\end{aligned} \tag{4.29}
$$

4.3.4 Nash-Stackelberg Model

Our system model with two mobile nodes, follows a 2-level Stackelberg model approach to reach an overall equilibrium stage. For different values of the (λ,μ) pair, the mobile nodes compete to maximize their utilities. If for any (λ,μ), nodes reach *Nash Equilibrium*, the strategies used are termed as equilibrium strategies. More precisely the strategies define a vector with equilibrium values of various parameters influencing the *Nash Equilibrium* state. In the current case, the equilibrium strategy vectors for nodes n_1 and n_2 (for any (λ,μ)) are given as: $\{P^*_{n1}(\lambda,\mu), g^*(\lambda,\mu), h^*(\lambda,\mu)\}$ and $\{P^*_{n2}(\lambda,\mu), m^*(\lambda,\mu), h^*(\lambda,\mu)\}$ respectively. However, it should be noted that there could be multiple *Nash Equilibria*. Hence, equilibrium strategy vectors should hold *Pareto Optimal* values of the considered parameters in order to become *Pareto Optimal Strategies* to reach the *Pareto-Superior Equilibrium*.

Access Point AP_1 at the higher level advertises different values for (λ,μ) pair. For every pair, nodes n_1 and n_2 follow AP_1 and reach an equilibrium state and return values of their

strategy parameters. This means that parameters (λ,μ), $\{P^*_{n1}, g^*, h^*\}$ and $\{P^*_{n2}, m^*, h^*\}$ for the two nodes (n_1 and n_2) on one hand and AP_1 on the other hand are inter-dependent. The optimization problem for AP_1 is maximizing its revenue. In terms of Stackelberg Leader-Followers model, AP_1, by using different (λ,μ) values, interacts with the nodes at the lower-level. A *Stackelberg Equilibrium* is said to be achieved when the network has succeeded to generate the maximum possible revenue. The equilibrium values for the pricing (λ) and reimbursement (μ) pair are given as (λ^*,μ^*).

The above explanation establishes a Nash-Stackelberg model for all the players in the considered game. In this model, players (mobile nodes) at the lower level reach the Nash Equilibrium which is followed by Stackelberg Equilibrium at upper level (network level).

4.3.4.1 Nash Equilibrium

Mobile nodes (n_1,n_2) compete to maximize their utility U_{n_i} for every (λ,μ) value through (4.27) and (4.29) respectively by using an optimal learning scheme (explained in Sec. 4.4). A strategy selected by such scheme achieves a Nash Equilibrium if no player has an incentive to deviate unilaterally from the point of equilibrium. Mathematically, for nodes n_1 and n_2, this can be expressed by the following inequalities:

$$U^*_{n1}[P^*_{n1}(\lambda,\mu), g^*(\lambda,\mu), h^*(\lambda,\mu)] \geq U_{n1}[P_{n1}(\lambda,\mu), g(\lambda,\mu), h(\lambda,\mu)] \tag{4.30}$$

and

$$U^*_{n2}[P^*_{n2}(\lambda,\mu), m^*(\lambda,\mu), h^*(\lambda,\mu)] \geq U_{n2}[P_{n2}(\lambda,\mu), m(\lambda,\mu), h(\lambda,\mu)] \tag{4.31}$$

4.3.4.2 Stackelberg Equilibrium

The leader (AP_1) in this Stackelberg game calculates its revenue using (4.25). For different values of (λ, μ), nodes n_1 and n_2 maximize their utilities until they reach the Nash Equilibrium Point (NEP). AP_1 receives the equilibrium values and maximizes its revenue as follows:

$$\max_{\lambda\geq 0,\mu\geq 0} R_{AP_1}[(\lambda,\mu), U^*_{n1}, U^*_{n2}] = \sum(U_{n1}, U_{n2}) - \sum(U_{n2}) \tag{4.32}$$

From (4.32), AP_1 which is the Stackelberg-Leader maximizes its revenue by tuning (λ, μ) values which in turn affects n_1 and n_2 (Stackelberg-Followers) utilities until the Stackelberg Equilibrium is achieved. The Stackelberg equilibrium is expressed as:

$$R^{SE} = arg \max_{\lambda,\mu} \max_{U_{n1},U_{n2}} R(U^*_{n1}, U^*_{n2}) \tag{4.33}$$

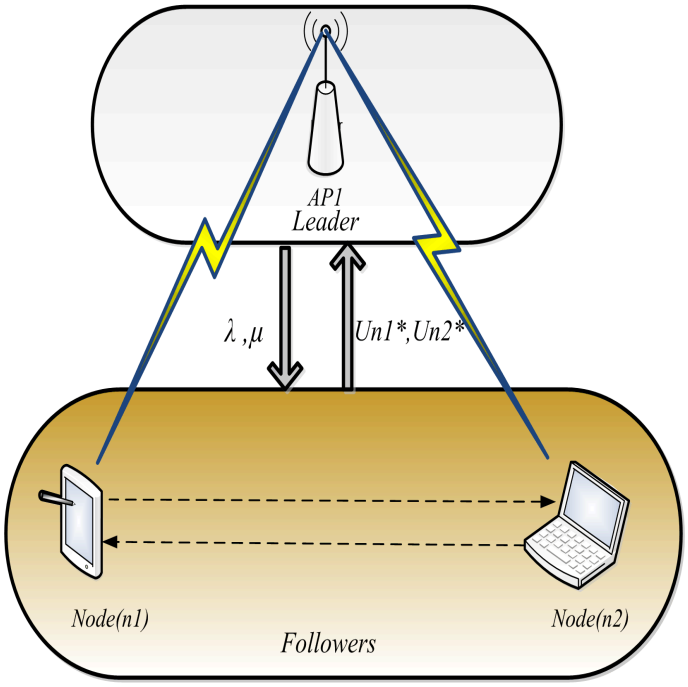


Figure 4.5: Nash-Stackelberg Leader-Followers Model.

4.4 Multiplicative Weighted-Imitative CODIPAS-RL Scheme

To reach the Stackelberg Equilibrium state in the considered game, the use of a learning scheme which can explore and exploit the interaction between the players is of the utmost importance. Multiplicative Weighted-Imitative CODIPAS-RL is one of the learning schemes proposed in [28]. A node in an active mode can use this scheme to update its strategy and estimate its utility using (4.34) and (4.35). (4.36) defines the internal clock of any node n_i with respect to network clock t as nodes are not always interacting and can either be in an active or a non-active state as defined in (4.18).

$$x_{n_i,t+1}(a_{n_i}) = x_{n_i,t}(a_{n_i}) + \mathbb{1}.x_{n_i,t}(a_{n_i}) \times \left(\frac{(1 + r_{\theta_{n_i}(t)})^{\hat{u}_{n_i,t}(a_{n_i})}}{\sum x_{n_i,t}(a'_{n_i})(1 + r_{\theta_{n_i}(t)})^{\hat{u}_{n_i,t}(a'_{n_i})}} - 1 \right) \tag{4.34}$$

$$\hat{u}_{n_i,t+1}(a_{n_i}) = \hat{u}_{n_i,t}(a_{n_i}) + q_{n_i,t}\mathbb{1}.(U_{n_i,t} - \hat{u}_{n_i,t}) \tag{4.35}$$

$$\theta_{n_i}(t+1) = \theta_{n_i}(t) + \mathbb{1}. \tag{4.36}$$

Variable x_{n_i} in (4.34) is used for the strategy selected by node n_i at time t to take action a_{n_i}. $x_{n_i,t+1}$ is the updated strategy at time $t+1$ for the same action a_{n_i} with an

increased or decreased probability. From (4.35), $\hat{u}_{n_i,t}(a_{n_i})$ is the estimated utility at time t for node n_i when action a_{n_i} is taken while $\hat{u}_{n_i,t+1}(a_{n_i})$ is the estimated utility at $t+1$. The observed utility $U_{n_i,t}$ is the value measured after an action is taken. The indicator function $\mathbb{I}$ corresponds to the active state of the node n_i.

The scheme is termed imitative as it imitates an action with a probability proportional to its previous use defined by the fraction part of (4.34). The learning rates $r_{n_i,t}$ and $q_{n_i,t}$ are assigned for the strategy update ($x_{n_i,t+1}$) and estimated utility ($\hat{u}_{n_i,t+1}(a_{n_i})$) functions respectively. The chosen values of the learning rates influence the learning behaviour of a node. Smaller values make the nodes learning at a slower rate, whereas using big values make them to converge at a faster rate towards the point of equilibrium. However, using moderate values of learning rates is advantageous as nodes can exploit as much as possible the available actions.

- At $t = 0$, node n_i chooses action $a_{n_{i,0}}$ and measures a noisy payoff that depends on the actions of the other nodes and state of the system. The utility estimation is then initialized to $\hat{u}_{n_{i,0}}$.

- At any time t, node n_i has an estimation of its utility. It chooses an action based on its experiences defined in (4.35) and experiments a new strategy. As a consequence, $U_{n_i,t}$ is measured based on which n_i updates the strategy function and estimates its utility using (4.34) and (4.35) for the next iteration.

The scenario considered in Sec. 4.3.1 follows the steps presented above. Initially, followers are unaware of each others presence. They discover each other when they interact. Unnecessary interaction is avoided to reduce the loss of battery life and signalling overhead. However, it is assumed that nodes in the network exhibit a selfish behaviour and restrain themselves from cooperating. In order to ensure a cooperative diversity and encouraging mobile nodes to show willingness to cooperate in the network, network operator announces incentives for cooperating nodes.

The objective for mobile nodes is to maximize their utilities by testing different strategies using a system of equations in (4.34) and (4.35), until a **NEP** is achieved. Using U^*_{n1} and U^*_{n2} values, the network maximizes its revenue as shown in Figure. 4.5. Considering **Phase I** in Sec 4.3.1, as soon as the interaction begins, different strategies are tested. Because of the random motion, it is evident that achieving a **NEP** is a hefty task. Both nodes attempt to maximize their utilities using (4.27) and (4.29) respecting the constraints. As discussed in Sec. 4.3.2, nodes n_1 and n_2 have action sets $\mathcal{A}_1$ and $\mathcal{A}_2$ respectively. These actions are tested by formulating different strategies using (4.34). The best strategy is defined as the one which selects the best available action producing the maximum return

value $U_{n_{i,t}}$. Since mobile nodes are always moving and most of the time in a random direction, it is difficult to select those strategies that maximize mobile nodes utilities as well as the network revenue. Another major factor is the RSS value which increases and decreases non-uniformly. Switching to different phases as described in Sec. 4.3.1 requires precision in making a decision. In order to tackle such situations, higher values of learning rates in (4.34) and (4.35) can be used to reach an equilibrium at an early stage to avoid data loss.

The action is imitated (repeated) with a higher probability for the next iteration at $t+1$ if at time t, the measured value of $U_{n_{i,t}}$ is better than its previous value at $t-1$. Otherwise the probability of imitating or repeating the same action is reduced.

4.5 Performance Evaluation

For the implementation, the settings have been kept simple. The action sets defined for both nodes have finite number of values. The actions selected for the chosen strategies for nodes n_1 and n_2 can either be the same or be different. The effect of choosing the same or different actions affect the overall performance of the proposed schemes. Learning rates in (4.34) and (4.35) are taken as $r = 0.1$ and $q = 0.6$ respectively. However, different values can be tried and tested depending upon various factors. For example, for a dynamic environment, where nodes are moving with a higher speed and in random directions, a higher values of r and q result in faster convergence of the learning scheme. Thus, the learning scheme can be used under various circumstances.

The case considered in our method is depicted in Figure. 4.3. It is assumed that, initially, mobile nodes n_1 and n_2 are unaware of the environment or the system. They carry limited information which include their velocity, RSS value, geographical location and battery life. Information about other nodes in the network is not known. The learning process is initialized by following the methods discussed in the previous section. It is assumed that before reaching a situation of data relaying or handover, mobile nodes have tried and tested different strategies to maximize their utilities.

4.5.1 Network Access and Interference Avoidance

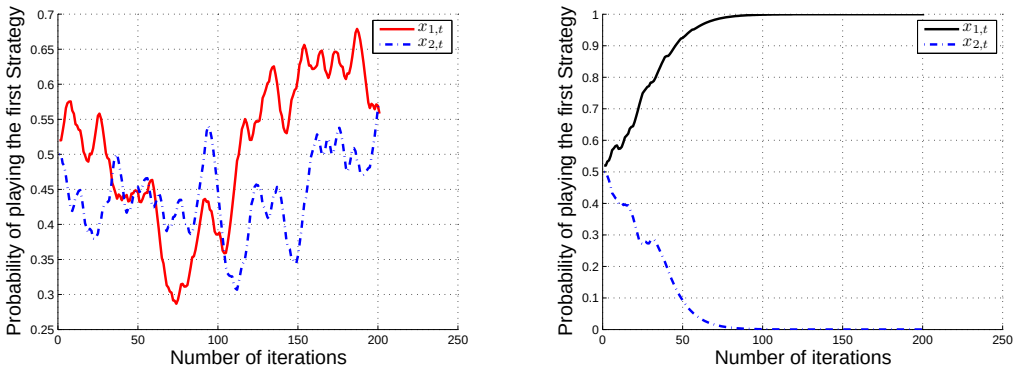
The assumption for selecting the actions from action sets $\mathcal{A}_1$ and $\mathcal{A}_2$ for nodes n_1 and n_2 respectively is based on the idea to avoid collisions or interference during transmission at the terminal point (AP_1). An action can be considered as a transmission of packets at a particular time slot or frequency. The interference can be avoided if two nodes do not transmit at the same time or frequency.

The scenario in Sec. 4.3.1, describes the procedure of two nodes interacting and learning by taking different actions and experimenting different strategies as mentioned in Sec. 4.4.

Both nodes access the network with two different strategies. The first strategy involves experimenting same set of actions, whereas in the second strategy, different set of actions are adopted. The results obtained show node behavior for both cases. The results basically represent the convergence to a state of equilibrium where nodes n_1 and n_2 achieve the most optimal utility values by following particular strategies $x_{1,t}$ and $x_{2,t}$ respectively. For selected strategies, different actions available in the action sets are tested.

Two different results are illustrated in the figures as (**a**) and (**b**). Result (**a**) represents the behaviour when same actions are selected whereas result (**b**) is for selecting different actions.

As mentioned earlier, both nodes access the network by taking some action. Accessing the network at the same time slot or different gives varied output. Figure.4.6 shows the probability of playing the first strategy by any of the two nodes. It is measured based on playing same and different actions respectively. Part (**a**) of the figure shows the result when same actions are selected and the outcome is not efficient because of the possible collisions when the same time slot or frequency is used by more than one node. Part (**b**) shows the convergence after few hundred iterations proving that the use of different time slots or frequencies results into an efficient outcome.

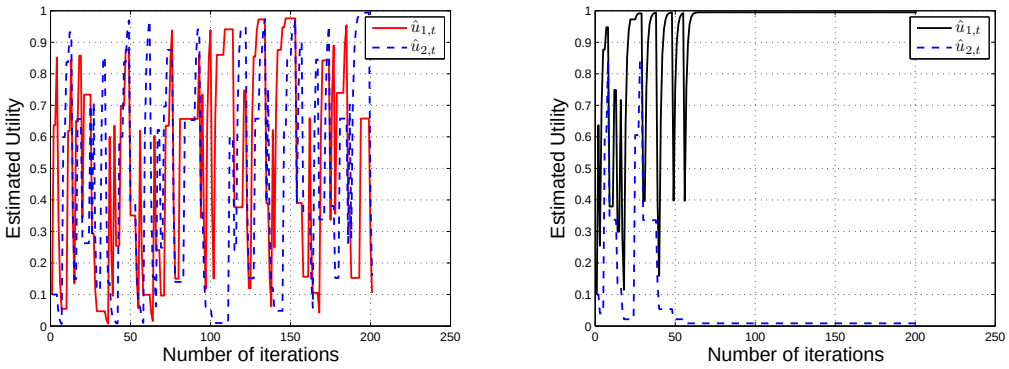


(a) same actions (b) different actions

Figure 4.6: Probability of playing the first strategy.

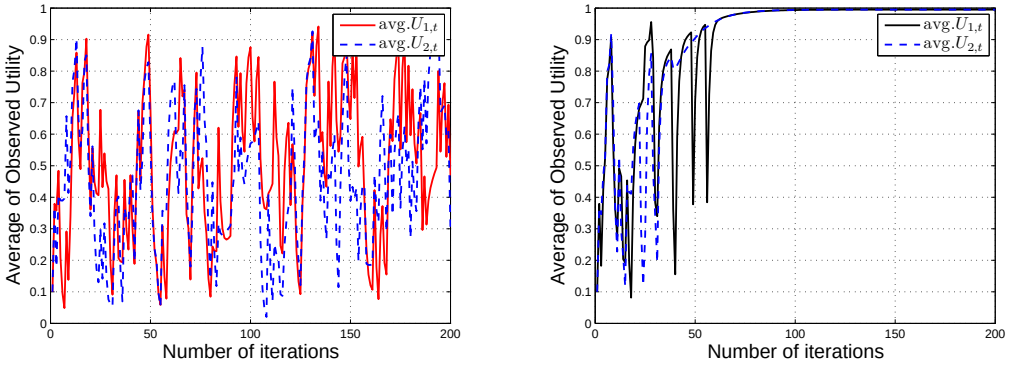
Figure. 4.7 shows the estimated utilities ($\hat{u}_{n_{1,t}}$,$\hat{u}_{n_{2,t}}$) for both nodes. Following the same set of actions result into perturbed estimated utilities unlike when different actions are chosen. Figure.4.8 represents the average of the observed utilities ($U_{1,t}$,$U_{2,t}$) with the set of results in which the collisions are resulted when the same actions are chosen whereas different actions lead to optimal results. The estimated values are used to estimate the possible observed utility values for both nodes. These values work as a reference to adopt future strategies by using (4.34). The estimated values of $\hat{u}_{n_{1,t}}$ and $\hat{u}_{n_{2,t}}$ encourage to repeat

the same strategy with a higher probability for future calculations. The effect of estimated utility function values in Figure. 4.7 can be seen on the observed utility in Figure. 4.8.



(a) same actions (b) different actions

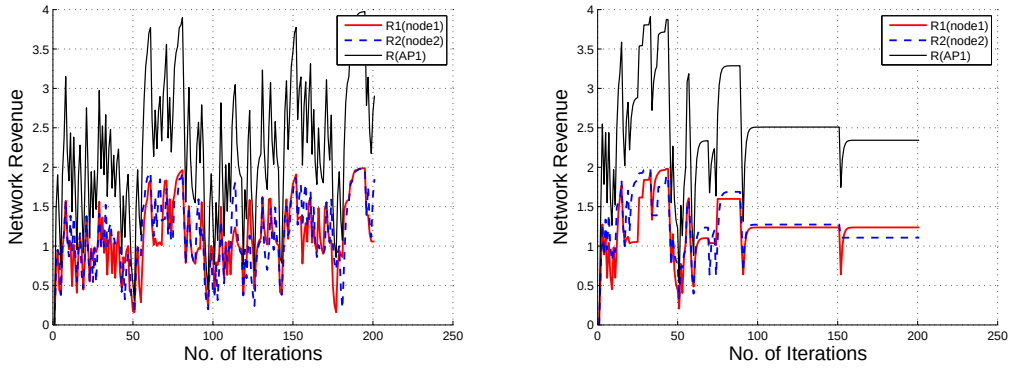
Figure 4.7: Estimated utility.



(a) same actions (b) different actions

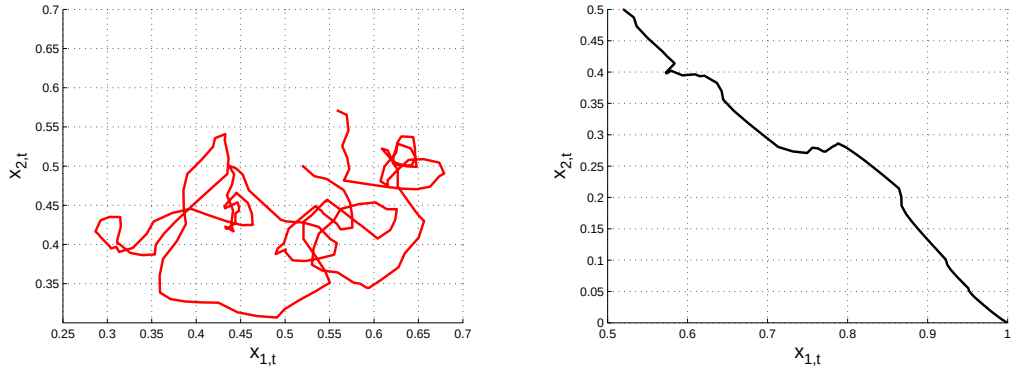
Figure 4.8: Average of the observed utility.

The leader of the game, the access point AP_1, calculates its overall revenue for each and every equilibrium state achieved at the lower level by nodes n_1 and n_2. There are two possible outcomes for the two sets of actions available mentioned earlier. In case of same set of actions, as a result of interference or collision during transmission, the overall throughput of nodes is greatly affected. This unstable situation destabilizes overall network revenue. It means that mobile nodes, as a result of their non-cooperating behavior, do not use network resources effectively. The experience learned by nodes because of interference and collisions, by adopting same set of actions, discourage cooperation. The result is that



(a) same actions (b) different actions

Figure 4.9: Network Revenue.



(a) same actions (b) different actions

Figure 4.10: Convergence to a Pure Global Optimum by Imitative CODISPAS-RL .

the overall calculated revenue is perturbed and never gets stable as shown in Figure. 4.9(a). The revenue calculation is directly proportional to the nodes behavior. Though, at some points there is a sudden increase in the network revenue but what matters here is a stable wireless network system which encourages all players to participate and cooperate. The unexpected outcomes are not favorable for long terms. In part (**a**), the same set of actions causes interference and collisions during transmission and the network never gets stable in terms of revenue generation.

The part (**b**) of the figure, on the other hand, shows stability in terms of revenue generation, as a result of collision and interference avoidance. It should be noted that the early stages of the result show instability. This is the initial learning phase where mobile nodes start to interact with their network environment. After a while, different strategies

are tried and tested and the one which returns optimal results is repeated with higher probability. Once mobile nodes reach at the point of equilibrium, network calculates its revenue. The process is repeated again and again to achieve the best possible outcome or payoff.

In order to show that the selected strategies $x_{1,t}$ and $x_{2,t}$ converge to a pure global optimum by using the Imitative CODIPAS-RL scheme. Figure. 4.10(a) illustrates the results for adopting the same set of actions. The strategies, $x_{1,t}$ and $x_{2,t}$, when adopt same set of actions, always arrive to a situation of instability. On the other hand, strategies with different set of actions proves to be stable and achieve global optimum as shown in Figure. 4.10(b).

4.6 Conclusion

This chapter has investigated the proposed Nash-Stackelberg Imitative CODIPAS-RL scheme in a wireless network scenario for data relaying and handover management. The algorithm in [92] was modeled mathematically with utility functions for mobile nodes and the wireless network. In order to achieve the Stackelberg Equilibrium, the Reinforcement Learning scheme was used to first achieve the Nash Equilibrium Point (NEP). The performance evaluation shows that the proposed algorithm performs well under the condition when different actions are chosen for the selected strategies. This eventually avoids chances of collisions during transmission and improves each nodes individual utility as well as results into the convergence to a pure global optimum of the selected strategies. As a consequence, the network maximizes its revenue under stable condition.

Chapter 5

Network of Information Email Service for the Future Internet

Contents

Users on the Internet always desire to experience the best Quality of Service (QoS). However, quality of experience (QoE) from users' perspective defines today's parameters for designing (new Web and Internet services) or redesigning (existing services). Among all the Web/Internet services available, Email is one of the widely used service.The basic function of this service is the transfer of text messages. Dedicated ports, servers and protocols are involved in this service which makes its installation and maintenance difficult. The situation becomes even more problematic when apart from these technical constraints, issues like spam emails, Internet viruses, privacy and security of users' emails (and email ID) and email spoofing further degrades the performance of this important service. These issues are still addressed in research communities.

This chapter presents a new email service based on the NetInf [113] architecture and uses the services NetInf provides in terms of routing, name resolution, storing/retrieving information and content distribution. The use of asymmetric key cryptography as a user ID is a unique feature introduced in the NetInf architecture. This ID, with no attached domain name, makes this server-less service scalable, secure and reliable. The major difference between the NetInf email service and the current one is its independence from dedicated servers, ports and protocols. Since each entity in NetInf is considered as an object with a unique ID, each email is considered as a separate object secured through asymmetric-key cryptography.

5.1 Motivation

This section addresses issues and problems related to the contemporary email service. Figure. 5.1 briefly explains the working principle.

The simple scenario explained here involves the steps required to send email message M from user $A@X$ (where A is the user name or identifier and X is the domain name) to user $B@Y$. Variations exist in the message format and the steps involved depending upon the email content, attachments (if any) and the security (if a security protocol is used). The details of the complete format of an email message is explained in [114] and [115]. Figure. 5.1 shows how an email message is forwarded from user $A@X$ to user $B@Y$ in the current Internet system. The steps followed are given below:

(a) User $A@X$ writes message M to user $B@Y$ using his/her Mail User Agent (MUA).

(b) The MUA of A forwards message M, using a submission protocol, to the Mail Submission Agent (MSA).

(c) The MSA resolves domain name Y of user B through DNS and receives an ID from the Mail Transfer Agent (MTA) for user B.

(d) MSA forwards the message to the MTA of B.

(e) MTA forwards the message to the Mail Delivery Agent (MDA) which further relays message M to user $B's$ inbox.

(f) User $B@Y$ retrieves message M using its MUA.

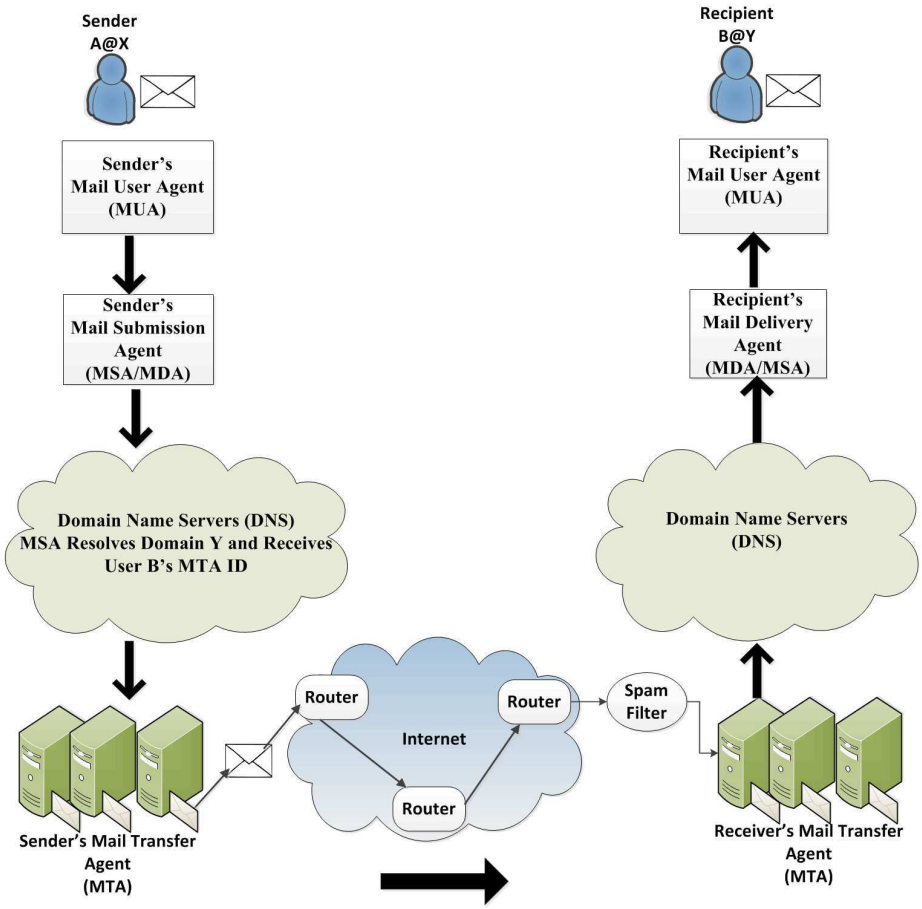


Figure 5.1: **Existing Email Service**

The email service today follows more or less the same procedure for sending and retrieving emails. Still, there are issues that degrade the overall performance of this whole system. For example, spam emails, computer worms, email spoofing, security and privacy (when the email address is used as a user identity for online services) are the issues that still affect the email service on the Internet. All these factors threaten the usefulness of email service.

5.1.1 Spamming and Network Worms

Spam email is one of the major problem email users deal on a daily basis. It can be defined as an undesirable number of emails that are circulating and are received over the Internet. Similarly, network worms, being independent in their nature, use emails as a way to replicate to vulnerable computers. Network worms mostly affect the network performance which is unlike computer viruses that target and harm files and folders.

5.1.2 Email Spoofing

Spoofed emails are actually fraudulent emails generated whenever the header information in an email is changed to make the message appearing as if received from a known sender. Such emails are often used to reveal the personal information of email users.

5.1.3 Privacy and Security

Privacy and security threats are major concern for email users. The possible reason for privacy breach is that most of the time messages are not encrypted. Similarly, an email has to pass through many intermediate levels (as mentioned earlier) to reach its final destination. This makes things easy for hackers to interfere and read/modify/delete messages. The use of an email address as an identifier for online services hosted on various web sites has raised privacy and security issues. Various online applications like Facebook, LinkedIn, Google Docs etc. demand users email addresses as identity. The irresponsible behaviour of the majority of users is also one of the reasons that introduced these security and privacy issues.

All these issues have been addressed and numerous possible solutions have been proposed. We are going to present an overview of these various efforts. In most cases, one problem has been addressed at a time. Solutions proposed in these research contributions either require architectural increments in the contemporary system architectures or software-based enhancements in the current services.

5.2 Related Work

The mechanism required to evaluate the reputation of an email system is known as the email reputation system. Basically, reputation systems [116] judge the high or low quality of the online content. With the success of the Internet, content creation and collaboration among millions of people around became quite easy. Earlier, a book, an encyclopedia or any other information used to have limited number of authors or creators. The Internet has made things easier. However, with the explosion of information and content available

online today, it is hard to justify which source of information is more reliable and authentic. Web sites like *Wikipedia* [117] are open collaborations but such systems are vulnerable. Reputation systems can be helpful in such a situation. Similarly, an email system can also be evaluated by a reputation system for improving its overall reputation in terms of spam and non-spam email filtering. In [118], a Collaboration-based Autonomous Reputation System for Email Service (**CARE**) have been proposed which is an autonomous email reputation system, based on the contemporary Internet architecture. Each domain, having an independent CARE system, works through inter-domain collaboration. Through cross examining the local email history database along with the email history databases of the collaborators, each domain rates remote domains either as spam or non-spam. This inter-domain cross-examination is done at a frequent rate. The evaluation showed the effectiveness of the proposed system. However, the overall increased overhead in terms of increased signalling as a result of frequent inter-domain cross-examination of email history databases can be problematic for such a system.

Context-based classification is another method to avoid spam emails. Instead of using statistical approaches, which take into account the frequency of the occurrence of a word in a document, context-based classification takes into account how the occurrence of one word will affect another word in the document. The statistical approaches for spam and non-spam email classification have strength as well as weaknesses, as discussed in [119]. The survey conducted between different statistical approaches for email filtering is summarized along with their performances. The current deployment of e-mail filtering on the web is mostly based on statistical approaches on text classification. However, the context-based approach is new and there is a wide scope to work on it in the near future. Using a context-based text classification can further improve spam filtering. The accuracy reported for statistical approaches is more than 90%. However, with the semantic description of the content on the web, it will become possible for machines to do the tedious task of differentiating between the right and wrong information. This means that a word having the same meaning but in a different context requires context classification methods. Such classification can be of worth examining for an application-like email in terms of spam filtering.

The use of public-key cryptography to provide security in email services is not a new concept. The use of Public Key Infrastructure (**PKI**) as a security framework for email systems is a set of hardware, policies, people, software, etc. This arrangement binds public keys with user identities by means of a Certificate Authority (**CA**). In Public-Key Cryptography, two separate keys are used, one is used to encrypt the data (email in this case) and the other to decrypt it. One key is defined as the Public Key while the other is called the Private Key. Normally, the public key is used for encrypting and the private one to unlock. Email systems like S/MIME (Secure/Multipurpose Internet Mail Extension) [120]

is one of the PKI-based security framework for content protection. However, it is not economical because of the high cost of the public-key certificate management infrastructure. PGP (Pretty good Privacy) [121] is also used for email protection. In [122], an identity-based solution is proposed that does not require public key management infrastructure. It is secure and cost effective. In [123], the identified drawback in PGP mails and S/MIME is that the email header is unauthentic which can encourage email spoofing. The proposed solution is based on XML technology. Its flexibility and its portability through web services gives it an edge over the existing system. The service is known as XML email and combines existing popular standards plus novel features in the area of XML and Web Services Security, making this approach efficient and more secure.

The above discussion that the problems faced by the email service can be mitigated by implementing various solutions proposed by the research community. It should be noted that most of the proposed solutions have targeted one or at most two issues. Some solutions are portable and are adaptive to any system environments and are backward compatible. Others require additional infrastructure, as in the case of reputation systems. The use of cryptography is also handy but encrypting the entire message every time is relatively expensive. All these methods are effective and have been implemented at some scale. Most of the time, some sort of subscription fee has to be paid to use these services. However, the point of argument which motivates the proposal of the NetInf email service is to incorporate all these solutions to one system. The NetInf Email architecture includes all the features required by an email service as discussed above.

5.3 Network of Information Email Architecture Framework

The use of the mobile Internet has increased overwhelmingly because of the availability of Smart phones and applications developed for such devices. This unprecedent increase of the Internet usage over mobile phones has degraded the performance of different services in terms of overall QoS, QoE, privacy and security. In the case of the Email service, available solutions mitigate these problems temporarily because all the provided services are compatible with the current TCP/IP-based Internet architecture. This Internet architecture has grown over the years unexpectedly and expecting that all new services will improve its overall efficiency is not possible. The contemporary Internet has different domains for different geographical regions. Thus, each domain has different requirements and problems that cannot be addressed by a single proposed solution. A policy implementable in one region can be unpractical in another region. A new architecture with a clean-slate approach, providing almost all possible solutions at all levels looks more promising in such a situation. The Network of Information (NetInf) has all these features.

5.3.1 NetInf Email Information Structure

As far as information is concerned on the Internet, it is not limited to simple text messages anymore. The complexity has become more visible after the four-decade-old client-server model failed to exploit different attributes related to an information. The NetInf Email service is based on the Network of Information architecture [113]. This service considers every object, either virtual (e.g. text files, songs, videos, etc.) or physical (e.g. monuments, buildings, man, machine, etc.) as an information, and defines them as an Information Object (IO) as shown in Figure. 5.2. The information object model of the NetInf Email service represents the hierarchy of the different types of objects. The higher level of this hierarchy is the generic representation of Information Objects IO. They are further classified into Semantic Information Objects (IO_s) (providing semantic details of an IO from the user perspective) and Management Information Objects (IO_m) (describing an IO from the NetInf point of view). The lower level of the hierarchy of the object model represents the information in terms of streams of bits and are defined as Bit-level Object (BO).

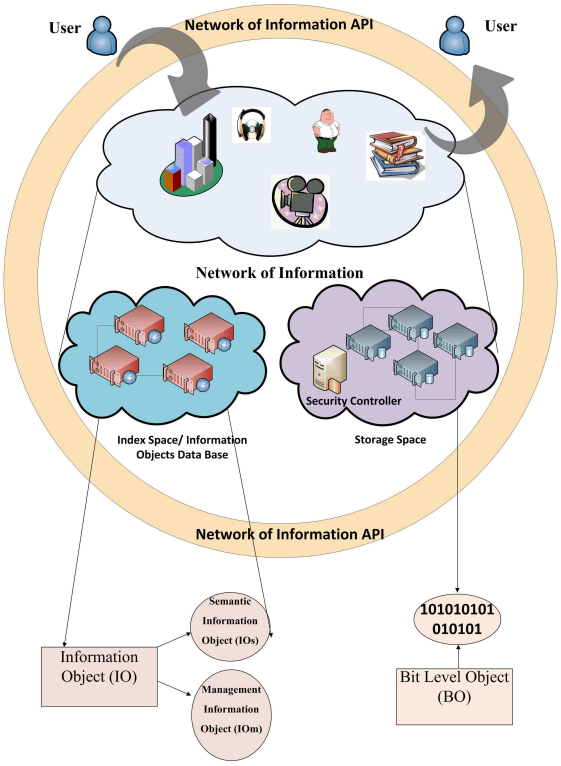


Figure 5.2: **Network of Information Email Service.**

5.3.1.1 Metalist Model for NetInf Email Service

The Metalist model in [36] explains *IO* management for the NetInf Email service. This model serves different aspects by representing the metadata related to *BO*. For example, tagging objects with attributes makes the semantic search easier for the end user. However, in order to reduce the metadata redundancy, existing metalist can be used in this model. The indirection method is used in cases where a metalist with a different format can be included as an external metadata in the NetInf Metalist model. Information (email in this case) in the NetInf Email service are managed by users and the NetInf Email architecture mutually. Users publish *IOs* with *Access Rights* (i.e. right to read, delete or modify the content of an information) defined and attached to it while the NetInf Email architecture manages them through information dissemination within the network (the detail of the NetInf Email structure and working principle is discussed in Sections 5.5).

Let's take an example for searching an information (email) through the NetInf Email service architecture. A user connected to NetInf looks for the relevant information by feeding semantic information through an interface. The NetInf Email system retrieves the result in the form of metadata tagged with the semantic information provided by the end user. Using this information, the required information can be accessed in NetInf. Here, the provided semantic information is IO_s. The metadata provided by the NetInf system is IO_m and the retrieved information is a set of *BO*.

5.3.1.2 Security and Privacy in the NetInf Email Service

Since privacy and security of information in NetInf are one of the prime objectives, the proposed Security Controller [124] for the NetInf architecture, facilitates any kind of secure information transfer as shown in Figure. 5.2. The owner of an information (email) has the authority to define the *Access rights* for his publishing contents. *Access Rights* define what rights receiving users have to access the content of the information. Generally, read only right is granted. The content of the information declared protected by its owner requires security check through Security Controller to access it. The user accessing encrypted content is challenged by the Security Controller to undergo verification tests. A simple example in this case can be considered for a user (email sender) who possesses a public key. This public key is distributed by its owner (email receiver) who owns this Public-Private key pair. The sender using this public key, encrypts his email and sends it via NetInf Email service. The receiver, also using the NetInf Email service, is challenged by the Security Controller for verification when accessing his emails. By providing the valid Private key, he validates the verification process and accesses the required information (emails).

5.4 NetInf Email Architecture Framework Description

The NetInf Email architecture framework has the following distinct components (illustrated in Figure. 5.2).

The **Storage Space** is the part of NetInf Email architecture where all the digital information i.e. *IOs* and *BOs* are stored. *IOs* are stored in the Index Space as well but in a different format. The Storage Space stores information as bit streams. The Security Controller [124] is located in the Storage Space taking care of accesses to both kinds of data objects. The Storage Space can be considered as a set of data centers. Information (email) once published in the NetInf Email system are managed by NetInf and may be replicated at different data centers depending on the user (receiver) network accessing activities. In NetInf, information are cached in data centers that come along the routing path adopted for delivery. This helps in data dissemination and reduction of the overall latency.

The **Index Space** indexes the IO_s. The semantic details of *BOs* are pushed in the Index Space by the users. The Metalist model is implemented in the Index space of NetInf. However, the semantic details can be indexed using other mechanisms, e.g. a RDF framework. This shows the flexibility of the NetInf Email architecture to quickly assimilate any change and to implement and apply it immediately. In NetInf, the XML format is used for tagging metadata. Hence, the Index Space maintains an XML-based database of IO.

The **NetInf API** was developed for NetInf and is based on the REST architectural style [125]. It is the only part through which users interact with NetInf. It provides basic functions to publish, retrieve, delete (and other functions) *IOs* using *Push*, *Put*, *Get*, *Publish*, *Search* and *Remove* operational commands. For example:

Push is used to store *BOs* in the storage part. Once *BOs* are stored, IDs generated by the Storage Space results in the creation of IO_m by NetInf which manages all these IDs. Upon storing this IO_m in the Storage Space, the generated ID is forwarded to the user.

Search is used for extracting the metalist of the requested information. In this case, the metalist is the list of addressed emails.

Get uses the same ID to retrieve IO_m. The contents of IO_m represents the IDs for *BOs* stored in the Storage Space. Using these IDs, the required information is extracted from the Storage Space.

Remove command can be used on both Storage and Index spaces for the removal of an object.

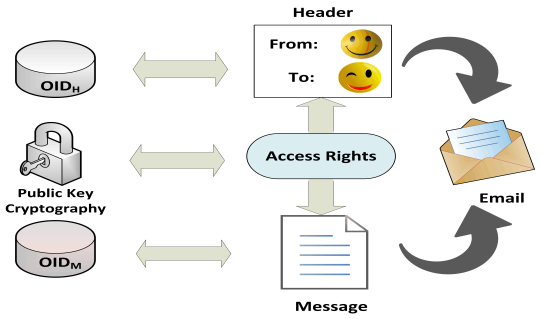


Figure 5.3: **NetInf Email Message Format.**

5.5 Network of Information Email Service

5.5.1 NetInf Email Message Format

Figure. 5.3 shows the NetInf Email message format. The header and the message parts are considered as Information objects. In the NetInf architecture, every object, virtual or physical, has a unique ID. OID_H and OID_M are the header ID and message ID respectively. The header part carries information about the sender, the recipient and the time-stamp. The message part contains the content or information for the email recipient. *Access Rights* attached to both objects define the level of permission granted by the sender to the recipient for accessing the email. The NetInf Email service architecture facilitates the management of all the components of the NetInf Email message which is explained in the next section.

5.5.2 NetInf Email Working Scenario

Here the scenario explains the end-to-end procedures involved in sending and receiving emails in NetInf. In summary, these procedures indulge all the three components of the architecture to be the part in the process of email exchange. From an end user to the storage space of NetInf, the data flow is ensured to be secured by the grace of asymmetric key cryptography. The NetInf interface provides simple commands like *Push*, *Put* and *Publish* for information insertion within the system. The retrieval involves a phase in which the user is challenged to prove his authenticity using his unique private key. The procedures, both at the sender and the receiver ends, are two-phase processes. The list of notations used during these processes is presented in Table 5.1.

M	Email Message
$\bar{M}$	Encrypted Email Message
H	Message Header
$\bar{H}$	Encrypted Message Header
PK	Asymmetric Public Key
$\bar{PK}$	Asymmetric Private Key
IO_{S}	Information Object(s)
IO_{m}	Management Information Object
IO_{m}^{M}	Management Information Object for Message M
IO_{s}	Information Object with Semantic Details
IO_{m}^{H}	Management Information Object for Message Header H
$AR\ \|_{D}^{R}$	Access Rights for Reading and Deleting
ID	Storage Space Generated Identifier
OID_{H}	Header Object Identifier
OID_{M}	Message Object Identifier

Table 5.1: Notations.

5.5.2.1 Sender End

In Figure. 5.4, the sending function format of an email is represented as $send(PK, H, M)$.

Phase #I
The message part ***(M)*** of ***send (PK,H,M)*** is processed.

(a) The email message consists of three fields: (i) The header ***(H)*** holds the details mentioned earlier; (ii) the receiver's Public Key ***(PK)*** and (iii) the content of the message ***(M)***.

(b) The Mail User Agent (MUA) encrypts the message ***(M)*** with the receiver's public key ***(PK)*** and *Access Rights* (Read, Delete) $\boldsymbol{(AR\ |_{D}^{R})}$ are attached to it. The encrypted message $\boldsymbol{(\bar{M})}$ along with *Access Rights* $\boldsymbol{(AR\ |_{D}^{R})}$ are pushed into NetInf through the NetInf API.

$$\boldsymbol{Push\ (\bar{M} = M \oplus PK\ ,\ AR\ |_{D}^{R})}$$

(c) Within the NetInf API, ***($\bar{M}$)*** is stored using the ***Put*** function. Each new object is acknowledged by NetInf which sends back a unique identifier ***(ID)*** generated by the Storage Space. This results in creating an IO_m^M managing the generated ***ID*** for ***($\bar{M}$)***.

(d) Through the ***Put (IO_m^M)*** operation within the Storage Space, NetInf returns IO_m^M ***ID***

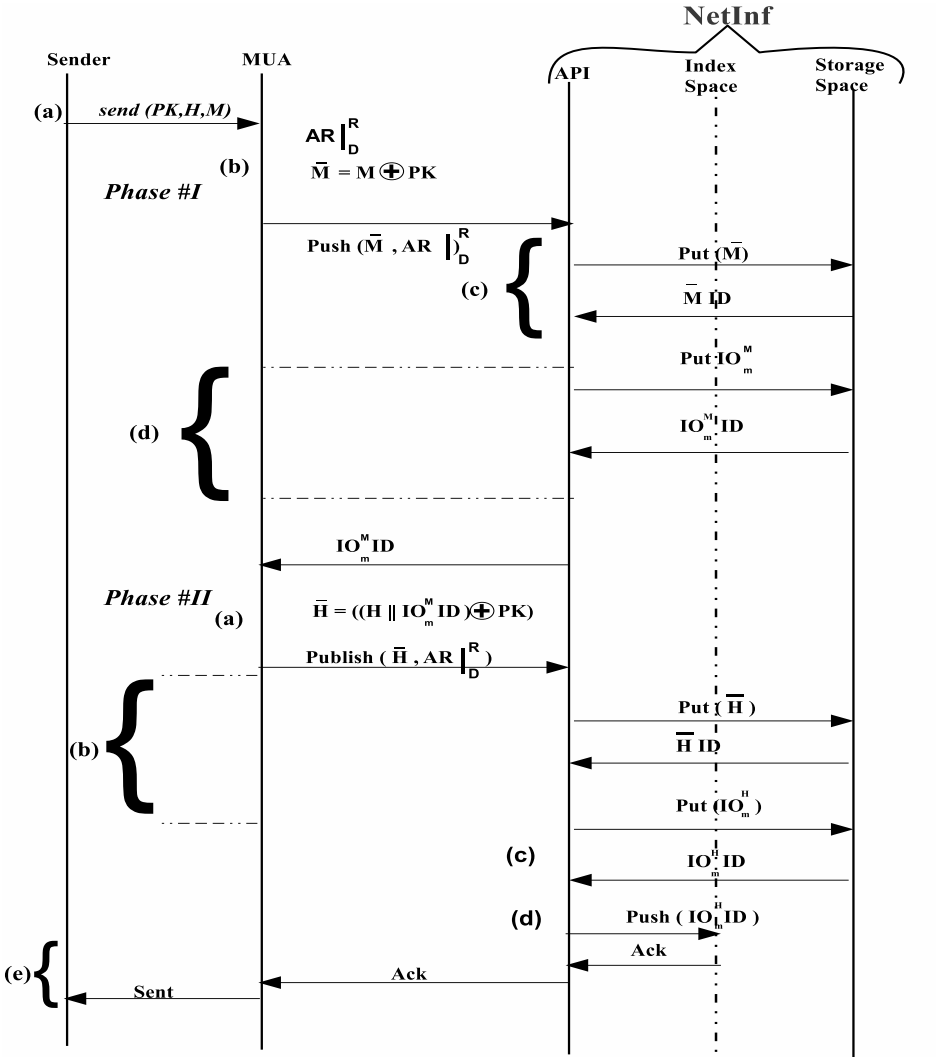


Figure 5.4: **NetInf Email Sending Procedure**

Phase #II

The Header ***(H)*** part of ***send(PK,H,M)*** is treated by NetInf.

(a) The header ***(H)*** is concatenated with $\boldsymbol{IO}\ {}^{M}_{m}\boldsymbol{ID}$ and is encrypted using ***PK*** as follows:

$$\bar{\boldsymbol{H}} = \boldsymbol{(H}\ ||\ \boldsymbol{IO}^{M}_{m}\ \boldsymbol{ID)} \bigoplus \boldsymbol{PK}$$

(b) The encrypted Header $\boldsymbol{(\bar{H})}$ along with $\boldsymbol{AR}\ |^{R}_{D}$ are published in NetInf. NetInf stores the encrypted header $\boldsymbol{(\bar{H})}$ in the Storage Space. For the management of the generated ***ID***, NetInf creates an $\boldsymbol{IO}^{H}_{m}$.

(c) The $\boldsymbol{Put\ (IO}^{H}_{m}\boldsymbol{)}$ operation returns the ***ID*** for $\boldsymbol{IO}^{H}_{m}$.

(d) $\boldsymbol{Push\ (IO}^{H}_{m}\boldsymbol{ID)}$ is the step where the whole email with all the contents is published in the Index space.

(e) At the end, an acknowledgement is received by the user at the sender end that the message has been sent successfully.

5.5.2.2 Receiver End

The receiver end user performs a request to identify the list of messages he has received, i.e. the list of objects that have been encrypted with his ***PK***. At the receiver end, an email message is retrieved in two phases as shown in Figure. 5.5. The first phase recovers the list of metalists of Semantic Information Objects $\boldsymbol{(IO_s)}$ with their ***(IDs)***. The $\boldsymbol{IO_s}$ is equivalent to $\boldsymbol{IO}^{H}_{m}$ as it recovers the header ***(H)*** of the email. The second phase allows to extract the message from the Header ***(H)*** depending upon the *Access Rights* ***(AR)***.

Phase #I

(a) The process is initiated when the user at the receiver end uses its MUA to fetch new messages. The MUA initializes the search for the emails addressed with the receiver's public key ***(PK)*** using the ***search (PK)*** function.

(b) NetInf relays this request on behalf of the MUA to the Index Space. The request is replied with the list of metalists of $\boldsymbol{IO_s}$ with the identifiers $\boldsymbol{(IO_sID)}$.

(c) Using the $\boldsymbol{Get(IO_sID)}$ command, the Storage Space is accessed to retrieve a specific message. The Security Controller challenges this access by testing users' credibility. If the user proves to be the actual owner of the public key ***(PK)*** by providing his $\bar{\boldsymbol{P}}\boldsymbol{K}$, the Security Controller grants access to the effective object. Otherwise, the access is denied.

(d) Using private key $\bar{\boldsymbol{P}}\boldsymbol{K}$, the user extracts $\boldsymbol{H}\ ||\boldsymbol{IO}^{M}_{m}\boldsymbol{ID}$

$$IO_s \oplus \bar{PK} \rightarrow H \,||IO_m^M ID$$

(e) Once decrypted, the user can access the Header ***(H)*** contents.

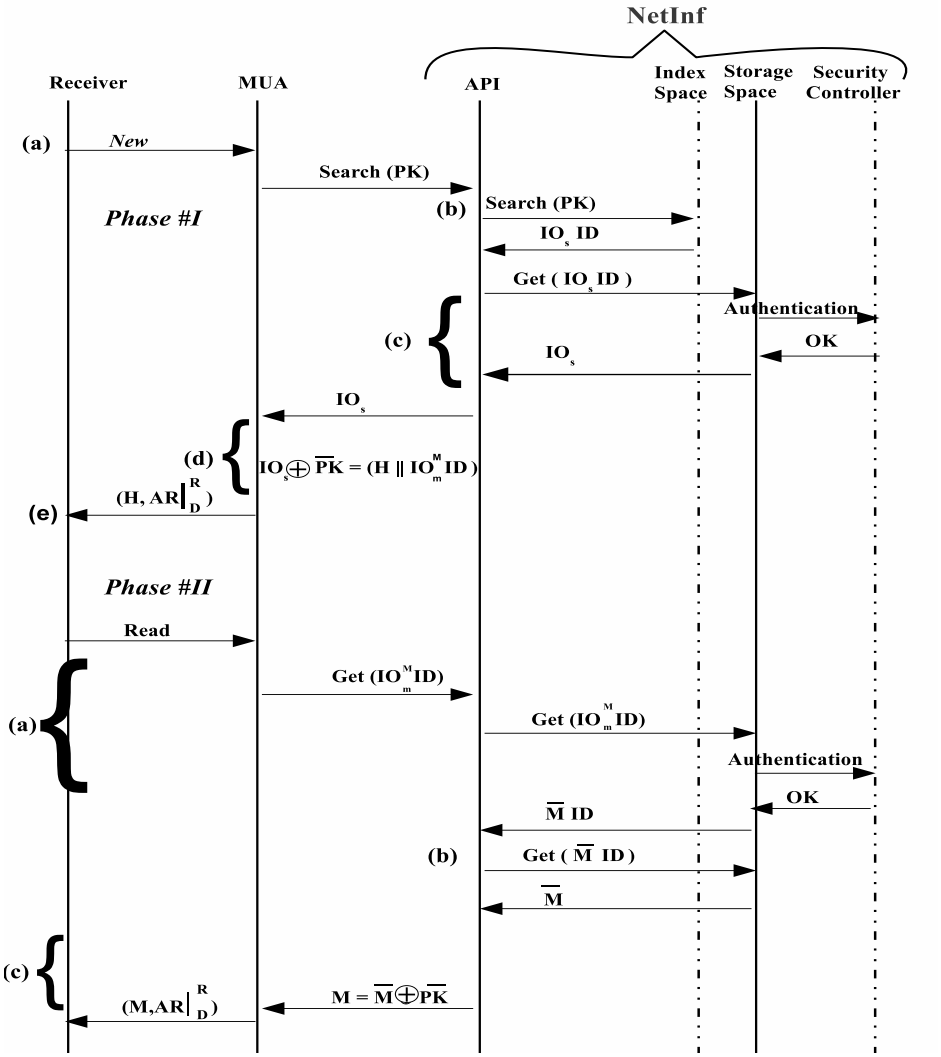


Figure 5.5: **NetInf Email Receiving Procedure.**

Phase #II

(a) In order to read the contents of the message ***M***, the user invokes the MUA to extract the bit-level object from the Storage Space. The ***Get(IO_m^M ID)*** command makes this happen by accessing the Storage Space. Here again, the Security Controller checks the user's authentication and returns $\bar{M}$ ***ID*** .

(b) With the ***Get($\bar{M}$ ID)*** operation, the encrypted $\bar{\boldsymbol{M}}$ is extracted.

(c) Finally, encrypted message $\bar{\boldsymbol{M}}$ becomes readable message $\boldsymbol{M}$ when decrypted using the receiver's private key $\bar{\boldsymbol{PK}}$.

$$\bar{\boldsymbol{M}} \bigoplus \bar{\boldsymbol{PK}} \rightarrow \boldsymbol{M}$$

5.6 Qualitative Evaluation

For the evaluation, we collected a sample of $3 * 10^5$ emails which were received between years 2006 and 2010. The size of emails varies, ranging from 100 bytes to 50 MB. Figure. 5.6 presents the distribution of the number of emails together with their cumulative distribution. It clearly shows that the size of emails follows the Gaussian law centered around 3 KB where 90% of emails have a size smaller than or equal to 30 KB. This means that most of the emails have a small size on average, i.e. they occupy small memory space. This qualitative analysis is an estimation to the possible outcome that will effect the size of the emails after encrypting emails and before sending. We estimate that since most of the emails being exchanged on the Internet have, on average, a small size, the performance of an email service in terms of latency will not be much affected. The performance criteria selected in this work is primarily privacy and security of users' identity and email content. The possible increase in the overall overhead that will be added after encrypting emails, by using the asymmetric keys, would be minimal.

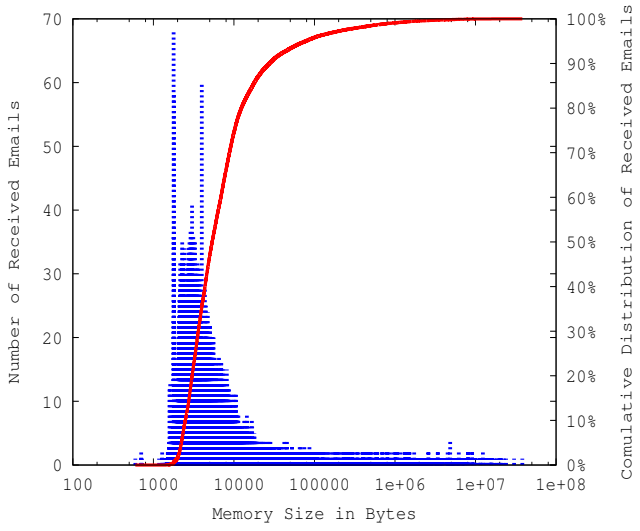


Figure 5.6: **Number of Received Emails with Different Memory Sizes.**

However, still in Figure. 5.6, there are a number of emails that require a large amount of memory space. This increase in the size can result into an overall system latency while sending or fetching an email. The NetInf architecture addresses this problem with better content distribution mechanism. Therefore, we envisage a trade-off between these two factors that are the system latency (which should be minimal) on one hand, and security and privacy of the content on the other hand.

5.7 Conclusion and Future Work

This chapter presented an innovative NetInf-based email service. The objective of this work is to introduce a novel email service. The unique features of this new service support the idea of having a network with no dedicated ports and servers. The approach used is based on the Future-Internet architecture known as the Network of Information. This email service works on top of the NetInf architecture. The privacy and the security of the email content are ensured by using asymmetric keys. This idea is not new for securing information data, but the use of public/private keys as user ID is a new concept in the context of NetInf. The NetInf Email format defines different components of the NetInf Email message. The components or rather IOs in the context of NetInf have unique IDs along with the defined *Access Rights*. The presented scenario for email exchange makes it easy to understand the mechanism of the proposed service. The evaluation section have discussed the analysis of the email data used and is linked with the system performance in terms of data size and the latency experienced while retrieving these data.

The proposed email service presented in this work is in the context of the NetInf architecture for the Future Internet. A service that is secure, reliable and where information is created, stored and retrieved without requiring a dedicated infrastructure unlike the contemporary email services. The basic NetInf prototype has already been implemented. Our ongoing work includes the development of an interface for the NetInf Email service. The performance evaluation of the service will include various tests with different parameters measuring latency, reducing the overall overhead due to the use of asymmetric key cryptography and reputation tests to get rid off Spam emails.

We now move towards the conclusion and future work for this thesis in the following chapter.

Chapter 6

Conclusion and Future Direction

The contemporary Internet architecture is the evolved structure that has sustained continuous development for the past four decades. So far, it has worked brilliantly, way beyond the expectations when it was first created. This characteristic has also been marked as disadvantageous as different challenges, highlighted inadequacies for this model. The amount of mobile data is growing tremendously and it will continue to grow. The addition of wireless technology is one of the main reasons behind this information explosion which worked as a catalyst for the fast evolution of the current Internet architecture. Solutions provided to mitigate the problems faced by this architecture proved to be successful temporarily as the unprecedented increase of mobile users was not expected.

The emerging demands for security, mobility and content distribution are difficult to be provided in a short duration and in an incremental way. The clean-slate approach is one approach where new design rules can be proposed and implemented to address all the current challenges. The notion of the Future Internet architecture has been addressed by different projects initiated all over the globe. The current Internet was based on the idea of host-to-host communication. Today this has been changed to a content distribution approach and demands that the architecture should be information centric rather than node centric. Most of the Future-Internet Projects have adopted this same idea for their architecture design. The idea is to move from IP-based addresses to content persistent content naming. This paradigm shift of Internetworking from being node centric to information centric is called *Information Centric Networking* (ICN).

Mobility management in ICNs has been addressed and different solutions have been proposed in this regard. The Network of Information (NetInf), one of the work packages of the 4WARD [3] project, is an information-centric architecture for the Future Internet and has addressed mobility issues. The key issue highlighted in this thesis is how to support

mobility in heterogeneous wireless networks in the context of both ICN (in general) and NetInf (in particular). This is the problem statement of our thesis that has been addressed through the contributions summarized below.

6.1 Summary of Contributions

In this thesis, primarily, the issue of mobility management is discussed. The NetInf Mobile Node architecture in Chap.3 is the first contribution addressing this issue where after discussing the architectural details, a VNL (Virtual Node Layer) algorithm is presented. In Chap.4, we focused on the non-cooperative behavior of mobile nodes in a network and proposed a solution that encourages cooperation among network entities. The NetInf email service is an application based on the NetInf architecture proposed in Chap.5. The NetInf email proposal is an example representing the NetInf architecture benefits in terms of security, users privacy and content management.

6.1.1 1st Contribution

Our first contribution is the proposal of a NetInf Mobile Node architecture framework. Its design leverages over the NetInf architecture and provides support for mobility, data relay and power management. The inclusion of VNL (Virtual Node Layer) in the architecture introduces three modules namely the Handover Module, the Data Relay Module and the Power Management Module providing support to mobile nodes in different situations. VNL is a programming abstraction (a generalization of a mobile agent) that can move from one node (by suspending the execution of a process) to another (by resuming the execution from the point where it was suspended). This ability of NetInf MNs enables it to support mobile nodes during handover events by avoiding link failures and maintaining ongoing sessions. The Data Relay Module works closely with the Handover Module and supports other mobile nodes by relaying data on their behalf during handover or when experiencing poor connectivity, whereas the Power Management Module manages power consumption of mobile nodes. NetInf MN in an inactive phase is tuned into an idle mode to economize the energy consumption. The basic purpose of mobility support through NetInf mobile nodes is to maintain the QoS (especially during handover). ILCTR and OLCTR are the routing capabilities provided by NetInf MN to route packets between NetInf and non-NetInf sites. They are also capable of buffering or storing the data temporarily working within challenged network environment.

Our proposal for mobility management involves a mutual contribution of network and mobile nodes. The mobility management in our considered case is neither network controlled nor mobile controlled entirely. The Central Control Unit (CCU) is a network entity

that coordinates with NetInf MN. Besides supporting NetInf MN during handovers, CCU has different units that observe and record different activities in the network such as mobility pattern of mobile nodes and mobile nodes movement prediction. The units for allocating mobility zones and VNL coordination are the main contributing units that work directly with NetInf MN. The Mobility Zone Allocation Unit allocates a virtual zone which represents the *locus of data relay and handover support*. The VNL Coordination Unit supports NetInf MN according to its relative position in the network. The presented VNL algorithm shows how each module of NetInf MN is turned on according to the situation with the help of cross-layer design supporting signals from both Network and MAC layers. The VNL working principle is explained through a handover scenario determining how multiple events are sequenced one after the other.

6.1.2 2nd Contribution

The second contribution in this thesis is the formulation of a game theoretical model for data relaying and handover management based on the Reinforcement Learning scheme. The **CO**mbined fully **D**istributed **PA**off and **S**trategy **R**einforcement **L**earning (CODIPAS-RL) scheme helps mobile nodes to learn about their network environment through experimenting different strategies and actions. Strategies and actions that return higher values of payoff are repeated with higher probability. The selected CODIPAS-RL scheme is Multiplicative-Weighted Imitative CODIPAS-RL in which the previous action is imitated with some probability depending upon the received payoff. The formulation of the mathematical model is based on Game Theory. The Stackelberg leadership model, a mathematical model for non-cooperative games, has been applied to a wireless network. In this wireless network model, a 2-level Stackelberg leader-follower model is applied to a set of three players. The leader is the Access Point (AP) and two nodes are its followers. The selfish behaviour of nodes discourage the mutual cooperation in the network. To solve this issue, the leader of the game advertises a reward parameter μ along with the price parameter λ. The reward is the reimbursement to a mobile node that cooperates with its neighbouring nodes during handover situations. The cooperation by a mobile node is defined as data relaying on behalf of a neighbouring mobile node. Parameter λ is the price every node has to pay for using network resources. This mutual cooperation between mobile nodes maximizes the overall network coverage through cooperative diversity. The game in this scenario is played between two mobile nodes and the access point. After the (λ,μ) advertisement, mobile nodes use the CODIPAS-RL scheme to learn the strategy that maximizes their individual payoff or utility functions. Once a point of equilibrium is achieved, the access point calculates its overall revenue. The leader of the game (the access point), use different (λ,μ) values to maximize its revenue.

The simulation results are based on using the same or different sets of actions by mobile nodes. The assumption for selecting actions from their respective action sets is based on the idea of avoiding collisions or interference during transmissions towards the access point. An action is considered as the transmission of a packet at a particular time slot or frequency. The simulation results show the convergence of the proposed model towards a global optimum for different sets of actions. Similarly, all the players maximize their payoff by choosing different actions.

6.1.3 3rd Contribution

The NetInf email service, based on the NetInf architecture, is a use case example and one of the proposed applications in the context of ICN. This application uses the NetInf service such as routing, name resolution and content distribution. The application uses asymmetric key cryptography as a user ID. The user ID is free of domain name and thus does not require dedicated servers and ports for email messaging which is different from the current email service. As every object in NetInf has a unique ID, every email and its components are considered as separate objects and are assigned unique IDs secured through asymmetric key cryptography. The NetInf email service architecture framework is based on the NetInf architecture. The main components of this framework are: Storage Space, Index Space and NetInf API. The storage space stores all kinds of objects in IO and BO formats. The Index Space defines the semantic description of the IOs where as the NetInf API provides an interface to end users to access this service. The NetInf email message format describes how the whole architecture and its components work together for sending and receiving email messages. Commands used to exercise these procedures are location independent and are based on the REST architectural style. Some of the commands used are: Push, Search, Get, and Remove.

The qualitative evaluation of this service is done on a huge sample of emails with variable sizes. It is observed that most of the emails, on average, have small size and the use of asymmetric key will not result into an overload issue. However, this increase can be huge and demanding for large-storage capability in the network. An additional problem can be the overall system latency while sending and receiving an email. NetInf addresses all these issues through a scalable content distribution mechanism. Each object when published in NetInf may be replicated to multiple sites.

6.2 Future Direction

The contribution in Chap.3 (NetInf Mobile Node Architecture) requires the implementation and the validation in a real environment. Currently we are working on NS-2 simulator to

evaluate the performance of the proposed VNL algorithm. We are using the ns-2.29 [108] patch developed by NIST to test our algorithm in a heterogeneous environment. For a real environment emulation, the Android OS [126] is a good platform for NetInf Mobile Node implementation. Similarly, the concept of Virtual Mobility Zones (VMZ) also need to be implemented in a real environment but the test bed for such an implementation requires dedicated mobile nodes in a dedicated area. A university campus is a good choice for such an experiment where users with mobile devices visit frequently and the campus WLAN facility can be used. However, such experiments require a big setup, a dedicated team and ample time.

As far as the mathematical model presented in Chap.4, it can be extended by including other parameters such as power optimization as well as optimal path selection. The discussed utility functions only maximize their throughput. However, it should be noted that adding more functions in the optimization problem will make it more complex. Other CODIPAS-RL schemes can be tested and compared with each other for a performance evaluation. It is not necessary to adopt Game Theory approach to formulate optimization problems. The relevance of Chap.3 and Chap.4 is evident from the fact that both contributions have addressed the same issues but with different approaches. Chap.3 has considered mobility issue with a practical approach at a network level, whereas Chap.4 has investigated the same approach on theoretical grounds.

The NetInf email application is left with implementation issue. The basic NetInf prototype is going to be used for the implementation of this service. The performance evaluation of this service requires different tests and experiments. The list include robustness of the service against malicious attacks, spam filtering, comparison with other email services through reputation systems against spam emails and the reduction of the system latency during email retrieval from the NetInf system.

References

[1] V. Jacobson, D. K. Smetters, J. D. Thornton, M. F. Plass, N. H. Briggs, and R. L. Braynard, "Networking named content," in *Proceedings of the 5th international conference on Emerging networking experiments and technologies*, ser. CoNEXT '09. New York, NY, USA: ACM, 2009, pp. 1–12. [Online]. Available: http://doi.acm.org/10.1145/1658939.1658941

[2] T. Koponen, M. Chawla, B.-G. Chun, A. Ermolinskiy, K. H. Kim, S. Shenker, and I. Stoica, "A data-oriented (and beyond) network architecture," *SIGCOMM Comput. Commun. Rev.*, vol. 37, no. 4, pp. 181–192, Aug. 2007. [Online]. Available: http://doi.acm.org/10.1145/1282427.1282402

[3] http://www.4ward project.eu.

[4] D. Lagutin, K. Visala, and S. Tarkoma, *Publish/Subscribe for Internet : PSIRP Perspective.* Valencia FIA book, 2010, vol. vol. 4.

[5] C. Perkins, "Ip mobility support for ipv4,," *RFC 3344*, 2002.

[6] D. Johnson, C. Perkins, and J.Arkko, "Mobility support in ipv6," *RFC 3775*, 2004.

[7] P. Nikander, J. Ylitalo, and J. Wall, "Integrating security, mobility, and multihoming in hip way," *Proc. NDSS, San Diego, CA*, pp. pp. 87–99, 2003.

[8] M. D'Ambrosio, P. Fasano, M. Marchisio, V. Vercellone, and M. Ullio, "Providing data dissemination services in the future internet," nov. 2008, pp. 1 –6.

[9] A. Eriksson and B. Ohlman, "Dynamic internetworking based on late locator construction," may. 2007, pp. 67 –72.

[10] N. Fotiou, P. Nikander, D. Trossen, and G. C. Polyzos, "Developing information networking further: From psirp to pursuit," *Proc. Intl. Conference on Broadband Communications, Networks and Systems*, 2010.

[11] "http://www.nets-fia.net/."

[12] "http://www.named-data.net."

[13] "http://mobilityfirst.winlab.rutgers.edu/."

[14] "http://nebula.cis.upenn.edu."

[15] "http://www.cs.cmu.edu/ xia/."

[16] "http://www.nets-find.net."

[17] "http://www.geni.net/."

[18] "http://cordis.europa.eu/fp7/ict/fire/."

[19] "http://akari-project.nict.go.jp/eng/index.html."

[20] "http://www.jgn.nict.go.jp/english/index.html."

[21] "http://www.edu.cn/english/."

[22] "http://www.cernet2.edu.cn/."

[23] S. Srinivasan, I. Rimac, V. Hilt, M. Steiner, and H. Schulzrinne, "Unveiling the content-centric features of tcp," in *Communications (ICC), 2011 IEEE International Conference on*, june 2011, pp. 1 –5.

[24] S. Gundavelli, K. Leung, V. Devarapalli, K. Chowdhury, and B. Patil, "Proxy mobile ipv6," *RFC 5213*, 2008.

[25] R. Koodli, "Fast handover for mobile ipv6," *IETF RFC 4068*, July, 2005.

[26] D. Fudenberg and J. Tirole, *Game Theory*, 1992.

[27] H. V. Stackelberg, "The theory of the market economy," *Oxford University Press, Oxford, England*, 1952.

[28] M. Khan, H. Tembine, and A. Vasilakos, "Game dynamics and cost of learning in heterogeneous 4g networks," *Selected Areas in Communications, IEEE Journal on*, vol. 30, no. 1, pp. 198 –213, january 2012.

[29] J. F. Nash, "Equilibrium points in n-points games," *Proc. of the National Academy of Science, vol. 36, no. 1*, pp. pp. 48–49, Jan. 1950.

[30] H. Schulze and K. Mochalski, "Ipoque internet study," *Technical Report*, vol. ipoque GmbH, 2008/2009.

[31] "http://www.fp7-pursuit.eu/."

[32] "http://www.sail-project.eu/."

[33] "http://www.akami.com/."

[34] "http://www.bittorrent.com/."

[35] K. Fall, "A delay-tolerant network architecture for challenged internets," in *Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications*, ser. SIGCOMM '03. New York, NY, USA: ACM, 2003, pp. 27–34. [Online]. Available: http://doi.acm.org/10.1145/863955.863960

[36] É. Renault and D. Zeghlache, "The metalist model: A simple and extensible information model for the future internet," in *EUNICE*, 2009, pp. 88–97.

[37] D. Farinacci, V. Fuller, D. Meyer, and D. Lewis, "Locator/id separation protocol (lisp)," *IETF - draft-ietf-lisp-23*, 2012.

[38] K. Pentikousis, F. Fitzek, and O. Mammela, "Cooperative multiaccess for wireless metropolitan area networks: An information-centric approach," in *Communications Workshops, 2009. ICC Workshops 2009. IEEE International Conference on*, june 2009, pp. 1 –5.

[39] R. D. et al, "Ipv6 dynamic host configuration protocol for ipv6 (dhcpv6)," *RFC 3315*, July 2003.

[40] S. Thomson and T. Narten, "Ipv6 stateless address autoconfiguration," *RFC 2462*, Dec. 1998.

[41] A. Jonsson, C. Perkins, and C. Perkins, "Mobile ipv4 regional registration," *RFC 4857*.

[42] a. C. C. H. Soliman, K. E. Malki, and L. Bellier, "Hierarchical mobile ipv6 mobility management (hmipv6)," *RFC 4140*, August 2005.

[43] A. Misra, S. Das, A. Dutta, A. McAuley, and S. Das, "Idmp-based fast handoffs and paging in ip-based 4g mobile networks," *Communications Magazine, IEEE*, vol. 40, no. 3, pp. 138 –145, mar 2002.

[44] A. Campbell, J. Gomez, S. Kim, A. Valko, C.-Y. Wan, and Z. Turanyi, "Design, implementation, and evaluation of cellular ip," *Personal Communications, IEEE*, vol. 7, no. 4, pp. 42 –49, aug 2000.

[45] R. Ramjee, T. La Porta, S. Thuel, K. Varadhan, and S. Wang, "Hawaii: a domain-based approach for supporting mobility in wide-area wireless networks," in *Network Protocols, 1999. (ICNP '99) Proceedings. Seventh International Conference on*, oct.-3 nov. 1999, pp. 283 – 292.

[46] D. Funato, K. Yasuda, and H. Tokuda, "Tcp-r: Tcp mobility support for continuous operation," in *Proceedings of the 1997 International Conference on Network Protocols (ICNP '97)*, ser. ICNP '97. Washington, DC, USA: IEEE Computer Society, 1997, pp. 229–. [Online]. Available: http://dl.acm.org/citation.cfm?id=850935.852432

[47] A. C. Snoeren and H. Balakrishnan, "An end-to-end approach to host mobility," in *Mobicom*, 2000, pp. 155–166.

[48] D. Maltz and P. Bhagwat, "Msocks: an architecture for transport layer mobility," in *INFOCOM '98. Seventeenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, vol. 3, mar-2 apr 1998, pp. 1037 –1045 vol.3.

[49] K. Brown and S. Singh, "M-udp: Udp for mobile cellular networks," *SIGCOMM Comput. Commun. Rev.*, vol. 26, no. 5, pp. 60–78, Oct. 1996. [Online]. Available: http://doi.acm.org/10.1145/242896.242901

[50] M. Riegel and E. M. Tuexen, "Mobile sctp," *IETF - Internet Draft*, Nov, 2007.

[51] R. Stewart, I. Arias-Rodriguez, K. Poon, A. Caro, and M. Tuexen, "Stream sontrol transmission protocol (sctp)," *IETF - Internet Draft - RFC 4460*, April 2006.

[52] J. R. et al, "Sip session initation protocol," *IETF - Internet Draft - RFC 3261*, June 2002.

[53] a. S. T. P. Vixie, Bellcore, Y. Rekhter, and J. Bound, "Dynamic updates in the domain name system (dns update)," *IETF - Internet Draft- RFC 2136*, April 1997.

[54] P. Eronen and Ed., "Ikev2 mobility and multihoming protocol (mobike)," *IETF - Internet Draft*, February 2006.

[55] F. Chiussi, D. Khotimsky, and S. Krishnan, "Mobility management in third-generation all-ip networks," *Communications Magazine, IEEE*, vol. 40, no. 9, pp. 124 – 135, sep 2002.

[56] A. Campbell, J. Gomez, S. Kim, C.-Y. Wan, Z. Turanyi, and A. Valko, "Comparison of ip micromobility protocols," *Wireless Communications, IEEE*, vol. 9, no. 1, pp. 72 – 82, feb. 2002.

[57] P. Roberts and J. Kempf, "Mobility architecture for the global internet," in *Proceedings of first ACM/IEEE international workshop on Mobility in the evolving internet architecture*, ser. MobiArch '06. New York, NY, USA: ACM, 2006, pp. 23–28. [Online]. Available: http://doi.acm.org/10.1145/1186699.1186709

[58] K.-S. Kong, W. Lee, Y.-H. Han, M.-K. Shin, and H. You, "Mobility management for all-ip mobile networks: mobile ipv6 vs. proxy mobile ipv6," *Wireless Communications, IEEE*, vol. 15, no. 2, pp. 36 –45, april 2008.

[59] S. J. . Koh and Q. Xie, "Mobile sctp with mobile ip for transport layer mobility," *IETF - Internet Draft*, June 2004.

[60] D. Crocker, "Multiple address service for transport (mast): An extended proposal," *IETF - Internet Draft*, Sept 2003.

[61] F. Teraoka, M. Ishiyama, and M. Kunishi, "Lin6: A solution to multihoming and mobility in ipv6," *IETF - Internet Draft*, Dec 2003.

[62] D. Lewis, a. D. F. D. Meyer, and V. Fuller, "Interworking lisp with ipv4 and ipv6," *draft-ietf-lisp-interworking-01*, Feb. 2010.

[63] L. Jakab, A. Cabellos-Aparicio, F. Coras, D. Saucez, and O. Bonaventure, "Lisp-tree: A dns hierarchy to support the lisp mapping system," vol. 28, no. 8, oct. 2010, pp. 1332 –1343.

[64] F. Foukalas, V. Gazis, and N. Alonistioti, "Cross-layer design proposals for wireless mobile networks: a survey and taxonomy," *Communications Surveys Tutorials, IEEE*, vol. 10, no. 1, pp. 70 –85, quarter 2008.

[65] S. Kashihara, K. Tsukamoto, and Y. Oie, "Service-oriented mobility management architecture for seamless handover in ubiquitous networks," *Wireless Communications, IEEE*, vol. 14, no. 2, pp. 28 –34, april 2007.

[66] L. Magagula and H. Chan, "Ieee 802.21-assisted cross-layer design and pmipv6 mobility management framework for next generation wireless networks," in *Networking and Communications, 2008. WIMOB '08. IEEE International Conference on Wireless and Mobile Computing,*, oct. 2008, pp. 159 –164.

[67] S. Mohanty and I. Akyildiz, "A cross-layer (layer 2 + 3) handoff management protocol for next-generation wireless systems," *Mobile Computing, IEEE Transactions on*, vol. 5, no. 10, pp. 1347 –1360, oct. 2006.

[68] P. Feder, R. Isukapalli, and S. Mizikovsky, "Wimax-evdo interworking using mobile ip," *Communications Magazine, IEEE*, vol. 47, no. 6, pp. 122 –131, june 2009.

[69] K. Munasinghe and A. Jamalipour, "Interworking of wlan-umts networks: an ims-based platform for session mobility," *Communications Magazine, IEEE*, vol. 46, no. 9, pp. 184 –191, september 2008.

[70] ——, "Interworked wimax-3g cellular data networks: An architecture for mobility management and performance evaluation," *Wireless Communications, IEEE Transactions on*, vol. 8, no. 4, pp. 1847 –1853, april 2009.

[71] P. Taaghol, A. Salkintzis, and J. Iyer, "Seamless integration of mobile wimax in 3gpp networks," *Communications Magazine, IEEE*, vol. 46, no. 10, pp. 74 –85, october 2008.

[72] S. Cherian, P. Feder, B. Sadeghi, and R. Wisenocker, "Integration of the ims/pcc framework into the mobile wimax network," *Communications Magazine, IEEE*, vol. 46, no. 10, pp. 66 –73, october 2008.

[73] E. Perera, R. Boreli, S. Herborn, M. Georgiades, J. Eisl, and E. Hepworth, "A mobility toolbox architecture for all-ip networks: an ambient networks approach," *Wireless Communications, IEEE*, vol. 15, no. 2, pp. 8 –16, april 2008.

[74] J.-Y. Hu and C.-C. Yang, "On the design of mobility management scheme for 802.16-based network environment," in *Vehicular Technology Conference, 2005. VTC-2005-Fall. 2005 IEEE 62nd*, vol. 2, sept., 2005, pp. 720 – 724.

[75] S. Schuetz, K. Zimmermann, G. Nunzi, S. Schmid, and M. Brunner, "Autonomic and decentralized management of wireless access networks," *Network and Service Management, IEEE Transactions on*, vol. 4, no. 2, pp. 96 –106, sept. 2007.

[76] S.-R. Yang and W.-T. Chen, "Sip multicast-based mobile quality-of-service support over heterogeneous ip multimedia subsystems," *Mobile Computing, IEEE Transactions on*, vol. 7, no. 11, pp. 1297 –1310, nov. 2008.

[77] "http://www.ieee802.org/21/."

[78] L. Eastwood, S. Migaldi, Q. Xie, and V. Gupta, "Mobility using ieee 802.21 in a heterogeneous ieee 802.16/802.11-based, imt-advanced (4g) network," *Wireless Communications, IEEE*, vol. 15, no. 2, pp. 26 –34, april 2008.

[79] P. Neves, F. Fontes, S. Sargento, M. Melo, and K. Pentikousis, "Enhanced media independent handover framework," in *Vehicular Technology Conference, 2009. VTC Spring 2009. IEEE 69th*, april 2009, pp. 1 –5.

[80] A. Pontes, D. dos Passos Silva, J. Jailton, O. Rodrigues, and K. Dias, "Handover management in integrated wlan and mobile wimax networks," *Wireless Communications, IEEE*, vol. 15, no. 5, pp. 86 –95, october 2008.

[81] S. Lee, K. Sriram, K. Kim, Y. H. Kim, and N. Golmie, "Vertical handoff decision algorithms for providing optimized performance in heterogeneous wireless networks," *Vehicular Technology, IEEE Transactions on*, vol. 58, no. 2, pp. 865 –881, feb. 2009.

[82] M. Brown, C. Newport, T. Nolte, N. Lynch, and M. Spindel, "The virtual node layer: A programming abstraction for wireless sensor networks ‚àó."

[83] N. Lynch, S. Mitra, and T. Nolte, "Motion coordination using virtual nodes," in *Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC '05. 44th IEEE Conference on*, dec. 2005, pp. 2823 – 2828.

[84] S. Dolev, S. Gilbert, N. A. Lynch, E. Schiller, A. A. Shvartsman, and J. L. Welch, "Virtual mobile nodes for mobile ad hoc networks," in *in DISC04*, 2004, pp. 230–244.

[85] F. Foukalas, V. Gazis, and N. Alonistioti, "Cross-layer design proposals for wireless mobile networks: a survey and taxonomy," *Communications Surveys Tutorials, IEEE*, vol. 10, no. 1, pp. 70 –85, quarter 2008.

[86] M. Abdelatif, G. Kalebaila, and H. Chan, "A cross-layer mobility management framework based on ieee802.21," in *Personal, Indoor and Mobile Radio Communications, 2007. PIMRC 2007. IEEE 18th International Symposium on*, sept. 2007, pp. 1 –6.

[87] S. Mohanty and I. Akyildiz, "A cross-layer (layer 2 + 3) handoff management protocol for next-generation wireless systems," *Mobile Computing, IEEE Transactions on*, vol. 5, no. 10, pp. 1347 –1360, oct. 2006.

[88] L. Magagula and H. Chan, "Ieee 802.21-assisted cross-layer design and pmipv6 mobility management framework for next generation wireless networks," in *Networking and Communications, 2008. WIMOB '08. IEEE International Conference on Wireless and Mobile Computing,*, oct. 2008, pp. 159 –164.

[89] M. Saleem, E. Renault, and D. Zeghlache, "Netlnf mobile node architecture and mobility management based on lisp mobile node," in *Consumer Communications and Networking Conference (CCNC), 2011 IEEE*, jan. 2011, pp. 981 –982.

[90] R. E. KALMAN, "A new approach to linear filtering and prediction problems," *Transactions of the ASME - Journal of Basic Engineering*, vol. 82 (Series D), pp. 35–45, 1960.

[91] R. H. Milocco and S. Boumerdassi, "Estimation and prediction for tracking trajectories in cellular networks using the recursive prediction error method," *WowMom, Montreal, QC, Canada*, pp. 1 –7, june 2010.

[92] M. Saleem and E. Renault, "Price-reward for data relaying and handover management in wireless networks," *The 13th International Symposium on Mobile Ad Hoc Networking and Computing,(MobiHoc)(Accepted)*, 2012.

[93] S. R.S., "Learning to predict by the method of temporal differences," *Machine Learning, 3*, pp. p.9–44, 1989.

[94] V. Srivastava, J. Neel, A. Mackenzie, R. Menon, L. Dasilva, J. Hicks, J. Reed, and R. Gilles, "Using game theory to analyze wireless ad hoc networks," *Communications Surveys Tutorials, IEEE*, vol. 7, no. 4, pp. 46 – 56, quarter 2005.

[95] M. Felegyhazi and J.-P. Hubaux, "Game theory in wireless networks : A tutorial," *Technical Report: LCA-REPORT-2006-002, EPFL,*, 2006.

[96] H. V. Stackelberg, "Marktform und gleichgewicht (market structure and equilibrium), vienna," 1934.

[97] E. Altman, T. Boulogne, R. El-Azouzi, T. Jiménez, and L. Wynter, "A survey on networking games in telecommunications," *Comput. Oper. Res.*, vol. 33, no. 2, pp. 286–311, Feb. 2006. [Online]. Available: http://dx.doi.org/10.1016/j.cor.2004.06.005

[98] R. A. Howard, *Dynamic Programming and Markov Processes.* The M.I.T. Press, 1960.

[99] R.Bellman, *Dynamic Programming.* Princeton University Press, Princeton, NJ., 1957.

[100] C. Watkins, "Learning from delayed rewards," *PhD Thesis, University of Cambridge, England.*, (1989).

[101] C. P.V.C. and M. Dorigo, "Training and delayed reinforcements in q-learning agents." *International Journal of Intelligent Systems, in press. (Also available as Tech. Rep. IRIDIA/94-14 Universit Libre de Bruxelles, Belgium.)*, (1997).

[102] B. P. Meuleau Nicolas, "Le dilemme exploration/exploitation dans les systems dapprentissage par reinforcement," *These Universit de Caen, Caen, FRANCE*, 1996.

[103] J. E. Moody, S. J. Hanson, S. B. Thrun, and K. Moller, "Active exploration in dynamic environments," 1992.

[104] J. Wyatt and D. I. N. Bu, "Exploration and inference in learning from reinforcement," 1997.

[105] Q. Zhu, H. Tembine, and T. Basar, "Distributed strategic learning with application to network security," *American Control Conference (ACC)*, 2011.

[106] H. Tembine, "Dynamic robust games in mimo systems," *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, vol. 41, no. 4, pp. 990 –1002, aug. 2011.

[107] Q. Zhu, H. Tembine, and T. Basar, "Heterogeneous learning in zero-sum stochastic games with incomplete information," pp. 219 –224, dec. 2010.

[108] *"The Network Simulator NS-2 - NIST add-on - Mac 802.11"*, National Institute of Standards and Technology (NIST), January 2007.

[109] D. Goodman and N. Mandayam, "Power control for wireless data," *IEEE International Workshop on Mobile Multimedia Communications, (MoMuC '99)*, 1999.

[110] C. Saraydar, N. Mandayam, and D. Goodman, "Pricing and power control in a multicell wireless data network," *Selected Areas in Communications, IEEE Journal on*, vol. 19, no. 10, pp. 1883 –1892, oct 2001.

[111] M. Xiao, N. Shroff, and E. Chong, "Utility-based power control in cellular wireless systems," vol. 1, pp. 412 –421 vol.1, 2001.

[112] A. MacKenzie and S. Wicker, "Game theory and the design of self-configuring, adaptive wireless networks," *Communications Magazine, IEEE*, vol. 39, no. 11, pp. 126 –131, nov 2001.

[113] M. Saleem, E. Renault, and D. Zeghlache, "Netlnf mobile node architecture and mobility management based on lisp mobile node," *IEEE Consumer Communications and Networking Conference (CCNC)*, 2011.

[114] E. P. Resnick, "Internet message format," *RFC: 5322*, October 2008.

[115] J. Klensin, "Simple mail transfer protocol," *RFC: 5321*, October 2008.

[116] L. De Alfaro, A. Kulshreshtha, I. Pye, and B. T. Adler, "Reputation systems for open collaboration," *Commun. ACM*, vol. 54, no. 8, pp. 81–87, Aug. 2011. [Online]. Available: http://doi.acm.org/10.1145/1978542.1978560

[117] "http://www.wikipedia.org."

[118] M. Xie and H. Wang, "A collaboration-based autonomous reputation system for email services," mar. 2010, pp. 1 –9.

[119] U. Pandey and S. Chakravarty, "A survey on text classification techniques for e-mail filtering," in *Proceedings of the 2010 Second International Conference on Machine Learning and Computing*, ser. ICMLC '10. Washington, DC, USA: IEEE Computer Society, 2010, pp. 32–36. [Online]. Available: http://dx.doi.org/10.1109/ICMLC.2010.61

[120] S. Turner, "Secure/multipurpose internet mail extensions," *IEEE Internet Computing,*, vol. 14, no. 5, pp. 82 –86, sep. 2010.

[121] J. Callas, L. Donnerhacke, H. Finney, and R. Thayer, "Openpgp message format," *IEFT RFC 2440*, Nov. 1998.

[122] T. Chen and S. Ma, "A secure email encryption proxy based on identity-based cryptography," dec. 2008, pp. 284 –286.

[123] L. Liao and J. Schwenk, "Secure emails in xml format using web services," in *Proceedings of the Fifth European Conference on Web Services*, ser. ECOWS '07. Washington, DC, USA: IEEE Computer Society, 2007, pp. 129–136. [Online]. Available: http://dx.doi.org/10.1109/ECOWS.2007.20

[124] E. Renault, A. Ahmad, and M. Abid, "Toward a security model for the future network of information," *Proceedings of the 4th International Conference on Ubiquitous Information Technologies Applications, 2009. ICUT '09.*, pp. 1 –6, dec. 2009.

[125] R. T. Fielding and R. N. Taylor, "Principled design of the modern web architecture," pp. 407–416, 2000. [Online]. Available: http://doi.acm.org/10.1145/337180.337228

[126] "http://www.android.com/."

List of figures

List of tables

Appendix A

Thesis Publications

Journal

- Muhammad Shoaib Saleem and Eric Renault, *A Nash-Stackelberg Reinforcement Learning Scheme for Handover Management in Information Centric Wireless Networks*, **Under Submission: Computer Communications Elsevier Journal**.

International Conferences

- Muhammad Shoaib Saleem and Eric Renault, **A Nash-Stackelberg Multiplicative Weighted Imitative CODIPAS-RL Scheme for Data Relaying and Handover Management in Wireless Networks**, 10th IEEE Consumer Communication Networking Conference, **CCNC**, 2013.Las Vegas, USA.(Accepted)
- Muhammad Shoaib Saleem and Eric Renault, **Price Reward for Data Relaying and Handover Management in Wireless Networks**, 13th ACM International Symposium on Mobile AdHoc Networking & Computing, **MobiHoc 2012**, Hilton Head Island, South Carolina, USA.
- Muhammad Shoaib Saleem and Eric Renault, **Towards a Secure Email Service for the Future Internet**, 2nd International Conference on Networking and Future Internet, **ICNFI**, 2012. Istanbul, Turkey.
- Muhammad Shoaib Saleem and Eric Renault, **Virtual Node Layer Mobility Management in Network of Information**, 5th ACM EuroSys Doctoral Workshop, EuroDW,2011, In conjunction with **ACM EuroSys 2011**, Salzburg, Austria.
- Muhammad Shoaib Saleem and Eric Renault, **Architecture and Design of Network of Information Mobile Node**, 1st International Conference on Networking and Future Internet, **ICNFI**, 2011. Paris, France.
- Muhammad Shoaib Saleem and Eric Renault, **NetInf Mobile Node Architecture and Mobility Management based on LISP Mobile Node**, 8th IEEE Consumer Communication Networking Conference, **CCNC**, 2011.Las Vegas, USA.
- Muhammad Shoaib Saleem and Eric Renault, *Information Centric Networking based Handover Support for QoS Maintenance in Cooperative Heterogeneous Wireless Networks*, Technical Report, 2011

Appendix B

Version Française

Résumé

L'écosystème Internet contemporain aujourd'hui a traversé série de changements évolutifs dans les quarante ou cinquante dernières années. Initialement conçu comme un réseau pour les nœuds fixes, il a réduit assez bien avec le développement de nouvelles technologies à la fois dans les réseaux fixes et sans fil. Cette architecture basée sur le modèle de communication du réseau téléphonique (1er réseau de nouvelle génération) est un modèle client-serveur sur lequel la communication des systèmes d'échanger des données plus dédié. Appelé comme le réseau 2ème génération, cette architecture au fil des ans a été contestée par de nombreux problèmes et des questions telles que la congestion du réseau, panne de chemin, les attaques DOS, les questions de mobilité pour les nœuds sans fil (pas directement pris en charge par l'architecture), les utilisateurs finaux, etc. demander le réseau un certain élément d'information quelle que soit l'endroit où il est stocké. Cette approche est la notion de base pour un réseau étaient d'information est considéré comme l'entité de premier remplaçant le nœud. Ces réseaux, en général, sont désignées comme des Réseau Centrique Information, dans lequel différents problèmes rencontrés par l'Internet actuel, mentionné ci-dessus, peuvent être traitées avec une approche unificatrice en mettant l'information au centre de l'architecture réseau. À l'échelle mondiale, cette conception de l'architecture réseau est qualifiée de l'Internet Centrique futurs d'information.

De même, l'utilisation d'Internet mobile a été augmentée massivement dans la dernière décennie. Il a été environ 1,2 milliard de mobiles à large bande des abonnements pour 2,4 milliards d'internautes en 2011. En raison de l'efficacité du spectre a augmenté et la disponibilité omniprésente de la connectivité cellulaire, la mobilité et la connectivité transparente est désormais considéré comme des produits de base la vie quotidienne. Toutefois, en cas d'Internet, les solutions IP de mobilité basées sur ne peut pas rattraper son retard dans la performance avec l'évolution rapide des réseaux cellulaires. Par conséquent, l'un des principaux objectifs de l'internet du futur est de concevoir des systèmes de gestion de mobilité qui permettent de surmonter les problèmes dans les réseaux sans fil tels que le transfert et la gestion de la localisation, les multi hébergement, sécurité, etc.

Dans cette thèse, nous avons proposé une solution de gestion de mobilité dans les réseaux sans fil dans le cadre du CII en général et dans le contexte de Réseau de l'information (NetInf) en particulier. Le NetInf est une architecture basée sur le CII internet du futur. Nous proposons une NetInf nœud mobile (MN NetInf) l'architecture qui est rétro-compatible avec l'architecture actuelle d'Internet ainsi. Ce soutien de l'architecture croix.

pour aide à la mobilité travaille en étroite collaboration avec l'Unité Centrale de Contrôle (UCC) (entité du réseau) pour des performances améliorées pour le

transfert de gestion en particulier. La couche virtuelle Node (VNL) algorithme définit la façon dont les différents modules de NetInf MN et des unités CCU travailler ensemble. Le jeu d'apprentissage théorique et Renforcement (CODIPAS-RL) modèle de plan mathématique à base montre comment relayé transfert de gestion et des données dans les réseaux sans fil peuvent augmenter la couverture du réseau à travers la diversité coopérative. Les résultats des simulations montrent que le modèle proposé réalise à la fois de Nash et de Stackelberg équilibres alors que l'sélectionnée CODIPAS-RL régime atteint un optimum global.

Enfin, comme un exemple de cas d'utilisation de l'architecture NetInf, nous proposons le service e-mail NetInf qui ne nécessite pas de serveurs dédiés et un port dédié à la différence du service de messagerie actuel. L'utilisation de clés asymétriques que l'ID de l'utilisateur est la caractéristique unique proposé pour ce service. Les détails NetInf email de services d'architecture suivre le principe de fonctionnement de ce service expliquant comment les différents travaux des éléments architecturaux. Nous discuter des défis et des exigences différentes relatives à ce service. Le prototype développé pour ce service sera utilisé pour la mise en œuvre de ce service.

Etat de l’Art

L'Internet est une architecture contemporaine décennie quatre vieilles, étendu avec le temps avec le développement continu d'une large gamme d'applications. Il a été remarqué récemment que, cette même infrastructure est maintenant principalement utilisée pour la distribution de contenu. La réponse de la communauté de la recherche dans ce qui concerne consiste à re-réfléchir à la conception et l'architecture de l'Internet actuel de telle manière que le nœud centrée sur l'architecture est remplacé par le information centrée sur l'architecture. Cela signifie que, contrairement à l'hôte existant pour héberger infrastructure de routage, un système basé sur le contenu doit être placé. Par conséquent, les applications qui impliquent à générer des demandes de contenu, sont acheminés à l'aide des identifiants de contenu. En d'autres termes, l'information est devenue de plus en plus important dans la communication et le réseautage. Cela est évident du fait que la plupart du trafic sur l'Internet d'aujourd'hui est liée à la distribution de contenu. Cette distribution de contenu aujourd'hui n'implique pas de bout en bout des échanges de données, mais plutôt la date divisée et nommé en tant qu'objets de l'information et est accessible dans une variété de façons. Accès aux données de noms d'objets au centre du réseau est unique. Cette approche Centrique Information a rendu les choses difficiles pour l'architecture de l'Internet actuel et a fait Centrique Information Networking domaine et domaine de recherche important.

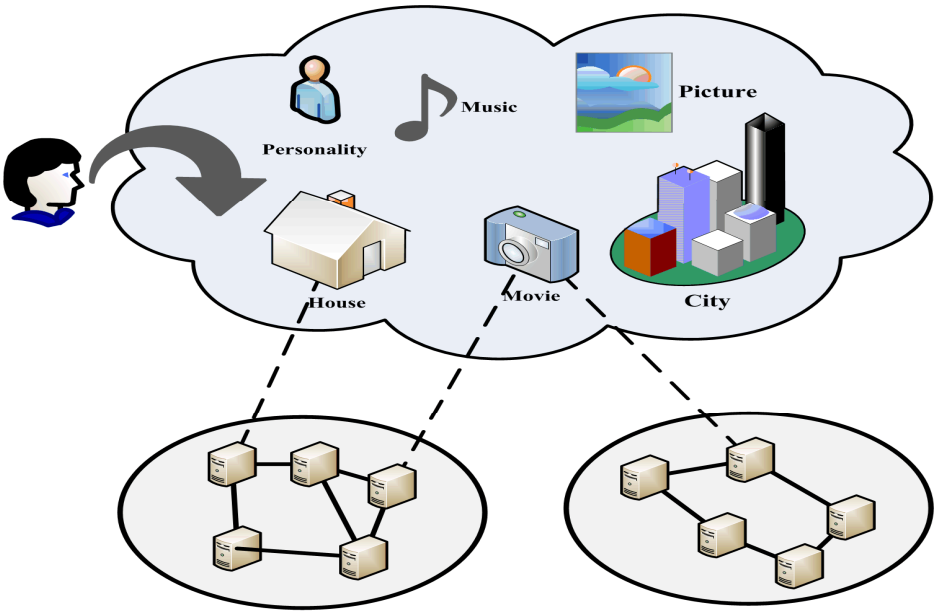


Figure 1:Réseau Centré autour D'Information

Comme mentionné dans le dernier paragraphe, dans réseau Centrique Information (CII) concept, le paradigme est décalée de bout en bout la communication entre les hôtes comme dans l'architecture actuelle d'Internet. La forte demande pour la distribution de contenu se réalise à travers des superpositions P2P pour une grande évolutivité et la distribution d'informations dans l'architecture actuelle de l'Internet. Dans ICN, l'information est considérée comme explicitement comme un {\ it entité de première classe} {Le CII} est un concept de l'architecture Internet du Futur qui a été adressée par un certain nombre de projets dans le monde entier comme PSIRP, 4WARD, PURSUIT, SAIL, DONA, NDN, etc. Ces projets ont proposé des architectures différentes, mais la pierre angulaire pour chaque architecture est de disposer d'informations que l'entité centrée. L'objectif est de rendre la distribution de contenu rapide et fiable en cas de perturbations lors de la communication. Une demande faite par un utilisateur pour une information est délivrée à partir de l'emplacement le plus proche possible plutôt que d'un hôte extrémité. La mise en cache joue un rôle important pour diffuser l'information au CII. Dans le cas des appareils mobiles, soutien à la mobilité dans le CII est un avantage supplémentaire. Dans le contexte actuel, la mobilité est le déplacement physique d'un dispositif mobile à travers son réseau d'accès. Dans une communication active, la mobilité peut sérieusement affecter la qualité de service du point de vue du réseau et la qualité utilisateur final de l'expérience.

Gestion de la mobilité dans les réseaux sans fil est un sujet vaste et beaucoup de travail a été fait dans ce domaine. Les solutions proposées par la communauté de recherche avec des approches diverses est la raison pour laquelle ce domaine de recherche est toujours activent. Gestion de la mobilité a été étudiée avec différents aspects qui incluent complètes architectures de gestion de mobilité ainsi que des outils et des techniques favorisant la mobilité. Par exemple il existe des solutions de

mobilité qui sont basées sur des couches Internet (comme nous le verrons dans ce chapitre) ainsi que des solutions transversales couchent gestion de la mobilité. Dans l'environnement réseau hétérogène, la gestion de la mobilité devient plus intéressante et stimulant. Il existe des solutions de mobilité qui traitent du problème des deux différentes technologies de réseau sans fil qui travaillent ensemble comme po Solutions pour les architectures de réseau spécifiques ont également été proposées. IEEE 802.21 Media Independent Handover (MIH) architectures de mobilité basées sur contribuent également à résoudre ce problème. Dans l'ensemble, la question de la mobilité a été abordée par la communauté de la recherche en tenant compte de différents critères, la méthodologie, les paramètres, etc.

Pour notre travail dans cette thèse, nous avons choisi d'aborder les solutions de mobilité qui sont liées à l'architecture actuelle d'Internet étant donné que presque tous les réseaux sans fil, directement ou indirectement, sont reliés à l'architecture de l'Internet.

Le court venant de solutions de mobilité dans les réseaux TCP / IP a été abordée par l'approche centrée information discuté dans la première partie de ce chapitre. Construit en solutions de mobilité plutôt que add-on une solution (à l'Internet actuel) dans l'architecture Internet semblent plus prometteuses. La contribution de cette thèse est un tel effort, où nous avons proposé des solutions de mobilité qui assure la coopération mutuelle entre le réseau et les nœuds mobiles. Dans un environnement réseau sans fil hétérogène, cette coopération peut être améliorée si le processus de décision (pour le transfert) est efficace. Cela peut être fait grâce à l'apprentissage individuel (sur le milieu environnant) des entités du réseau et de travail pour leurs intérêts individuels (approche la théorie des jeux).

Chapitre 3

Dans ce chapitre, nous présentons une architecture nœud mobile pour le réseau de l'information nommée NetInf nœud mobile. Il s'agit d'une extension de l'architecture nœud de base proposé dans l'architecture NetInf et compatible avec le protocole TCP / IP des réseaux basés sur aussi bien. La couche nœud virtuel et de ses modules introduit dans l'architecture nœud mobile proposé de fournir le soutien suivant:

(i) un transfert sans heurt, avec une latence minimale.

(ii) de relayer des données entre les nœuds voisins dans le réseau afin d'éviter la perte de données lors de la mobilité.

(iii) Gestion de l'alimentation pour éviter la perte de la vie de la batterie nœud mobile.

Interne / externe Routeurs construction de localisation (ILCTR / OLCTR) sont deux fonctions de routage introduits dans NetInf nœuds mobiles pour NetInf et non-NetInf interaction des sites. L'objectif fondamental de NetInf nœud mobile est de maintenir la qualité de service lors d'événements de mobilité. Les événements sont rétrocession des situations critiques où les chances de mobilité au cours de la dégradation de la qualité de service d'une session en cours sont plus élevés. Dans ce travail, l'algorithme proposé VNL est présenté avec l'aide d'un scénario où une session en cours est soutenue au cours du transfert entre les deux réseaux sans fil.

Afin de rendre l'information en réseau centrée suffisamment concurrentiel pour travailler avec l'héritage protocole TCP / IP, il doit fournir des solutions pour le problème existant dans All-IP (Next Generation Networks) réseaux. La nouvelle architecture doit être compatible avec l'architecture Internet existante. L'éventail des questions que la société contemporaine les réseaux IP sont face à de nombreux plis et gestion de la mobilité est l'un d'entre eux. L'Internet actuel n'a pas été conçu pour répondre à chaque problème, donc, la plupart des solutions d'aujourd'hui sont éphémères. Avec l'avancement de la téléphonie mobile, de nouveaux protocoles ont été développés pour gérer la mobilité. Les performances de ces protocoles est resté stable dans les premières années. Cependant, dans la dernière décennie, le développement de nouvelles applications écrasantes accessibles sur les réseaux sans fil au moyen de dispositifs intelligents a exhortés à avoir des algorithmes plus efficaces et les protocoles. Il ya deux solutions possibles. Soit, comme d'habitude, de fournir des solutions correctif ou de concevoir une architecture de la table rase qui fournit des solutions construite en à tous les problèmes et les enjeux auxquels sont confrontés les réseau TCP / IP basé sur.

Il ya une histoire de travail considérable accompli pour la gestion de la mobilité dans les réseaux sans fil. Cependant, le domaine est toujours vivant et très actif en tant que nouvelles technologies émergentes dans de domaine sans fil toujours de nouveaux défis. A chaque couche IP, les exigences en matière de gestion de la mobilité exige des approches différentes. En outre, l'avancement rapide de la communication mobile encourage à développer de nouvelles idées et des cadres. Ubiquitous QoS soutien dans l'environnement réseau sans fil est un grand défi. Dans les zones urbaines, la congestion du trafic de données, la décoloration de canal, et le résultat intolérable ingérence dans les dis-connectivité, la couverture est mauvaise et le manque de qualité de service requise. Nous introduisons la couche virtuelle Node (VNL) concept dans notre étude. Dans notre cadre de travail proposé, nous introduisons:

(i) Une couche de nœuds virtuels (VNL) dans le MN NetInf. Cette VNL est une abstraction de programmation. Le concept a été utilisé mais dans le contexte de la mise en réseau centrée sur l'information, l'idée est nouvelle.

(ii) Nous introduisons une entité centrale connue sous le nom Unité Centrale de Contrôle (UCC) sur les sites du réseau de soutien de coordination entre les différents

VNL technologies de réseau sans fil en particulier au cours du transfert des scénarios. CCU enregistre et met à jour le schéma de mobilité des nœuds mobiles, prévoir le mouvement nœud mobile et la répartition des zones de mobilité pour les nœuds virtuels (expliqué dans les sections suivantes).

(iii) Notre proposition met l'accent sur la collaboration du réseau et le terminal utilisateur final mobile. Les événements de mobilité abordés ne sont ni contrôlés ni réseau nœud mobile contrôlé. Le contrôle est partiellement partagé entre deux entités. Le VNL avec ses modules, a expliqué dans les sections suivantes, avec la coopération couche de fond, donner seamless handover avec une probabilité minimum de toute défaillance.

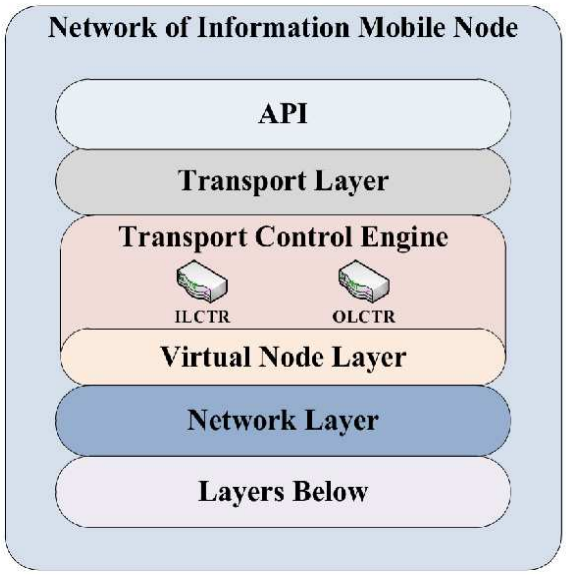


Figure 2: Network of Information Mobile Node (NetInf MN)

Ce chapitre présente un bref aperçu de la solution de mobilité proposé pour le Réseau de l'information suivie de l'introduction à notre architecture NetInf nœud proposé mobile. Les détails décrits dans les paragraphes plus tard mis en évidence des caractéristiques différentes exposées par NetInf nœud mobile. L'abstraction VNL est la généralisation de l'agent mobile avec des fonctionnalités supplémentaires représentées par la remise, la puissance et des modules de relais de données. L'ILCTR et les fonctions OLCTR dans le moteur de contrôle des transports (TCE) facilite le transfert de données entre NetInf et non-NetInf sites. NetInf nœud mobile travaille en étroite collaboration avec l'Unité Centrale de Contrôle (UCC) pour l'exécuter la fonction. Pendant le transfert et de relayer des données, les nœuds mobiles sont mutuellement assistée par le réseau et NetInf nœuds mobiles. Les unités CCU fonctionner pour effectuer des tâches diverses, par exemple, les schémas de mobilité Stockage de nœuds mobiles et la prédiction des mouvements nœud mobile. Cependant, les jeux de rôle de base CCU est la répartition des zones de mobilité virtuelle (VMZs) et d'aider NetInf nœuds mobiles pour activer VNL. Ceci

est rendu possible en proposant l'algorithme VNL, s'explique par le scénario de transfert présentée.

Chapitre 4

Ce chapitre présente un schéma d'apprentissage Prix-Récompense pour encourager la coordination mutuelle entre les nœuds mobiles et de leurs réseaux sans fil. Afin de maximiser la couverture globale du réseau à travers la diversité coopérative, une Nash-Stackelberg multiplicative pondéré imitation CODIPAS-RL régime est proposé. Le réseau sans fil met en œuvre un jeu à 2 niveaux de Stackelberg en introduisant des prix de récompense (lambda, mu) alors que le paramètres Apprentissage par renforcement (RL) régime ouvre la voie pour les nœuds mobiles pour atteindre un état d'équilibre de Nash-.

L'évaluation des performances du système d'apprentissage pour le scénario présenté prouve une convergence rapide vers la solution optimale en adoptant divers ensembles d'actions pour les stratégies choisies. Cela garantit la durabilité de QoS au cours du transfert des situations en relayant des données et évite les collisions entre les nœuds mobiles tout en accédant aux ressources réseau.

L'utilisation de l'Internet dans les réseaux sans fil toujours une demande pour une meilleure qualité de service au cours de la mobilité. Nous avons proposé un algorithme de relais de transfert de données et qui a montré comment la diversité de coopération dans les réseaux sans fil augmente la couverture du réseau et les revenus et assure la fiabilité de connexion pendant le transfert des situations. Dans un réseau sans fil où les nœuds mobiles sont très dynamiques et évoluent dans des directions aléatoires, un nœud mobile permettra d'éviter de coopérer avec d'autres nœuds en raison de: (a) la perte de la vie inutile de la batterie et (b) la perturbation ou le retard des données personnelles. Par conséquent, il est nécessaire à l'appareil une stratégie qui devrait encourager les nœuds mobiles à coopérer dans un réseau sans fil.

Chaque nœud mobile dans un réseau sans fil en concurrence afin de maximiser son rendement. Une stratégie de prix-récompense peut être adoptée qui tente nœuds de coopérer. Dans ce chapitre, nous avons étendu notre effort pour montrer comment le réseau et les nœuds peuvent travailler de façon coopérative afin de maximiser leurs utilités individuelles. Ceci est réalisé par des stratégies suivantes qui finissent par entraîner vers des conditions d'équilibre de Nash-Stackelberg. Nous avons développé un scénario pour modéliser notre problème où un réseau agit comme un chef de file et les nœuds mobiles comme disciples. Dans ce jeu 2-niveau Stackelberg, le chef de file (point d'accès WLAN dans notre cas) propose des prix de

récompense (lambda, mu) paramètres du réseau. Les nœuds mobiles, sachant que le réseau encourage le relais évité de suivre leurs intérêts égoïstes.

Dans un environnement dynamique où un nœud mobile suit des trajectoires aléatoires, sans connaître les probabilités de transition d'un état à un autre nécessite un apprentissage par renforcement (RL) régime. Dans les régimes RL, les apprenants d'interagir avec leur environnement et d'utiliser leur expérience pour choisir ou éviter certaines actions en fonction de leurs conséquences. Les actions qui ont conduit à gains élevés dans une certaine situation (ou état) tendent à se répéter chaque fois la même situation (état) revient, alors que les choix qui ont conduit à gains relativement faibles tendent à être évités.

Sur la base de la discussion ci-dessus, le projet de Nash-Stackelberg multiplicative pondéré imitation CODIPAS-RL régime est conçu pour une convergence rapide pour arriver à condition d'équilibre de Nash suivie par l'état d'équilibre de Stackelberg.

Les contributions présentées dans ce chapitre sont les suivantes:

(i) le relais de données et la remise algorithme de gestion pour assurer une mobilité transparente.

(ii) Proposition d'un Nash-Stackelberg multiplicative pondéré imitation CODIPAS-RL régime avec un taux de convergence rapide vers des solutions souhaitables pour tous les joueurs dans le jeu.

(iii) amélioration du chiffre d'affaires global du réseau et la couverture grâce à la coopération entre les nœuds mobiles et le réseau sans fil.

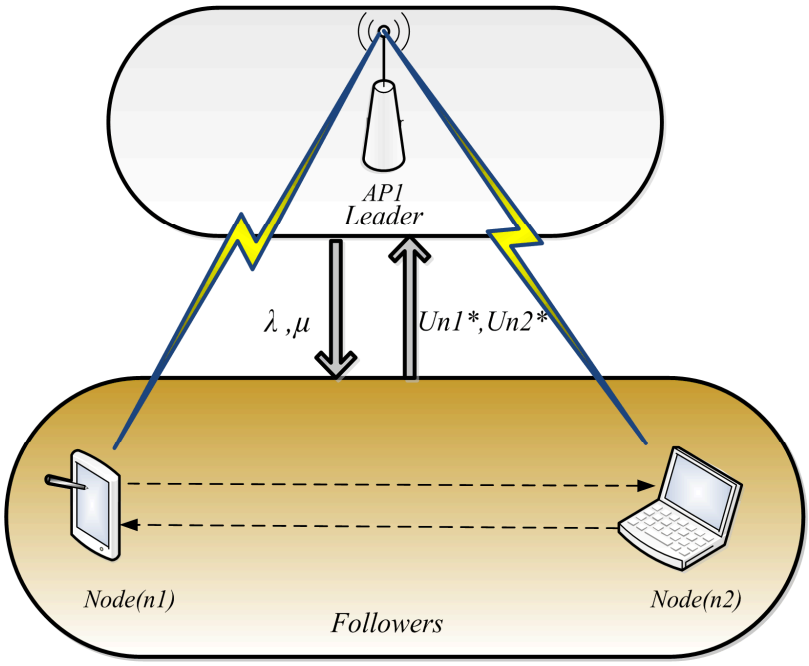


Figure 3: Stackelberg Leadership Model

Dans ce chapitre, nous avons étudié le projet de Nash-Stackelberg imitation CODIPAS-RL régime dans un scénario de réseau sans fil pour relayer des données et la remise de gestion. L'algorithme dans est modélisée mathématiquement dans lequel les fonctions d'utilité pour les nœuds mobiles et le réseau sans fil sont formulées. Afin d'atteindre l'équilibre de Stackelberg, le régime apprentissage par renforcement est utilisé pour d'abord atteindre le point d'équilibre de Nash (NEP). L'évaluation des performances montre que l'algorithme proposé fonctionne bien sous la condition lorsque des actions différentes sont choisies pour des stratégies choisies. Cela évite éventuellement les chances de collisions pendant la transmission et améliore chaque nœud utilité individuelle ainsi que les résultats dans la convergence vers un optimum global de pur les stratégies choisies. En conséquence, le réseau optimise ses revenus dans des conditions stables.

Chapitre 5

L'idée derrière le développement de l'architecture Internet du futur est d'améliorer la disponibilité de l'information à l'échelle mondiale. Il ya eu beaucoup d'efforts pour développer l'architecture de l'Internet tout en redéfinissant l'avenir des blocs de construction de base de l'architecture de l'Internet.

Le Réseau de l'information est un Centrique Information Networking architecture basée sur l'Internet du futur, où une information est considérée comme l'unité de Premier du réseau. Les nœuds, contrairement à l'architecture contemporaine sur Internet, ne sont pas considérés comme la source de l'information, mais comme des machines de stockage, de traitement et de transmission des données. L'Internet TCP / IP est un réseau de réseaux de nœuds interconnectés. Toutefois, le principal intérêt des utilisateurs finaux est une information, pas leur emplacement (à savoir le serveur ou l'adresse IP où il est stocké).

Les utilisateurs de l'Internet toujours envie de profiter du meilleur de la qualité de service (QoS). Cependant, la qualité de l'expérience (QoE) du point de vue des utilisateurs d'aujourd'hui définit «les paramètres pour la conception (nouveau site Web et de services Internet) ou de la refonte (services existants). Parmi tous les services Web / Internet disponibles, e-mail est l'un des services largement utilisés sur Internet.

Ce chapitre présente un nouveau service de messagerie basé sur le NetInf architecture et utilise le NetInf services fournit en termes de routage, de résolution de nom, le stockage/récupération de l'information et de distribution de contenu.

L'utilisation de la cryptographie à clé asymétrique en tant qu'utilisateur ID est une caractéristique unique introduit dans l'architecture NetInf. Cet ID, sans nom de domaine attaché, qui rend ce service sans serveur évolutive, sécurisée et fiable.

La principale différence entre le service de messagerie NetInf et l'actuel, c'est son indépendance à partir de serveurs dédiés, les ports et les protocoles. Étant donné que chaque entité dans NetInf est considérée comme un objet avec un unique ID, chaque e-mail est considéré comme un objet distinct garanti par la cryptographie à clé asymétrique. Il ya quelques idées utiles qui peuvent aider à réduire la gestion et à une augmentation globale de la surcharge a entraîné à l'aide de clés de cryptographie, comme celui présenté dans

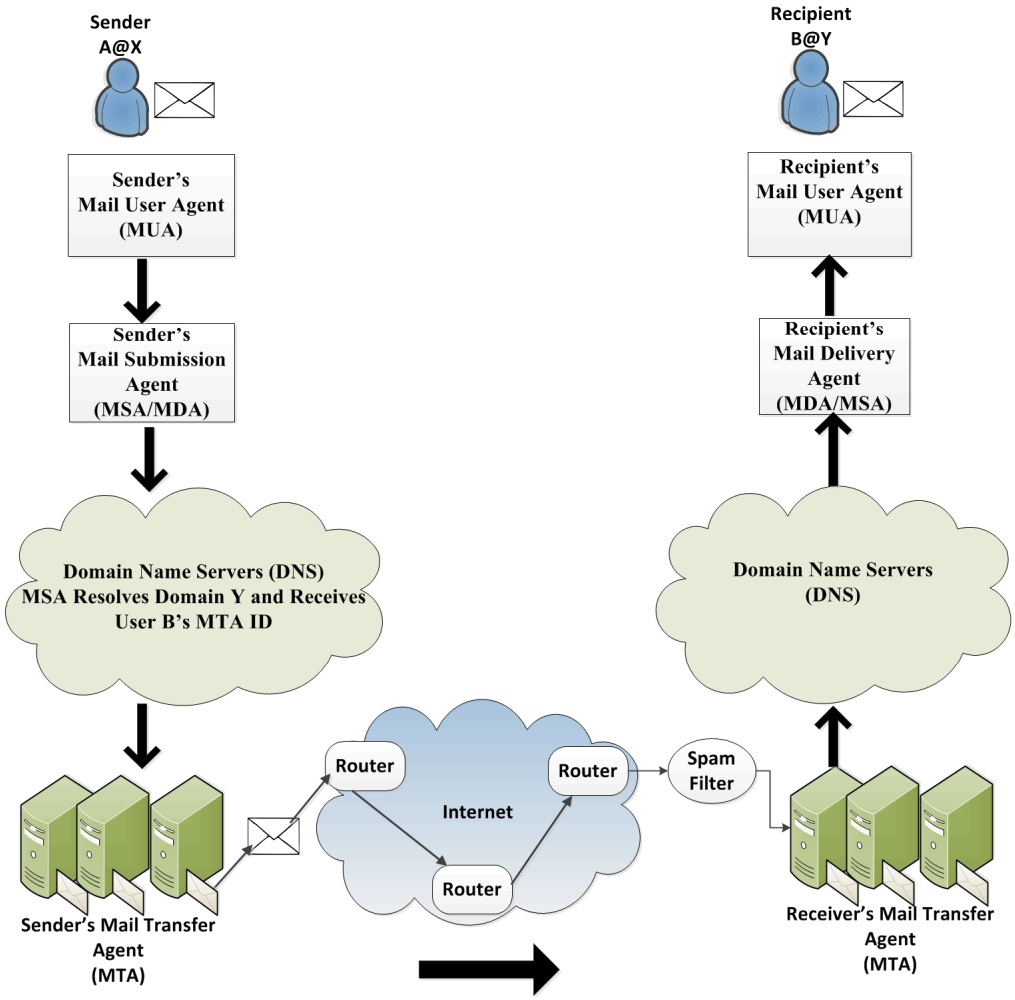


Figure 4: Email Service Today

Les services de courrier électronique d'aujourd'hui de suivre plus ou moins la même procédure pour l'envoi et la récupération des e-mails comme on le verra dans l'exemple ci-dessus. Pourtant, il ya des questions qui dégradent la performance

globale de ce système dans son ensemble. Pour les vers par exemple, le spamming et l'ordinateur, e-mail spoofing questions, de sécurité et de la vie privée en utilisant l'adresse e-mail comme une identité pour les services en ligne sont les questions qui restent des services de messagerie effet sur l'Internet. Tous ces facteurs mettent en péril l'utilité du service de courriel.

Spam est l'un des préoccupations majeures aujourd'hui. Il peut être défini comme le nombre de courriels indésirables qui sont distribués et reçus sur Internet sur des bases quotidiennes. De même les vers informatiques en étant indépendants dans leur e-mail usage de la nature comme un moyen de répliquer à des ordinateurs vulnérables. Les vers informatiques principalement affecter la performance du réseau qui est à la différence des virus informatiques qui portent atteinte aux fichiers et dossiers.

Les e-mails frauduleux qui sont en réalité des emails usurpés sont générés lorsque l'information d'en-tête dans un courriel est modifié pour rendre le message apparaîtra comme si elle est reçue d'un expéditeur connu. Ces emails sont souvent utilisés pour révéler les renseignements personnels.

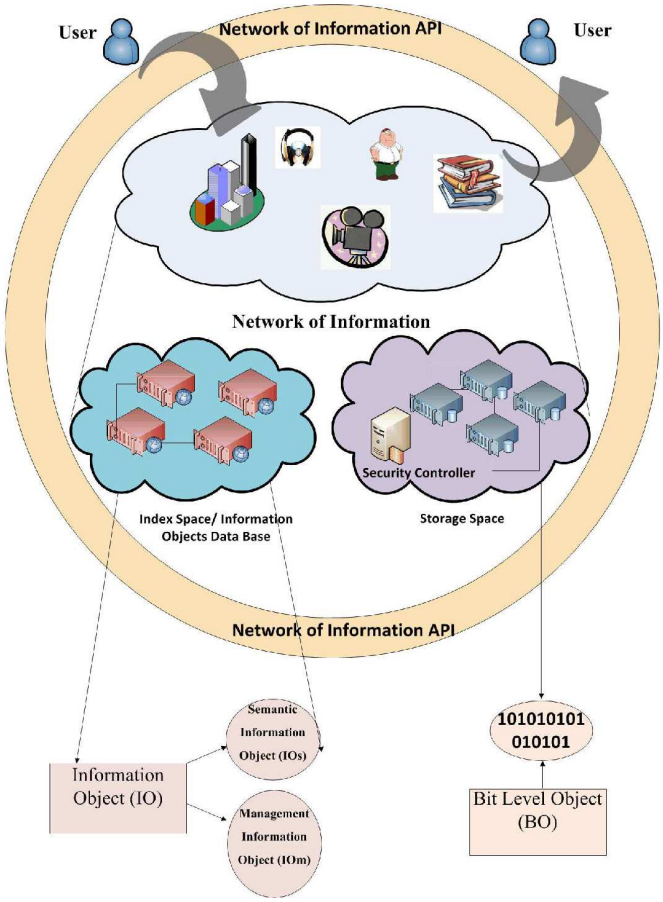


Figure 5: NetInf Email Service Architecture

La vie privée et les menaces de sécurité sont préoccupation majeure pour les utilisateurs de messagerie. La raison possible pour atteinte à la vie, c'est que la plupart du temps les messages ne sont pas cryptés. De même, un e-mail doit passer par un niveau intermédiaire de nombreux (comme mentionné plus haut) pour atteindre sa destination finale. Cela rend les choses faciles pour les pirates d'intervenir et de lire / modifier / supprimer le message. L'utilisation de l'adresse e-mail comme identifiant pour les services en ligne hébergés sur divers sites Web a soulevé la vie privée et les questions de sécurité. Divers application en ligne comme Facebook, LinkedIn, Google Docs, etc., exige adresses email des utilisateurs que l'identité. Le comportement irresponsable de la majorité des utilisateurs est également l'une des raisons qui introduisent ces questions de sécurité et de confidentialité.

Ce chapitre a présenté un service de messagerie innovante basée sur NetInf. L'objectif de ce travail est d'introduire un service de messagerie roman. Les caractéristiques uniques de ce nouveau service prennent en charge l'idée d'avoir un réseau sans ports dédiés et des serveurs. L'approche utilisée est basée sur l'architecture Internet du futur réseau connu sous le nom de l'information. Ce service de messagerie fonctionne sur le dessus de l'architecture NetInf. La confidentialité et la sécurité du contenu e-mail est assurée par l'utilisation des clés asymétriques. Cette idée n'est pas nouvelle pour la sécurisation des données d'information, mais l'utilisation de la de clés publique / privée en tant qu'utilisateur ID est un nouveau concept dans le contexte de NetInf. Le format e-mail NetInf définit les différentes composantes du message e-mail NetInf. Les composants ou plutôt IO dans le cadre de NetInf ont unique ID avec définis les droits d'accès. Le scénario de l'échange électronique présenté le rend facile à comprendre le mécanisme du service proposé.

La section d'évaluation a examiné l'analyse des données de messagerie utilisée et est liée à la performance du système en termes de taille et de la latence connu lors de la récupération de ces données.

Le service de messagerie proposé dans le présent travail s'inscrit dans le contexte de l'architecture NetInf pour l'internet du futur. Un service qui est sûr, fiable et où l'information est créée, stockée et récupérée sans nécessiter une infrastructure dédiée à la différence des services de courrier électronique contemporains. Le prototype NetInf de base a déjà été mis en œuvre. Notre travail en cours comprend le développement d'une interface pour le service e-mail NetInf. L'évaluation des performances du service comprendra divers tests avec différents paramètres de mesure de latence, ce qui réduit la surcharge globale due à l'utilisation de la cryptographie à clé asymétrique et essais réputation de prendre soin de spam.

Conclusion

L'architecture contemporaine est l'Internet la structure a évolué qui a soutenu le développement continu pour les quatre dernières décennies. Jusqu'à présent, il a travaillé avec brio, bien au-delà des attentes quand il a d'abord été créé. Cette caractéristique a également été marquée aussi mauvaises que des défis différents, avec le passage du temps, a souligné les insuffisances de ce modèle. La quantité de données mobiles connaît une croissance extraordinaire et il va continuer à croître. L'ajout de la technologie sans fil est l'un des principales raisons derrière cette explosion de l'information qui a travaillé en tant que catalyseur de l'évolution rapide de l'architecture actuelle d'Internet. Les solutions proposées pour atténuer les problèmes rencontrés par cette architecture s'est révélée fructueuse temporairement l'augmentation sans précédent des utilisateurs de téléphones mobiles n'a pas été prévu.

Les nouvelles demandes pour la distribution de la sécurité, la mobilité et le contenu, etc. sont difficiles à être fournis dans une courte durée et de façon incrémentale. L'approche de la table rase est une approche où les nouvelles règles de conception peuvent être proposés et mis en œuvre pour relever tous les défis actuels. La notion de l'architecture Internet du futur a été abordée par les différents projets initiés dans le monde entier. L'Internet d'aujourd'hui est fondé sur l'idée de l'hôte à hôte de communication. Aujourd'hui cela a eu des changements à l'approche de distribution de contenu et les exigences que l'architecture devrait être centrée sur l'information plutôt que centrée sur nœud. La plupart des projets internet du futur ont adopté cette même idée de leur conception de l'architecture. L'idée est de passer d'adresses IP à base de nommage au contenu persistant. Ce changement de paradigme de la forme Internetworking nœud centrée étant de centrée sur l'information est appelé information en réseau Centrique.

Gestion de la mobilité dans les ICNs a été adressée et des solutions différentes ont été proposées à cet égard. Le Réseau de l'information (NetInf) work package de 4WARD est une architecture centrée sur l'information pour l'Internet du futur et a abordé les questions de mobilité. La question clé mis en évidence dans cette thèse est de savoir comment favoriser la mobilité dans les réseaux sans fil hétérogènes dans le cadre de deux CII (en général) et NetInf (en particulier). Cette déclaration le problème de cette thèse qui a été adressée.

Dans cette thèse, principalement la question de la gestion de la mobilité est discutée. L'architecture NetInf nœud mobile dans le chapitre 3 est la première contribution de cette question, où après avoir discuté les détails architecturaux, VNL (couche de nœuds virtuels) algorithme est présenté. Dans le chapitre 4, nous nous sommes concentrés sur le comportement non-coopératif de nœuds mobiles dans un réseau et a proposé une solution qui encourage la coopération entre les entités du réseau. Le service de courriel NetInf est une application basée dans l'architecture NetInf proposé dans le chapitre 5. La proposition email NetInf est un exemple d'architecture représentant avantages NetInf en termes de sécurité, la confidentialité des utilisateurs et de gestion de contenu.

Première Contribution

Notre première contribution est la proposition d'une architecture Mobile Node NetInf. Sa conception s'appuie sur l'architecture au cours NetInf et fournit un soutien à la mobilité, de relais de données et de gestion de l'alimentation. L'inclusion de VNL (couche de nœuds virtuels) dans l'architecture de définir trois modules, à savoir, de passation Module, module de relais de données et le module de gestion de l'alimentation assurant ces fonctions qui prennent en charge les nœuds mobiles dans des situations différentes. VNL est l'abstraction de programmation (une généralisation de l'agent mobile) qui peut se déplacer d'un nœud (en suspendant l'exécution d'un processus) vers un autre nœud (par reprise de l'exécution du point où elle a été suspendue). Cette capacité de NetInf MN lui permet de supporter les nœuds mobiles lors d'événements de mobilité en évitant toute défaillance d'un nœud et le maintien d'une session en cours. Le module de relais de données travaille en étroite collaboration avec remise module et soutient d'autres nœuds mobiles par relayer des données sur leur nom au cours des événements lorsque des événements voisins mobiles nœuds expérience comme la mobilité (handover) ou de la connectivité des pauvres. Le module de gestion de puissance gère la consommation d'énergie des nœuds mobiles. NetInf MN dans une phase d'inactivité est réglé dans un mode veille pour économiser la consommation d'énergie. L'objectif fondamental de soutien à la mobilité par le biais NetInf nœud mobile est de maintenir la qualité de service (en particulier au cours du transfert). ILCTR et OLCTR sont les capacités de routage fournis par NetInf MN pour router les paquets entre NetInf et non-NetInf sites. Ils sont également capables de tamponner les données temporairement si l'on travaille dans un environnement en question.

Notre proposition de gestion de la mobilité implique la contribution mutuelle du réseau et des nœuds mobiles. La gestion de la mobilité dans notre cas considéré n'est ni contrôlée ni mobile réseau entièrement contrôlé. L'unité de contrôle centrale (CCU) est une entité du réseau qui coordonne avec NetInf MN. En plus de soutenir

NetInf MN lors des passations, CCU a différentes unités qui observent et consignent des différentes activités dans le réseau tel que le profil de mobilité des nœuds mobiles et mobiles de prédiction mouvement nœuds. L'unité de répartition des zones de mobilité et de l'unité de coordination VNL sont les principales unités contribuant qui travaillent directement avec NetInf MN. L'unité d'allocation de mobilité point alloue zone virtuelle qui représentent la (lieu de relais de données et la remise soutien). L'unité de coordination VNL soutient NetInf MN en fonction de sa position relative dans le réseau. L'algorithme présenté VNL montre comment chaque module de NetInf MN est activée en fonction de la situation avec l'aide de la couche traverse de support sous forme de signaux de réseau et la couche MAC. Le principe de travail est VNL a expliqué à travers un scénario de transfert.

Deuxième Contribution

La deuxième contribution de cette thèse est la formulation d'un modèle de jeu théorique pour relayer des données et la remise de gestion fondé sur un schéma Apprentissage par renforcement. Le Combined pleinement Distributed Payoff et Strategy Reinforcement Learning gain (CODIPAS-RL) régime permet nœuds mobiles d'apprendre au sujet de leur environnement réseau à travers l'expérimentation des stratégies et des actions différentes. Stratégies et actions qui retournent des valeurs plus élevées de gain sont répétées avec une plus grande probabilité. Le sélectionnée CODIPAS-RL schéma est multiplicatif pondéré imitation CODIPAS-RL dans lequel l'action précédente est imité avec une certaine probabilité en fonction de la récompense reçue. La formulation du modèle mathématique est basée sur la théorie des jeux. Le leadership de Stackelberg modèle mathématique pour le modèle basé sur jeux non-coopératifs est appliqué à un réseau sans fil. Dans ce modèle de réseau sans fil d'un 2-niveau Stackelberg leader-suiveur modèle est appliqué à un ensemble de trois joueurs. Le chef de file est le Point d'Accès (AP) et deux nœuds sont ses disciples. Le comportement égoïste des nœuds décourage la coopération mutuelle dans le réseau. Pour résoudre ce problème, le meneur de jeu annonce un paramètre récompense mu avec le paramètre de prix lambda. La récompense est le remboursement à un nœud mobile qui coopère avec ses nœuds voisins au cours du transfert situation. La coopération par un nœud mobile est définie comme relayer des données pour le compte d'un nœud voisin mobile. Le paramètre lambda est le prix tous les nœuds doit payer pour l'utilisation des ressources réseau. Cette coopération mutuelle entre les nœuds mobiles maximise la couverture globale du réseau à travers la diversité coopérative. Le jeu dans ce scénario se joue entre deux nœuds mobiles et le point d'accès. Après (lambda, mu) la publicité, les nœuds mobiles utilisent CODIPAS-RL régime d'apprendre la stratégie de maximiser leur gain individuel ou des fonctions d'utilité. Une fois un point d'équilibre est atteint, un point d'accès calcule son chiffre d'affaires global. Le meneur de jeu (point d'accès), utilisation différente (lambda, mu) des valeurs afin de maximiser son chiffre d'affaires.

Les résultats des simulations sont basés sur l'utilisation des ensembles identiques ou différents des mesures prises par les nœuds mobiles. L'hypothèse de la sélection des actions de leurs jeux d'action respectifs se fonde sur l'idée d'éviter les collisions ou d'interférence lors de la transmission vers le point d'accès. L'action est jugée comme transmission de paquets à un intervalle de temps particulier ou de la fréquence. Le résultat montre la convergence du modèle proposé vers un optimum global pour divers ensembles d'actions. De même, tous les joueurs de maximiser leur gain en choisissant différentes actions.

Troisième contribution

Le service de courriel NetInf est basé sur une architecture NetInf et est un cas d'utilisation comme l'une des applications proposées pour le CII. Cette application utilise le service NetInf telles que le routage, la résolution de noms et de distribution de contenu. L'application utilise la cryptographie asymétrique clé comme ID utilisateur. L'ID utilisateur est libre à partir du nom de domaine et ne nécessite donc pas de serveurs dédiés et des ports pour la messagerie électronique qui est différent du service de messagerie actuel. Comme chaque objet dans NetInf possède un ID unique, chaque e-mail et de ses composants sont considérés comme des objets distincts sont attribués à des identifiants uniques obtenus grâce à la cryptographie à clé asymétrique. L'architecture NetInf service de courrier électronique est basé sur une architecture NetInf. Les principales composantes de ce service sont les suivants: l'espace de stockage, Index de l'espace et NetInf API.

L'espace de stockage stocke toutes sortes d'objets en IO et les formats de BO. L'espace indice définit la description sémantique de l'OI, où que NetInf API fournit une interface pour les utilisateurs finaux d'accéder à ce service. Le format NetInf message décrit comment les différentes composantes d'un message e-mail sont considérées comme des objets dans le contexte de NetInf. Le scénario email NetInf travail explique comment ces composants sont envoyées et reçues. Les trois blocs de construction de service de messagerie NetInf, à savoir, l'espace de stockage, Indice de l'espace et NetInf API mis leur part au cours de ces procédures. Les ensembles de commandes utilisées sont indépendante de la localisation et sont basés sur l'architecture de style REST. Parmi les exemples courants sont Push, Recherche, obtenir, supprimer, etc.

L'évaluation qualitative de ce service se fait sur un vaste échantillon d'e-mails avec des tailles variables. Il est observé que la plupart des e-mails, en moyenne, de petite taille et l'utilisation de la clé asymétrique n'entraînera pas d'être une surcharge.

Cependant, collectivement cette augmentation peut être énorme et les demandes de grande capacité de stockage dans le réseau. Un problème supplémentaire peut être latence générale du système lors de l'envoi et la réception d'un email. NetInf aborde cette question à travers le mécanisme de distribution de contenu évolutive. Chaque objet lors de la publication dans NetInf est répliqué sur plusieurs sites. Ainsi, nous voyons un compromis entre la latence du réseau et de sécurité de contenu (en utilisant les touches de cryptographie).

Nous concluons notre thèse en mentionnant quelques-unes des directions de recherche futures.

La contribution au chapitre 3 (NetInf nœud mobile Architecture) nécessite un cadre complet pour la mise en œuvre dans un environnement réel. Actuellement, nous travaillons sur simulateur NS-2 pour évaluer la performance de l'algorithme proposé VNL. Nous utilisons ns-2.29 correctif développé par le NIST pour tester notre algorithme dans un environnement hétérogène. Pour l'émulation environnement réel, Androïde OS est une meilleure plateforme pour la mise en œuvre NetInf nœud mobile. De même, le concept des zones de mobilité virtuelle (VMZ) doivent également être mises en œuvre dans un environnement réel, mais le banc d'essai pour la mise en œuvre exige une telle dédiés nœuds mobiles avec un espace dédié. Un campus universitaire est un bon choix pour une telle expérience où les utilisateurs de périphériques mobiles visitez fréquemment et installation sur le campus de WLAN peut utiliser. Toutefois, de telles expériences nécessitent une configuration grande, une équipe dédiée et beaucoup de temps.

En ce qui concerne le modèle mathématique présenté au chapitre 4, il peut être étendu en y incluant d'autres paramètres tels que l'optimisation de la puissance ainsi que la sélection de chemin optimale. Les fonctions d'utilité discutée maximisé leur débit seulement. Toutefois, il convient de noter que l'ajout de plus de fonction dans le problème d'optimisation, il sera plus complexe. Autres CODIPAS-RL régimes peuvent être testés et comparés les uns avec les autres comme une comparaison des performances. Il n'est pas nécessaire d'adopter une approche théorie des jeux à formuler des problèmes d'optimisation. Cela peut être une approche intéressante. L'intégration du chapitre 3 et chapitre 4 est une forme évidente du fait que les deux contributions ont abordé les questions mêmes, mais avec des approches différentes. Chapitre 3 a examiné question de la mobilité avec une approche pratique où le chapitre 4 a enquêté sur la même approche sur des bases théoriques.

L'application de messagerie NetInf se retrouve avec problème d'implémentation. Le prototype NetInf de base va être utilisé pour la mise en œuvre de ce service. L'évaluation des performances de ce service nécessite des tests et des expériences différentes. La liste comprend la robustesse du service contre les attaques malveillantes, le filtrage des spam, la comparaison avec d'autres services de courrier électronique par le biais des systèmes de réputation contre les spam et la réduction de la latence du système pendant la récupération email.