Context-Aware Handoff Support for Wireless Access Networks

By:

Lekometsa MOKHESI

Supervisor:
Dr. Antoine BAGULA

THESIS PRESENTED FOR THE DEGREE OF MASTER OF SCIENCE
IN THE DEPARTMENT OF COMPUTER SCIENCE
UNIVERSITY OF CAPE TOWN

-NOVEMBER 2010-

I know the meaning of plagiarism and declare that all of the work in the document, save for which is properly acknowledged, is my own.

Lekometsa Mokhesi

DEDICATION

To my late brother Tota "Storsh" Mokhesi
To my Mother, 'Malipolelo Mokhesi
To my Sister, 'Lipolelo Mokhesi
To my Nieces, Mpho Mokhesi and Tumeliso Motsomi

ACKNOWLEDGEMENTS

To the source of life, Elohim, The ultimate King, Jesus Christ
and The Governor, The Holy Spirit for the awesome privilege of Life.
Thy Kingdom Come!
To My supervisor for his guidance and support throughout this research.
To My family, for their continual love and support.
To My friends for their encouragement .
To Luca Foschini for his assistance and support with MUM prototype.
To Anders Madsen for his assistance with Hugin Expert.
To the Government of the Kingdom of Lesotho for their financial assistance

PUBLICATIONS

L. Mokhesi and A. Bagula. *Context-Aware Handoff Decision for Wireless Access Networks using Bayesian Networks*. In Proceedings of the Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists (SAICSIT'09), Pages 104 - 111. ACM International Series, 2009.

L. Mokhesi and A. Bagula. *Personalised Handoff Decision for Seamless Roaming in Next Generation of Wireless Networks*. In Proceedings of the 6th International Conference on Network and Service Management (CNSM2010), Pages 362 - 365. IEEE Communications, 2010.

ABSTRACT

The phenomenal emergence of several heterogeneous wireless networks and technologies has allowed users to have access to IP services anywhere, at anytime, from any network and with whatever terminal they use. This computing platform has been driven by the rapid evolution of mobile devices that are equipped with multiple network interfaces and the development of IP based applications. One of the challenging tasks with this Next-Generation Networks (NGN) computing platform is service continuity when users roam around different wireless networks e.g. Wi-Fi, Bluetooth and Cellular networks. This challenge is further elevated when dealing with applications that distribute time continuous data with stringent Quality of Service (QoS) requirements. One of the adaptation methods to ensure service continuity by minimizing flow interruptions when users are mobile is session handoff.

The main contribution of the thesis is to present a handoff support system which implements a handoff decision engine using a Multi-Criteria Decision Making (MCDM) method based on a Bayesian Belief Networks (BBN) and a handoff execution procedure based on buffering and doublecasting techniques. The handoff support system is built around the following features: 1) It utilises a proxy-based middleware architecture, 2) It uses a BBN based MCDM for handoff decision, 3) It is able to represent the full context information which represents the execution environment, 4) It is able to perform decision making under both certainty and uncertainty, 5) It is able to decide correctly on the target network under dynamic context, 6) It performs decision making in the midst of conflicting, interdependent and constraint criteria, and 7) It uses a profile-based handoff decision to offer personalisation to users.

The experimental results showed that when compared with Analytical Hierarchy Process (AHP), the handoff decision method based on BBN performs better on: 1) Modelling of the handoff decision problem and the full representation of the context information, 2) Decision making under uncertainty, 3) Modelling of constraints and interdependent criteria and 4) Support for user preferences. When evaluating the handoff execution, further results revealed that the underlying handoff management strategies provide service continuity by minimising handoff latency and packet losses.

Keywords: Handoff Decision, Multi-Criteria Decision Making, Handoff Management, Service Continuity

TABLE OF CONTENTS

1 Introduction **1**
1.1 Background Information . . . 1
1.2 Objective . . . 4
1.3 Contribution . . . 5
1.4 Scope and Limitation . . . 5
1.5 Methodology . . . 6
1.6 Thesis Organisation . . . 7

2 Background and Related Work **8**
2.1 Evolution of Wireless Networks . . . 8
2.1.1 Wireless Local Area Network . . . 8
2.1.2 Wireless Metropolitan Area Network . . . 9
2.1.3 Mobile Cellular Networks . . . 9
2.1.4 Personal Area Networks . . . 11
2.1.5 Wireless Sensor Networks . . . 11
2.2 Pervasive and Ubiquitous Computing . . . 12
2.3 Handoff . . . 13
2.3.1 Handoff Process . . . 13
2.3.2 Handoff Types . . . 14
2.3.3 Handoff Management . . . 18
2.4 Mobility Support in Wireless Networks . . . 18
2.4.1 Overview . . . 18
2.4.2 Mobile IP . . . 19
2.5 Context Awareness . . . 20
2.6 Related Work . . . 22
2.6.1 Handoff Management . . . 22
2.6.2 Handoff Decision . . . 24
2.7 Summary . . . 25

3 Multiple-Criteria Decision Making **27**
3.1 Bayesian Belief Networks . . . 28
3.1.1 Influence Diagrams . . . 29
3.1.2 Using Bayesian Belief Networks as a MCDM Tool . . . 30
3.2 Handoff Decision using BBN . . . 31
3.2.1 Building the Network . . . 32
3.2.2 Generating Conditional Probability Tables . . . 33
3.2.3 Differentiated Profiling on Synthetic Criteria . . . 36
3.2.4 Utility Function . . . 37
3.2.5 Handoff Decision Algorithm based on BBN . . . 38
3.3 Creating User Profiles . . . 39
3.4 Dealing with table size for BBN model . . . 41
3.4.1 Reducing the BBN Model for Scalability . . . 42

3.5 Analytical Hierarchy Process 44
3.5.1 Building Criteria Priorities 44
3.5.2 Building Goal Priorities 45
3.6 Summary 47

4 Design and Implementation 49
4.1 Architecture 50
4.1.1 Distributed Systems Architecture 50
4.1.2 Proxy-Based Architecture 50
4.2 System Components 51
4.2.1 Server 51
4.2.2 Proxy 51
4.2.3 Handoff Components 53
4.2.4 Client 54
4.3 Emulation 55
4.3.1 Emulated Environment 55
4.3.2 Wireless Card Emulation 55
4.3.3 Handoff Trigger Emulation 56
4.4 Implementation 56
4.4.1 Client 56
4.4.2 Proxy 60
4.5 Summary 67

5 Evaluation 68
5.1 Resources 70
5.2 Performance of the handoff decision algorithm 70
5.2.1 Comparison between BBN and AHP structures 70
5.2.2 Handoff decision correctness in terms of network ranking 74
5.2.3 Handoff decision correctness in terms of network ranking under constraint criteria 79
5.2.4 Handoff decision correctness in terms of network ranking for reduced BBN models 81
5.2.5 Time required for handoff decision completion 83
5.2.6 Time required for handoff decision completion for reduced BBN models 84
5.3 Performance of the handoff management strategies 85
5.3.1 Time required for handoff completion for hard handoff strategy 86
5.3.2 Time required for handoff completion for soft handoff strategy 87
5.3.3 Time required for vertical handoff completion between Bluetooth and Wi-Fi 88
5.3.4 Buffer utilisation 89
5.4 Performance of the overall handoff support system 91
5.5 Summary 93

6 Conclusion 95

References 98

List of Tables

2.1 Context Model for handoff decision 21

3.1 MCDM variables in the handoff decision problem 32
3.2 CPT of node Wi-Fi_Interface Power Consumption 36
3.3 CPT of node Wi-Fi Cost 36
3.4 CPT of node Wi-Fi_User Preferences 36
3.5 Service Differentiation for QoS and SLS 37
3.6 Level Differentiation for Cost and IPC 37
3.7 Part of a sample utility table 38
3.8 Part of the utility table for Profile A 41
3.9 Part of the utility table for Profile B 41
3.10 Part of the utility table for Profile C 42
3.11 Part of the utility table for reduced Profile A 44
3.12 Part of the utility table for reduced Profile B 44
3.13 Cost Priority Matrix 45
3.14 Pairwise Comparison Scale 46
3.15 Cost Priority Vector Determination 46
3.16 Overall Criteria Priority Matrix 46
3.17 Decision vector 47

4.1 sendHandoffTrigger Method 60
4.2 softHandoff Method 60
4.3 Node specification 63
4.4 CPT for node UMTS_Cost specification 63
4.5 Part of the utility table 63
4.6 Compiling the model using Hugin API 64
4.7 Propagating the evidence 65
4.8 Example of case file entries 66
4.9 Retrieving utilities from decision nodes 66

5.1 QoS Priority Matrix 1 72
5.2 QoS Priority Matrix 2 73
5.3 QoS Priority Matrix 3 73
5.4 QoS Priority Matrix 4 73
5.5 Goal Priority Matrix for Profile A1 75
5.6 Goal Priority Matrix for Profile B1 75
5.7 Utilities from BBN(Profile A) and AHP(Profile A1) 77
5.8 Utilities from BBN(Profile B) and AHP (Profile B1) 78
5.9 Utilities for BBN (profile C) 80

List of Figures

2.1 Evolution of Wireless Networks 9
2.2 Seamless Roaming in 4G networks 12
2.3 Horizontal vs. Vertical Handoff 15
2.4 Hard vs. Soft Handoff 16
2.5 Micro, Macro and Global Handoff 17
2.6 Components of MobileIP 19

3.1 MCDM Hierarchy 27
3.2 BBN for handoff trigger 29
3.3 Influence Diagram for handoff trigger 30
3.4 Handoff LIMID as a MCDM 31
3.5 Handoff Decision LIMID 33
3.6 Handoff Sequence 33
3.7 Implemented Handoff Decision LIMID 35
3.8 Handoff decision algorithm based on BBN 40
3.9 Reduced handoff decision LIMID for Profile A 43
3.10 Reduced handoff decision LIMID for Profile B 43
3.11 AHP hierarchy for handoff decision 45
3.12 AHP main steps 47

4.1 High Level Middleware Components 50
4.2 Proxy and client interaction 52
4.3 Flow of requests from the client, proxy and server 53
4.4 Handoff decision and execution components 54
4.5 Client class interaction 56
4.6 Handoff decision and execution message flow 59
4.7 Proxy class interaction 61

5.1 AHP hierarchy 71
5.2 BBN hierarchy 72
5.3 Utility table growth vs. Increase in criteria 74
5.4 Expected utilities for BBN(Profile A) and AHP(Profile A1) 76
5.5 Expected utilities for BBN(Profile B) and AHP(Profile B1) 79
5.6 Expected utilities for BBN(Profile A) and BBN(Profile A') 82
5.7 Expected utilities for BBN(Profile B) and BBN(Profile B') 83
5.8 Execution time for both AHP and BBN 84
5.9 Execution times for full BBN model vs. reduced BBN model 85
5.10 Overall Hard Handoff Execution Time 86
5.11 Overall Soft Handoff Execution Time 88
5.12 Overall Soft Handoff Execution Time between Bluetooth and Wi-Fi 89
5.13 Proxy buffer usage during hard handoff 90
5.14 Monitored bandwidth for all networks. 92
5.15 Monitored delay for all networks. 92

5.16 Monitored jitter for all networks. 93

List of Acronyms

1G First Generation of Mobile Telecommunication

2G Second Generation of Mobile Telecommunication

3G Third Generation of Mobile Telecommunication

4G Fourth Generation of Mobile Telecommunication

AP Access Point

AHP Analytical Hierarchy Process

BBN Bayesian Belief Network

CPT Conditional Probability Table

DAG Directed Acyclic Graph

EDGE Enhanced Data Rates for Global Evolution

GA Genetic Algorithms

GPRS General Packet Radio Services

GSM Global System for Mobile communication

HSCSD High-Speed Circuit-Switched Data

HSDPA High-Speed Downlink Packet Access

IP Internet Protocol

IPC Interface Power Consumption

LAN Local Area Network

LIMID Limited Memory Influence Diagram

MAN Metropolitan Area Network

MCDM Multiple Criteria Decision Making

NGN Next-Generation Networks

OSI Open System Interconnection

PDA Personal Digital Assistant

RSS Received Signal Strength

SLS Service Level Specification

UMTS Universal Mobile Telecommunications System

WAN Wide Area Network

WCDMA Wideband Code Division Multiple Access

Wi-Fi Wireless Fidelity

WiMAX Worldwide Interoperability for Microwave Access

WSN Wireless Sensor Networks

1 INTRODUCTION

1.1 Background Information

Mobile devices and wireless network technologies are evolving towards a *universal wireless access* computing model, where users remain connected wherever they are, and have access to services with whatever terminal they use. One of the drivers for the universal wireless access computing model is the Fourth Generation of Wireless Communications (4G). This family of next-generation of wireless systems represents a heterogeneous networking environment with different access network technologies converging into one IP-backbone. These networks however differ in bandwidth, latency, cost and coverage [6]. Currently, there is little or no integration between the existing heterogeneous wireless network technologies and systems. The emerging wireless technologies are complementary to each other; as a result, they serve different needs of the mobile user. For example, wireless local area networks (WLAN) provide high data rates to the users, even though they are limited in coverage. Current and next-generation cellular networks provide wide coverage for mobile users but with decreased data rates. Finally, satellite networks provide worldwide coverage but also with even more decreased data rates.

The complementary nature of these technologies implies that their integration will enable the user to be best connected to the appropriate network depending on their needs [3]. Furthermore, the integration of these disparate wireless technologies will ensure a seamless end-to-end connection to the user. The challenge with the next-generation of wireless networks is to provide service continuity while users roam over these networks.

Service continuity is one of the aspects addressed in mobility management in wireless access networks. Mobility management is the capability of a network to cope with the mobility of users. Mobility can be distinguished into two categories [47]: Seamless mobility and Nomadic mobility. Seamless mobility renders continuous, uninterrupted service while a user moves between different points of attachment of different wireless networks. Nomadic mobility describes the ability of a mobile node to attach to a network at different locations, but does not expect continuous and seamless services. Seamless mobility has two aspects related to it: Reachability and Service Continuity. There have been several studies conducted both in academia and industry in order to deal with the intricacies of seamless mobility. One of the widely known solutions is Mobile IP which will be discussed in more detail in chapter 2. Mobile IP deals with two IP addresses. One address, the home address, is used to identify the device while the second address, called Care-of-Address (CoA) describes the current location of a device. Packets destined to a mobile device are first sent to its home address, where a home agent provides forwarding functionality to the current care-of-address. Mobile IP however concentrates significantly on reachability of the mobile host and does not address effectively service continuity. Service continuity when users roam around different wireless networks e.g. Wi-Fi, Bluetooth, Cellular networks is a challenging task. This challenge is further escalated when dealing with applications that distribute time continuous data with stringent Quality of Service (QoS) requirements such as bandwidth, jitter, delay and packet loss, for example multimedia streaming applications.

One of the adaptation methods to ensure service continuity by minimizing flow interruptions when users are mobile is *session handoff* also referred to as session handover. There are two types of session handoffs: Horizontal handoff occurs when a mobile user moves from one access point (AP) to a target access point of the same network. Vertical handoff occurs when a mobile user with a mobile device that supports multiple interfaces moves between APs of different networks. The primary aim of most handoff solutions is to reduce the handoff latency. These solutions span different layers of the OSI model (i.e. Data link, Network, Transport and Application layer). In general, in application layer handoff support systems, the handoff process can be divided into three main steps handoff detection, handoff decision, and handoff execution.

Handoff detection is the system's recognition that handoff is needed. Previous research and solutions on handoff detection are based only on recognition of disconnections (signal strength), and this can take a significantly long time to discover. Hence, may result into great packet loss which may undermine service continuity when dealing with applications that send continuous data. This thesis proposes that handoff can be triggered by a wider set of context changes; these includes (i) users moving out of network coverage (disconnection), or (ii) network QoS changes to a level unacceptable for the application, or (iii) users entering preferred networks (support for user preference for communication networks)[6].

Handoff decision decides on the target wireless network to connect to. More researches are exploiting context information that represents the deployment scenario to decide on the target wireless network. Context information includes: User devices and their capabilities, user context information, application QoS requirements, user Service Level Specification (SLS), user location, network coverage and network QoS. Recent mobile devices such as laptops, Personal Digital Assistants (PDAs) and smart phones have evolved to support more than one networking interface. This evolution has made it easier to realise the concept of vertical handoff. The primary role of vertical handoff is to redirect the communication stream to a different network interface on the mobile device.

As the context information increases i.e. number of networks, decisions criteria; the decision on which network to choose becomes much more complicated. Thus, a suitable handoff decision algorithm becomes important for heterogeneous network access. Furthermore, the decision model should be able to incorporate user preferences in the decision process whereby users can specify their preferences on certain criteria and networks. Facing multiple criteria during handoff decision, a single criterion can no longer be used to rank the candidate networks [64]. Recent solutions have proposed a Multiple Criteria Decision Making (MCDM) approach to solve the target network selection problem. One of the popular MCDM methods which that been used extensively is the Analytical Hierarchy Process (AHP). As with the other MCDM methods proposed in the literature, the AHP has significant short-comings such as decision making only under certainty. For this reason, this thesis proposes an alternative decision model based on the Bayesian Belief Networks (BBN) MCDM framework.

Lastly, *Handoff Execution* redirects the flow of data to the new wireless network interface. This process involves switching from the current interface that the mobile device is using to the interface selected by the handoff decision. Handoff execution differs significantly from one solution to the other. Most solutions adopt a mobile IP strategy, which involves a home agent and a foreign agent. Recently, proxy-based solutions have emerged whereby there is an introduction of an additional network entity called a *proxy* that facilitates the handoff execution. However, these solutions still suffer from significant packet losses during the client's disconnections and do not deal with the additional measures to ensure that the service provisioning still fits the new deployment scenario. These additional measures are particularly important when dealing with multimedia applications whereby the QoS mapping is required for content adaptation, when moving between networks that differ in bandwidth to avoid network congestion. The thesis employs handoff management strategies outlined in [9] (Soft/Hard Handoff, Data Buffering) for handoff execution.

1.2 Objective

The handoff decision and execution problem in wireless networking still remains a challenge. In order to realise the advent of seamless mobility in the next generation of wireless networking, effective handoff support solutions have to be in place. This thesis focuses mainly on the handoff decision making and execution process. The objective is twofold, firstly to develop a new handoff decision algorithm with the following features:

- Perform handoff decision using multiple criteria. Previous handoff decision solutions only considered one criterion, which is signal strength. In the current generation of wireless networking where context awareness is important, considering one criterion for handoff decision is limited,
- Perform handoff decision under uncertainty. Previous handoff decision solutions only considered handoff decision under certainty. However, in wireless networking, not all context information is available during the decision making process, but yet a decision has to be made under such situations,
- Perform handoff decision in the presence of interdependent, conflicting and constraint criteria. Previous handoff decision solutions treated criteria to be independent of each other. However, in wireless networking, some criteria are interdependent.
- Perform handoff decision under static and dynamic context. Context information in wireless networking is dynamic. A handoff decision that adapts to context changes and makes correct decisions is necessary.
- Support user preferences in the handoff decision process. In 4G systems, personalisation is important where services offered to a user are tailored towards their preferences.

Secondly, provide a handoff execution solution that:

- Minimises handoff latency and packet losses.
- Is flexible and extensible
- Provides additional adaptation to the user's session in the new execution environment after handoff.

The handoff support solution presented in this thesis is based on the controlled handoff decision procedure, which employs a proxy-based middleware support. The proxy which acts on behalf of the resource limited mobile clients is deployed at the edge of the wired network and autonomously mediates the communication between the client and the server. The handoff decision and execution together with the QoS mapping techniques are handled by the proxy. Each client participating in a session has a corresponding proxy instance on the wired network. All networks under one domain are serviced by one proxy.

1.3 Contribution

The thesis re-visits the problem of mobility in wireless networks, particularly zeroing on handoff decision. The main contributions of the thesis are outlined below:

- Firstly, building upon a BBN framework, the work presented in the thesis proposes a novel handoff decision making algorithm for the emerging generation wireless networks with the expectation of making access more ubiquitous.
- Secondly, the thesis describes the extensions made on a previously proposed testbed to achieve multi-criteria aware handoff decision making and use these extensions to evaluate the performance of the proposed algorithm.
- Thirdly, the thesis presents profile-based handoff decision to offer personalisation by creating differentiated service offering for different users based on their preferences.
- Finally, the thesis presents a novel strategy to achieve scalability during handoff decision and execution by mapping the BBN model into smaller profile-based models that reduce BBN complexity while achieving better computational speed.

BBN have been used extensively in expert systems and artificial intelligence for knowledge representation and reasoning under uncertainty. However, to the best of our knowledge at the time of writing this thesis, the work presented in this thesis is the first to use BBN MCDM framework for handoff decision. Most of the MCDM methods used in the literature (see chapter 2) focus on decisions under certainty and furthermore treat decision criteria as independent of each other. This might be in disagreement with wireless networking where during handoffs, some information may be uncertain but yet a decision still needs to be made. Furthermore, handoff decision making may require criteria and their interrelations to be often interdependent. For instance the tolerable delay of an application running on the client relates to the handoff latency of the network to execute the appropriate handoff execution strategy.

1.4 Scope and Limitation

As previously mentioned, this thesis focuses mainly on the handoff decision making and execution process. The complexity of service continuity in wireless networking presents a challenge in building a comprehensive solution. A major challenge is dealing with different types of handoff occurrences (macro, global), which require rebinding and re-addressing which will include challenging incorporations of additional techniques, such as DHCP and DHCP Relay for client node re-configuration. As a result, this thesis focuses only on inter-subnet handoff occurrences. To thoroughly test the effectiveness of the system, an appropriate infrastructure is needed that has several APs from several wireless networks: Wi-Fi, UMTS/GPRS, and Bluetooth which can be categorised into several domains and subnets. There is currently only one wireless network technology, Wi-Fi, within the Department of Computer Science with three APs. As many other research proposals in this area, emulation is a natural option. Furthermore, due to the diversity of mobile device operating systems and platforms, the client stub was only developed for laptop computers. This limitation can be addressed in future work to cater for other categories of mobile devices such as PDAs, smartphones and other handheld devices.

1.5 Methodology

This research involved developing an experimental handoff support system for wireless access networks. The focus of this work is on developing new handoff decision algorithm which can be deployed in handoff support system regardless of the architecture (i.e. proxy-based and non-proxy based). The handoff decision is part of the handoff process that starts with initiation and ends with execution. This research re-uses a policy-based handoff initiation technique and employs handoff management strategies described in [9].

The experimental handoff support system is evaluated in an emulated environment. The emulation consists of: Wireless card emulator to emulate the three wireless networks under consideration (Wi-Fi, UMTS and Bluetooth), handoff trigger and context changer emulators. The three networks were chosen as they represent today's widely deployed networks. The developed system has the ability to initiate handoff based on varying context changes and perform handoff decision correctly: under uncertainty and using constraint criteria. The system furthermore performs handoff execution using two strategies that minimise packets loss. The system then performs additional content adaptation to make sure the content fits the new environment. Experiments were carried out to evaluate the system. The experiments conducted were grouped into two main categories. The first set of experiments involved evaluating the performance of the new handoff decision algorithm. The aim was to determine the ability to fully represent the handoff problem, the correctness of the decision based on context changes and the overall execution time. The new decision algorithm was compared with an already existing algorithm based on Analytical Hierarchy Process (AHP). The last set of experiments involved evaluating performance of the handoff management strategies. The aim was to determine the overall handoff time required for handoff execution and how the strategies utilise resources.

1.6 Thesis Organisation

The structure of the thesis is as follows: Chapter 2 describes the background and related concepts of the handoff in wireless networking. A discussion and comparison of the available handoff support solutions and the various handoff decision algorithms are presented in the related work. Chapter 3 describes the multi-criteria decision making and the two MCDM methods compared in this thesis. Chapter 4 provides the design and implementation of the experimental handoff support system. Chapter 5 details the experiments conducted to evaluate the new handoff decision method proposed and the overall handoff support system. Chapter 6 details the concluding remarks including the problems and limitations together with further work on the developed handoff support system.

2 Background and Related Work

The past three decades have experienced a phenomenal emergence of several heterogeneous wireless networks. This rapid evolution (refer to Figure 2.1) was driven by increasing demand by users for convergent voice and data services and the need of more advanced internet services such i.e. multimedia to be available to the wireless domain. Furthermore, wireless networking introduced the most important feature of mobility; a process that allows the mobile devices to remain untethered from the wired network, while having access to services offered by the wired network [63]. Mobility is enabled by handoff also referred to as handover from one cell to the other;

The mobility-enabling nature of the wireless internet provides an opportunity for multimedia delivery on the wireless internet domain. The next generation of wireless networks (4G) is characterised by heterogeneous networks that differ in transmission cost, bandwidth and coverage. As part of the handoff support mechanism discussed in this thesis, this chapter first presents an overview of relevant concepts in the evolution of wireless networks, the advent of pervasive and ubiquitous computing, handoff and handoff support systems before discussing related work.

2.1 Evolution of Wireless Networks

Figure 2.1 provides a view of the evolution of wireless networks and technologies from the 1st to the 4th generation. Each generation is distinguished by the transmission speed and the supported applications. The wireless networks that have emerged up-to-date can be categorised as: Wireless LANs, Wireless MANs, Mobile Cellular Networks, Personal Area Networks (PAN) and Wireless sensor networks (WSN).

2.1.1 Wireless Local Area Network

Wireless local area network (WLAN) extends a local area network (LAN) through wireless radio connections which are provided by APs. WLAN are implemented under the IEEE 802.11a,b,g,n

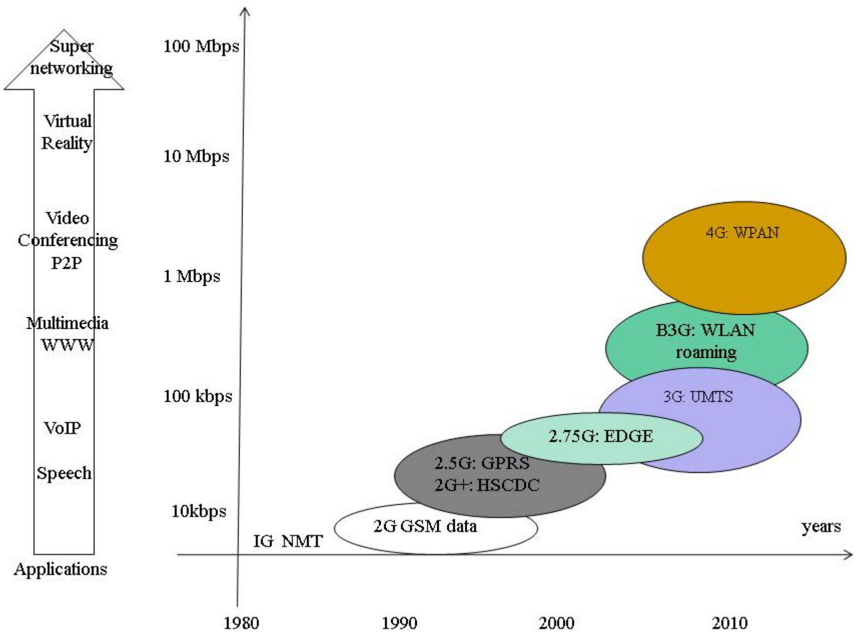

Figure 2.1: Evolution of Wireless Networks

standards. WLAN provides data rates from 2Mbps to 54Mbps with distance range of 50m to 500m for indoor and outdoor usage respectively [12]. Wireless Fidelity (Wi-Fi) is a commonly deployed example of a WLAN.

2.1.2 Wireless Metropolitan Area Network

Wireless Metropolitan Area Network (WMAN) is a type of wireless network that connects several WLANs. An example of a WMAN is Worldwide Interoperability for Microwave Access (WiMAX). WiMAX covers a wider range than Wi-Fi, usually up to 50km with data rate of 40Mbps.

2.1.3 Mobile Cellular Networks

A Mobile cellular network can be characterised as xG, where x refers to the number corresponding to the level and G is the generation. Mobile cellular networks can also be classified as either connection oriented or packet switched. In connection oriented networks, the sender of the information establishes a connection with the receiver before any data can be send. In packet switched networks, a device can send and receive packets without the need to establish a connection with the receiver. The types of cellular networks are [22, 26] :

- **1G** - First generation. 1G refers to the initial category of mobile wireless networks that used only analog technology and were developed primarily for voice services.

- **2G** - Second generation. 2G refers generically to a category of mobile wireless networks and services that use digital technology. 2G wireless networks introduce support for data services. These networks have rather low speed. Examples include:
  - GSM - GSM (Global System for Mobile communication) offers data transmission speeds of up to 9,600 bps.
  - AMPS/DAMPS - sometimes spelled DAMPS, is a digital version of Advanced Mobile Phone Service.
  - HSCSD - High-Speed Circuit-Switched Data (HSCSD) is circuit-switched wireless data transmission for mobile users at data rates up to 38.4 kbps, four times faster than the standard data rates of GSM.
- **2.5G+** - Second generation plus. 2.5G refers generically to a category of mobile wireless networks that have a packet data overlay built on top of the circuit-switched voice network to support higher data rates than 2G mobile networks (2G networks support data in a circuit-switched model). Examples include:
  - GPRS - General Packet Radio Services (GPRS) is a packet-based wireless communication service that promises data rates from 56 up to 114 Kbps and continuous connection to the Internet for mobile phone and computer users. GPRS is based on GSM and complements existing services such circuit-switched cellular phone connections and the Short Message Service (SMS).
  - EDGE - Enhanced Data Rates for Global Evolution, a faster version of the GSM wireless service, is designed to deliver data at rates up to 384 kbps and enables the delivery of multimedia and other broadband applications to mobile phone and computer users. EDGE is a packet-oriented bearer delivering much higher speed than GPRS, anticipating the speed advances of UMTS.
- **3G** - Third generation. 3G refers generically to a category of next-generation mobile networks which operate at a higher frequency bandwidth (typically 2.1 GHz and higher) and have a larger channel bandwidth. This enables 3G networks to support very high data rates, up to 2 Mbps. With the higher bandwidth, more data and multimedia services are possible. 3G refers to the radio network and RF technology, and does not affect the switching core.
- **3.5G** - Third generation Plus. 3.5G refers to an enhancement of 3G which allow networks based on UMTS to have higher data transfer speeds and capacity. High-Speed Downlink

Packet Access (HSDPA) provides the deployment of 3.5G. It is an evolution and improvement on Wideband Code Division Multiple Access (WCDMA) a 3G protocol . HSDPA improves the data transfer rate by a factor of at least five over W-CDMA [62].

- **4G** - Fourth Generation. 4G is also referred to as beyond 3G (B3G) or ”Wireless Broadband”. 4G networks are envisioned to provide better QoS, higher bandwidth and cheaper cost. 4G will be completely packet-switched and based on one IP-Backbone for all networks. 4G includes other network technologies other than cellular i.e. WLAN and Bluetooth in its infrastructure to allow seamless roaming between different network technologies. Figure 2.2 depicts seamless connectivity and global roaming in 4G networks.

2.1.4 Personal Area Networks

A personal area network (PAN) is the interconnection of information technology devices within the range of an individual person, typically within a range of 10 meters. PAN bearers include:

- Bluetooth - Bluetooth is an industrial specification for wireless personal area networks (PANs). Bluetooth provides a way to connect and exchange information between devices such as mobile phones, laptops, PCs, printers, digital cameras, and video game consoles over a secure, globally unlicensed short-range radio frequency [11].

- Infrared - Wireless Infrared is the use of wireless technology in devices or systems that convey data through infrared (IR) radiation. Portable devices include laptop computers, personal digital assistants, and portable telephones. Infrared is electromagnetic energy at a wavelength or wavelengths somewhat longer than those of red light. [13].

2.1.5 Wireless Sensor Networks

Recent advances in micro-electro-mechanical systems (MEMS) technology, wireless communications, and digital electronics have enabled the development of low-cost, low-power, multifunctional sensor nodes that are small in size and communicate untethered in short distances. These sensor nodes consist of sensing, data processing, and communicating components [2]. A wireless sensor network (WSN) generally consists of a base station (or ”gateway”) that can communicate with a number of wireless sensors via a radio link. Data is collected at the wireless sensor node, compressed, and transmitted to the gateway directly or, if required, uses other wireless sensor nodes to forward data to the gateway [56].

This rapid evolution of several wireless networks and technologies described above is providing steadily the realisation of the universal wireless access. Two computing paradigms which describe this universal wireless access are described in the next section.

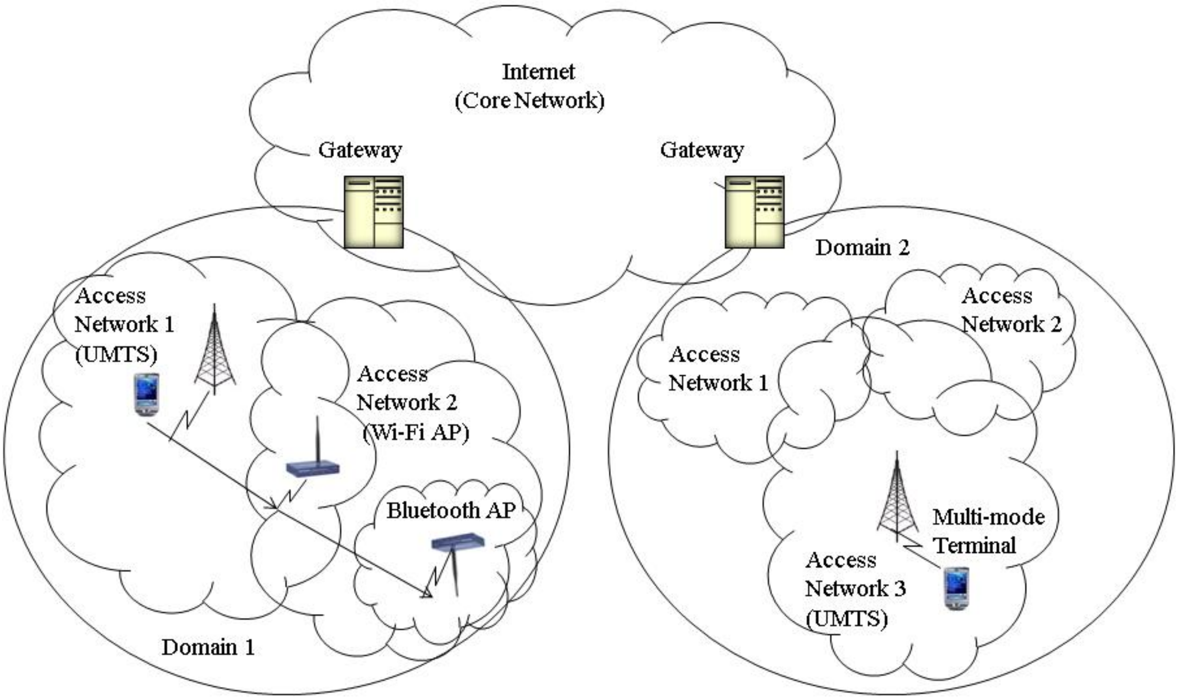

Figure 2.2: Seamless Roaming in 4G networks

2.2 Pervasive and Ubiquitous Computing

Mobile networks are more and more widespread in our daily life. This rapid evolution has slowly provided hope for two types of computing paradigms, namely Ubiquitous Computing and Pervasive computing. These terms have been and are continued to be used interchangeably. However, these two concepts are not the same even though they are described by the same attributes [48]. More commonly, these terms are used to define the internet of things. Ubiquitous Computing was first proposed by Mark Weiser [61] at the Xerox's Palo Alto Research Center in the late 1980s. According to Mark, Ubiquitous Computing names the third wave in computing. Starting by a first generation of mainframes, each shared by lots of people, the evolution led to the personal computing era with person and machine staring at each other across the desktop followed by the ubiquitous computing wave, or the age of calm technology, where technology recedes into the background of our lives. Mark was envisioning a computing model where computing becomes ubiquitous by computerising life as it is by computers that got smaller and smaller and in the end did not look like computers anymore and were everywhere [48]. This scenario could be described as an execution environment whereby people are surrounded by computational resources and tiny networked devices [10]. This could be illustrated by a car navigation system that, by accessing satellite pictures, alerts us to a traffic jam ahead, or an oven that shuts off when our food is cooked [37]. Pervasive Computing on the other hand emerged from IBM in the late 1990s. IBM termed this concept to refer to a computing paradigm that enabled information to be delivered to people anywhere at anytime time when they need it. While one year later Sun Microsystems' pervasive computing philosophy was that ”the computer is the network”, IBM was more focussed on the device and the appliances, such as handhelds, wireless computers, mobile phones [61].

Pervasive computing would make information available everywhere; ubiquitous computing would require information everywhere [25]. Therefore, ubiquitous computing depends on pervasive computing and can only be realised once pervasive computing is in place. For the sake of this document, these terms will be used interchangeably to refer to the universal wireless access, where users remain connected wherever they are, and have access to services with whatever terminal they use. In order to provide this universal wireless access, all wireless networks have to support handoff. The next section describes in detail the concept of handoff in wireless communication.

2.3 Handoff

Handoff occurs when a mobile device in an active session switches from one network access point (AP) to another network access point. This normally results from the transition of a mobile device from the radio range of the current AP into the radio range of the target AP. In wireless networking, an AP works as a bridge between the mobile devices and the internal wired network. The access point relays packets between the mobile device and the network [53].

2.3.1 Handoff Process

The handoff process refers to the exchange of messages between the AP and mobile device leading to the transfer of physical layer connectivity from one AP to the other [39]. During the handoff process, there are at least three participating entities, namely the mobile device, the current AP, which the mobile device is currently connected to before handoff and the target AP, which is the AP that the mobile device seeks to establish connectivity to after handoff.

During the handoff period, the mobile device and the AP exchange a series of management messages. These messages may sometimes include the authentication information specific to the mobile device across to the distribution system. As a result, there is a time window referred to as the handoff latency where the mobile device cannot send nor receive any packets from the AP [39, 55]. The handoff latency is the main challenge in the wireless internet domain especially when dealing with continuous services; applications that distribute continuous data with strict QoS requirements.

At the Media Access Control (MAC) layer, the handoff process can be divided into the following sequence of steps [23]: Discovery, Authentication and Re-association.

1. Discovery: As the mobile device is in motion, the signal strength and the Signal to Noise Ratio (SNR) of the AP that the mobile device is currently communicating with might drop to a level below the handoff triggering threshold. Consequently, the mobile device might lose connectivity with the current AP, hence triggering it to scan for potential APs in its vicinity. This is achieved by the mobile device listening to beacon messages sent periodically by the APs on assigned channels and prioritising them based on the Received Signal Strength (RSS) [39].

   There are two methods of scanning: Active and Passive. In passive scanning, the mobile device scans each channel on the channel list and waits for beacon messages sent by the AP. The beacons are buffered and the information about the APs that sent them is extracted [23]. In active scanning, the mobile device in addition to listening to beacon messages on channels, sends a probe request to the APs.

2. Re-authentication: Once the scanning phase is complete, the mobile device will initiate Re-authentication using the priority list based on RSS. This process involves the transfer of context information which includes the authentication information from the old AP to the new target AP. The authentication can be achieved through a protocol such as Inter-Access Point Protocol (IAPP).

3. Re-association: is the process for transferring associations from one AP to another. Once the mobile device has been authenticated with the new AP, the re-association can be started [35]. Association is a recordkeeping procedure that allows the distribution system to track the location of each mobile device, so that frames destined for the mobile device can be forwarded to the correct AP [23].

2.3.2 Handoff Types

Handoff can be classified into numerous types depending on: i) networks and technologies involved. ii) datalink layer procedure. iii) geography mobility of mobile devices.

When considering the number of networks and technologies involved, there are two handoff types: Horizontal and Vertical Handoff (refer to Figure 2.3)

- Horizontal handoff, also referred to as ”Inter-Technology handover” occurs when a mobile device in an active session switches from one AP to the next AP of the same network technology. E.g. a laptop equipped with a Wi-Fi interface switching between two Wi-Fi APs

- Vertical handoff or ”Inter-Domain handover” occurs when a mobile device that supports multiple wireless network interfaces switches between APs of different network technologies. E.g. a laptop equipped with both a Wi-Fi and 3G interfaces, switching from a Wi-Fi AP to a 3G AP or vice versa.

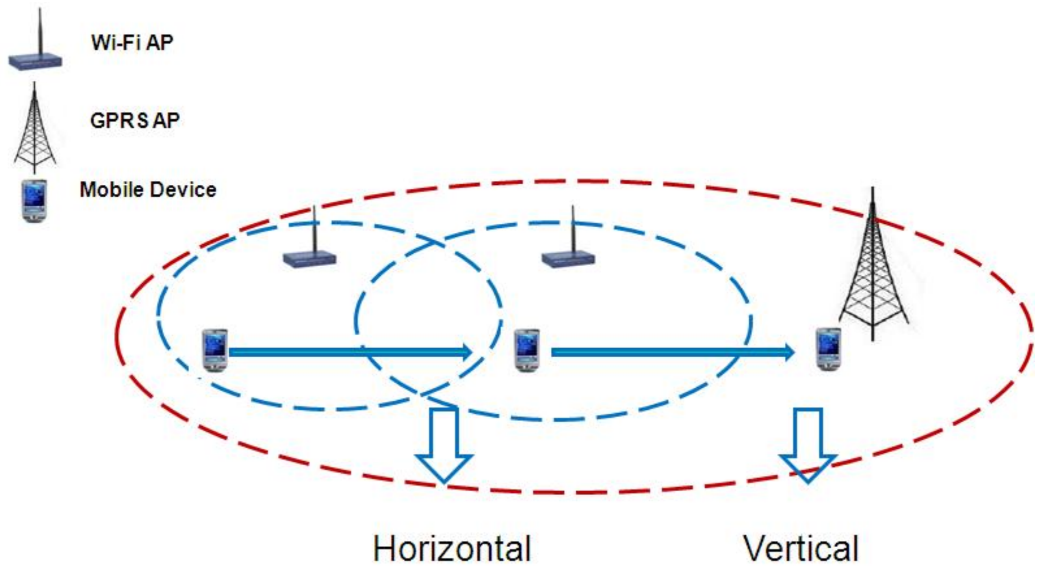

Figure 2.3: Horizontal vs. Vertical Handoff

Horizontal handoff has been fully implemented and is supported by most wireless network technologies. However, horizontal handoff alone is not sufficient to provide the mobile user with the full benefits of the wireless internet. Vertical handoff on the contrary, has not been fully implemented and is not supported by most wireless network technologies. The rapid evolution of several heterogeneous wireless networks and technologies has provided users with a variety of choices between these networks based on their different needs. For example, WLAN provides high data rates to the users, with limited coverage only to a campus, or an enterprise. Current and Next-generation cellular networks such as GPRS and UMTS provide wide coverage for mobile users traversing cities but with decreased data rates. Though there has been a rapid evolution of several heterogeneous wireless networks to date, none of them can provide high bandwidth, low latency, low power consumption and wide area coverage [54]. As a result, support for vertical handoff provides the prospect of integrating these disparate networks to provide users with best available network that meets their needs and also provide a means of roaming seamlessly across these networks.

From the datalink layer perspective (refer to Figure 2.4), handoff can be classified into:

- Hard handoff, also called *Break-before-make* handover, means the mobile device disconnects from the old AP and then connects to the target AP in a relay mode. While this is the simplest kind of handoff, it can also lead to problems of latency and data loss. It is also the most basic case, because, in wireless communications, losing the old AP will generally happen without advance notice.
- Soft Handoff, also called *Make-before-break* handover requires a mobile device that has multiple interfaces to simultaneously connect to both target and old APs. In this way, data can be delivered over both paths at the same time and packet loss can be reduced to virtually zero. The practical problem is that in many mobile systems, devices are limited and cannot connect to two base stations at the same time, for example because they use different frequencies.

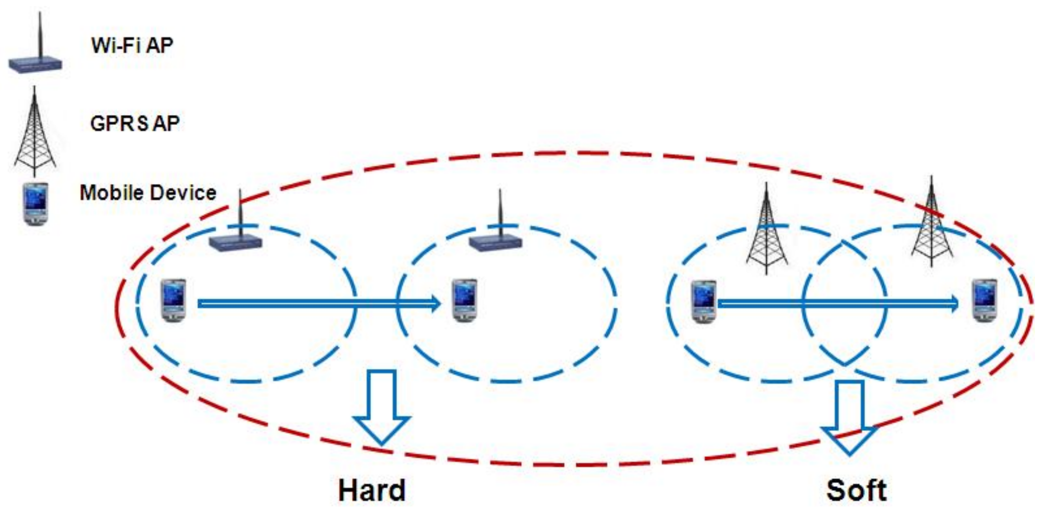

Figure 2.4: Hard vs. Soft Handoff

These definitions usually apply at the datalink layer; nonetheless, they are used in this thesis also in the application layer, to differentiate between solutions that allow multiple use of network interfaces (soft handoff), from those that use only one network interface at a time (hard handoff) [9].

Wi-Fi implements hard handoff, while other networks, such as UMTS, implement soft handoff [8]. This difference in the data link handoff implementation significantly contributes to the challenges of seamless communication across these heterogeneous networks.

The current and future generation of wireless networks cover a large geographical area and or an administrative domain such as a university campus. Each network is usually comprised of several sub-networks (subnets). Handoff occurrences due to the mobility of mobile devices can be categorised as follows [51] (refer to Figure 2.5):

- Micro Handoff - also called inter-subnet handoff. This refers to mobile device that roam around between cells in the same subnet. This mobility implies that the device maintains its IP address.
- Macro Handoff - also called intra-domain handoff. This refers to mobile device that roam around between cells of different subnets. The mobility in this case incurs an IP address change.
- Global Handoff - This refers to the mobile device that roams around cells of different domains. This scenario does not only incur change of the mobile device's IP, but additionally, authentication, authorisation and accounting operations.

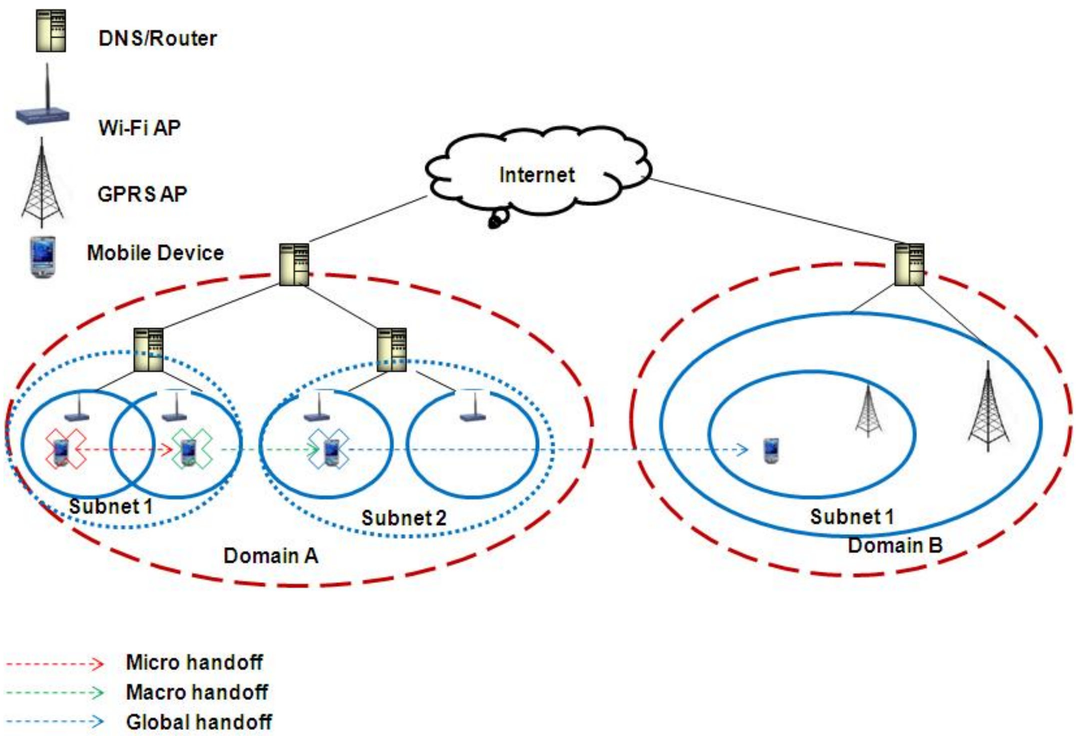

Figure 2.5: Micro, Macro and Global Handoff

In order to manage handoff described above, many solutions spanning different layers have emerged. Handoff management is a key aspect in NGN for developing mobility-aware applications [7]. Handoff management at the application layer will be discussed next.

2.3.3 Handoff Management

In general, at the application layer, handoff management process can be divided into three main steps including handoff initiation, handoff decision, and handoff execution [9, 21]. Handoff initiation is the system's recognition that handoff is needed. Previous solutions on handoff detection are based only on recognition of disconnections (signal strength), and this can take a significantly long time to discover. This may result into great packet loss which may undermine service continuity when dealing with applications that send continuous data. Handoff can be triggered by a wider set of context changes; this includes (i) users moving out of network coverage (disconnection), or (ii) network QoS changes to a level unacceptable for the application, or (iii) users entering preferred networks (support for user preference for communication networks)[6].

Regardless of handoff types, the handoff process control or the handoff decision mechanism can be located in a network entity or in the mobile device itself. The handoff decision usually involves some sort of measurements and information about when and where to perform handoff and obtained from one entity or both. The different handoff schemes are as follows [16]:

- Network-Controlled HandOver (NCHO) - the network entity has the primary control over the handover.
- Mobile-Controlled HandOver (MCHO) - the mobile device must take its own measurement and make the handover decision on its own.
- In Mobile-Assisted HandOver (MAHO), - the information and measurements from the mobile device are used by the network to decide.
- In Network-Assisted Hand-Over (NAHO), - the network collects information that can be used by the mobile in a handover decision.

Handoff management is an aspect of mobility support discussed next.

2.4 Mobility Support in Wireless Networks

2.4.1 Overview

Mobility is the main feature provided by mobile communication and wireless networking. The ability for a mobile station to remain untethered to the wired network yet still have access to the services offered by the wired network has been a significant milestone is wireless networking. However, the issue of mobility support in wireless networking has recently become an important consideration among researchers. This is due to the rapid emergence of heterogeneous wireless networks and technologies together with the evolution of several mobile devices with increased capabilities. The prospect of integrating this heterogeneous wireless networks and technologies together which is driven by the Fourth Generation of Wireless Communications (4G) to provide seamless mobility for the user introduced a challenge for mobility support.

Mobility support contains two distinct but related components: Location management and handoff management. Location management deals with how to locate a mobile node, track its movement, and update the location information. In other words, location management is largely focused on reachability of the mobile node as is roams around different wireless networks. Handoff management on the other hand deals mainly with the control of the change of a mobile nodes AP during active data transmission session [29]. As a result of the increasing demand for mobility support, the Internet Engineering Task Force (IETF) developed Mobile IP.

2.4.2 Mobile IP

Mobile IP is an internet protocol developed by the IETF in order to support mobility of globally mobile users who wish to connect their mobile nodes to the internet and maintain connectivity as they roam around different networks [51]. Mobile IP supports location-independence by maintaining an unchanged home address of the mobile node regardless of the current location of the node, refer to Figure 2.6.

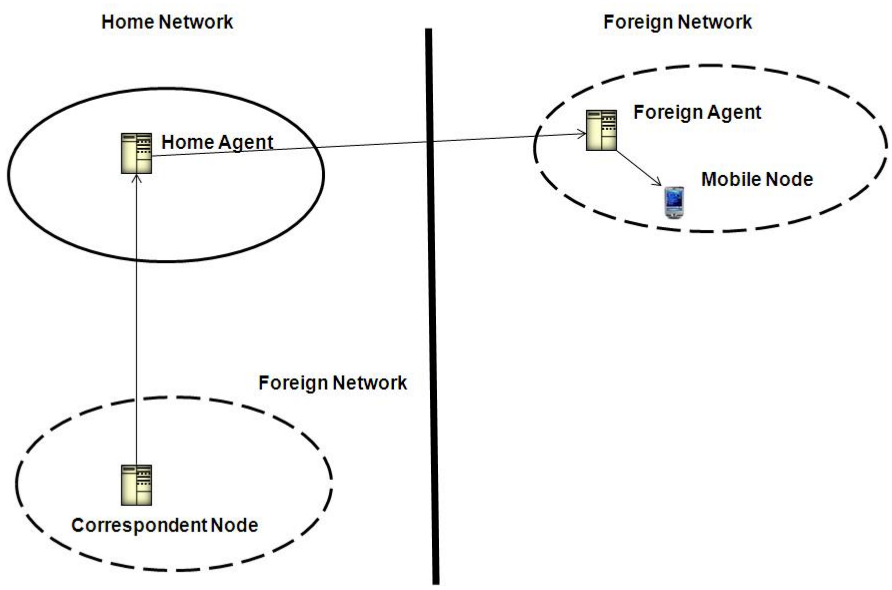

Figure 2.6: Components of MobileIP

With Mobile IP, the mobile node has two IP addresses associated with it [44, 24, 57] : 1) home address, which is the permanent IP address of the node. This IP address remains unchanged. 2) Care-of-Address (CoA) is associated with the mobile node while roaming in a foreign network. Two entities exist in Mobile IP: 1) Home Agent (HA), which maintains a list of registered mobile nodes in the home network and forwards the packets addressed to the mobile node when it is away from home. 2) Foreign Agent (FA), which maintains a list of roaming nodes in its network and communicates with the Home Agent to deliver packets to the mobile node through the mobility bindings of the care-of-address and the home address of the node.

When the mobile node is within the home network, it is being serviced by the Home Agent. Once the mobile node moves out of coverage of its home network into a foreign network, a foreign agent takes over the delivering of packets to the mobile node. Mobile IP employs tunnelling techniques to deliver packets to the mobile node once it's in a foreign network. The HA intercepts all packets intended for the mobile node's home address and forward them to the mobile node using its CoA though the HA. This is normally referred to as Triangular Routing.

The main drawback of Mobile IP as regards service continuity is its lack of proper mechanism for handling handoffs, as all location updates have to go through the HA. This incurs significant packet delivery delay as it increases the round-trip time between the correspondent node and the mobile node. Mobile IP's main focus is on reachability of the mobile node when roaming around different networks. Mobile IP has limitations regarding handoff support. These will be discussed in related work section.

Recent advances in both mobility and handoff support systems have considered context awareness to be important in providing solutions that are aware of the execution environment in order to dynamically adapt to changes in this context [6]. The next section discusses context-awareness and how it is applied in handoff support systems.

2.5 Context Awareness

Context-awareness is a term that emerged after the introduction of pervasive computing. In pervasive computing, it refers to the notion that computers or computer systems should be able to sense or be aware of the situation (context) within their environment, and be able to dynamically react to changes in the context. Context was first defined by [52] as location, identities of nearby people and objects and changes to these objects. In the context of handoff, context information defines the information that can be utilised to support handoffs. Context awareness is highly essential in mobility management particularly handoff management in order to optimise the decision and execution process. Without context-awareness, the handoff management does not yield informed decisions as it is not aware of the execution environment. Implementing a context-aware handoff solution however poses the following challenges [60]:

- The heterogeneity of the wireless networks and technologies e.g. WLANs, 2G and 3G cellular networks, PANs and the upcoming 4G networks.
- The diversity of different applications which differ significantly as regards QoS requirements i.e. real time and non-real time applications.
- The fact that in the next generation of wireless networks, the user should take the centre stage. User preferences should be incorporated into the decision models. Furthermore, the models should be intuitive enough for users to use them with ease
- Location information and mobility information which demands advanced location management.

In order to deal with context management, [6] developed a context model useful in handoff support systems. Context information can be categorised based on how often it changes. The two categories are: static and dynamic context information [6, 46]. Context information that does not change very often or at least not during an active user session is defined as static, on the other hand, context information that changes quite frequently and may lose accuracy over the duration of the user's session is called dynamic [1]. Context information can be gleaned both from the mobile terminal's and the network's side.

Table 2.1: Context Model for handoff decision

Context Type	Terminal Side	Network Side
Static	Device capabilities service types, QoS requirements of services, user preferences	Provider's profile
Dynamic	Running application, type reachable access points (APs)	Current QoS parameters of APs

On the terminal side, context information includes: device capabilities (display size, resolution, battery life, memory, processor speed, and available interfaces). All applications running on the terminal are classified into three application types, namely conversational/real-time, interactive, and streaming applications, where each of them has its own QoS requirements. User preferences are grouped as interface preferences for multimode terminal and service preferences (precedence of service types, expected QoS, and cost constraints). Running application types defines the service profile of currently running applications. Reachable APs identifies currently available networks and addresses of the APs. On the network side, service provider's profiles consist of provider's identity and charging models. Current QoS parameters define the current status of the available network QoS parameters [6, 1].

2.6 Related Work

This thesis focuses on handoff support systems (also referred to as handoff management solutions) placing more emphasis on the handoff decision process. As topics which are closely related to our work, related works on Handoff management and Handoff decision are presented in this section.

2.6.1 Handoff Management

The numerous prospects of the next generation of wireless networks have stimulated a lot of research on Mobility management. Handoff management, a component in Mobility management has attracted significant attention from researchers both from academia and industry. The proposals on handoff management can be categorised according to their target on level on the OSI protocol stack. Different proposals concentrate their solutions on the data-link, network, transport or finally the application layer. It is worth noting that solutions at the datalink, network, and transport layers aim to improve handoff latency and packet loss through static optimization of low-level parameters (i.e. timeouts, number of probes for handoff detection) [9]. However, these lower level solutions are inflexible as they lack the proper context-awareness needed to provide service continuity. They also suffer from QoS violation for time sensitive applications. Furthermore, these solutions are rigid and do not address the current user demands on wireless networking such as user profiling. Lastly, this lower level of concentration does not provide a universal extensible solution as they require protocol stack changes.

[58] is a datalink solution for IEEE 802.11 systems that eliminates the handoff association process between the mobile station and the new AP by transferring the old association states between the mobile station and the old AP to the new AP, therefore allowing a new AP to pretend as the old AP for the mobile station. [3, 34] propose the use of the link layer information to reduce the handoff requirement detection delay. Mobile IP provides mobile support at the network layer. Mobile IP has a problem of triangular routing, and increased network load. Triangular routing furthermore suffers from increased impact of possible network partition, increased load

of on the network due to continuous location update transfer between the HA and FA. Several enhancements on Mobile IP have been proposed, namely Hierarchical MIP (HMIP), Fast handover for MIP(FMIP) and Seamless MIP (S-MIP). HMPI [14] uses a hierarchical network management by introducing an additional entity called Mobile Anchor Point (MAP). The MAP sits at the edge of the network and maintains a binding between itself and mobile nodes currently visiting its network domain. The goal of HMIP is to reduce handoff latency by reducing the network registration time. However, HMIP does not deal with packet losses during disconnections. FMIP employs a proactive scheme to handle mobile node's movement by forwarding packets to the target access router where the mobile node will be connecting at. FMIP however is not QoS and does not provide any adaptation mechanism at the new attachment point. S-MIP [28] combines the concepts of HMIP and FMIP to reduce handoff latency and introduces decision engine. S-MIP has the same problem as FMIP

Recently, more solutions emerged to support handoff at the application layer. [6] is one of the first works to introduce context aware handoff decision on the proxy based handoff management. The proposal has two components: Adaptability manager and a proxy for handoff execution. The proxies are deployed at every network and the Adaptability manager manages a domain which contains multiple networks. The Adaptability manager performs the handoff decision and coordinates with the proxy for handoff execution. This proposal is similar to the one presented in this thesis in that it performs content buffering and doublecasting. The difference is that the double casting is from the adaptability manager to the mobile node and the proxy where to mobile node is migrating to. This approach however is still similar to a single buffering technique as only one interface is active during the handoff execution. As a result, this approach minimises packet loss but not necessarily QoS violation. Furthermore, the proposal does not have any component running on the client device. This could however delay handoff triggering hence incur QoS violation. This thesis is based on handoff support presented in [9]. However, [9] considers only handoff between Wi-Fi and Bluetooth. This work extends the model to include UMTS network. Furthermore [9] does not implement any handoff decision algorithm/engine.

[15] proposes a Universal Seamless Handoff Architecture (USHA) based on the IP tunnelling (IP encapsulation) technique. USHA consists of a handoff server which serves several mobile hosts. An IP tunnel is maintained between the handoff server and every mobile host. All communication is made and bound to the IP tunnel and not to the actual physical interface and packets transmitted through UDP protocol. During handoff, the mobile host changes the physical connection to its new interface and notifies the handoff server, consequently changing the IP tunnel settings. The subsequent packets will be routed to the new physical link. [15]

decreases handoff latency by binding the application to the IP tunnel's interface instead of the physical interface, however, the solution does not provide any packet loss during the clients disconnection with the handoff server.

2.6.2 Handoff Decision

As already outlined above, the introduction of context information into handoff management escalates the handoff problem even further. The emergence of various wireless technologies (3G/UMTS, WLAN, WMAN, etc.), with the evolution of multi-interface Mobile devices has provided users to be best connected wherever they are, with whatever terminal they use as they roam around these heterogeneous networks. New handoff mechanisms do not only consider the received signal strength but include more importantly context information that characterizes the deployment scenario [40]. As the context information increases i.e. number of networks, a Multiple Criteria Decision Making (MCDM) method is required to perform the decision process.

Handoff detection precedes handoff decision. Handoff detection was previously based only on RSS, but recent advancements in wireless networking with integration of diverse technologies has proved that handoff can be triggered by a wider set of context changes. This work has adopted the policy-based handoff triggering based on [6]. There are numerous MCDM techniques that have been utilised for the handoff decision context. The most popular being the Analytical Hierarchy Process (AHP). Other common MCDM methods are: Simple Additive Weighting (SAW) and Technique for Order Preference by Similarity to Ideal Solution (TOPSIS). These methods are described below:

- In SAW, the overall score of an alternative is computed as the weighted sum of all the attribute values.
- AHP on the other hand, develops a goal of hierarchy to solve the decision problem with a large number of attributes. It requires pairwise comparison between alternatives for each attribute in each hierarchy and the consistency check.
- TOPSIS is based on the principle that the chosen alternative should have the shortest distance from the ideal solution and the farthest distance from the negative-ideal solution.

Most MCDM methods including the ones described above are limited as they either consider only decisions under certainty or treat decision criteria as independent of each other or cannot easily model constraints in their decision process. However, in handoff decision making, criteria and their interrelations are often interdependent. Furthermore, some context information is uncertain during the client's active connection and lastly, some criteria such as the application requirements can modelled as a constraint.

AHP has been proposed in numerous researches for target network selection including [6, 1, 30, 43]. The general process includes calculating the objective weights from the pairwise comparison matrix, then calculating each alternative network scoring with respect to each objective. AHP considers decisions under certainty and does not model constraints.

[42] proposes a Fuzzy Multiple Attribute Decision Making Algorithm (FMADM) that selects the optimum target access network by utilising a maximisation of the an objective or fitness function evaluated for each network. The objective/fitness function is later optimised using Genetic Algorithms (GA). The objective function employed here mimics the AHP by using pairwise comparison, hence, suffers from considering decision making only under certainty and does not consider interdependence of criteria. The objective function is further optimised using GA. This process also incurs an additional overhead in terms of time as GA are intrinsically slow, which is undesirable in wireless networking which already suffers from numerous overheads such as: high network delays, intermittent connectivity. The additional optimisation using GA can also undermine service continuity when dealing with time sensitive applications.

[65] employs a constraint cost function for target network selection. The cost function results into a constraint optimisation problem. The function includes a constraint factor which indicates whether a constraint for a particular service requested by the user can be satisfied by the network under evaluation. For instance, a user may request a service with a specified minimum delay and minimum power consumption requirement. If the mobile node has a low battery life, then the power consumption criteria takes a higher weighting than the delay constraint. [65] also suffers from decision making only under certainty and does not model constraints.

2.7 Summary

The emergence of several wireless networks and technologies and the rapid evolution of mobile have made the wireless internet possible. These heterogeneous networks are complimentary and meet different needs of the user. For instance, Wi-Fi provides high data rates but is limited in terms of coverage. 3G cellular networks on the other hand provide more coverage but decreased data rates. This evolution has provided users the prospect of being always connected to the best network wherever they are and with whatever terminal they use. With the prospect of universal access comes the challenge of mobility management as users roam around different networks. Mobility Management deals with location management and handoff management.

Location management focuses on reachability of the mobile node while handoff management deals with service continuity when a mobile node switches APs while roaming around different networks. two types of handoff types exist: Horizontal and Vertical handoff. Horizontal handoff is supported by many wireless networks and topologies. Vertical handoff still remains a challenge in the current and future wireless networks. Different solutions to handoff management exist in literature. These solutions focus on different layers of the OSI protocol stack. Solutions at Lower layers (Datalink, Transport, and Network) are more rigid and require significant change in the protocols stack. Application layer solutions have emerged recently to address the handoff problem. These solutions provide more flexibility and extensibility. Recent research has begun to incorporate context awareness into handoff decision (handoff triggering and target network selection) process. The decision is based on different criteria such as operator's cost, network QoS, client application requirements etc. As the criteria increases, the decision process becomes complicated, hence, requires a MCDM approach. Many MCDM have been proposed in literature spanning fuzzy logic techniques, cost based functions and probabilistic methods. Even genetic algorithm based techniques have been coupled with other cost function methods to optimise the handoff decision process. Because of the limitation of the current MCDM on target network selection, this thesis proposes a handoff decision method based on Bayesian Belief Networks (BBN) MCDM framework. The next chapter discusses multi-criteria decision making and the two MCDM methods compared in this thesis.

3 Multiple-Criteria Decision Making

As previously mentioned, this thesis proposes a handoff support system that employs on a Multi-Criteria Decision Making (MCDM) method based on Bayesian Belief Networks (BBN) for handoff decision, also referred to as target wireless network selection. MCDM is analytic method to evaluate the advantages and disadvantages of decision alternatives on multiple criteria [45]. As depicted by Figure 3.1, a MCDM problem is, in essence, formed into a hierarchy of three elements: the goal/objective, the criteria, and the alternatives. These elements can be pre-

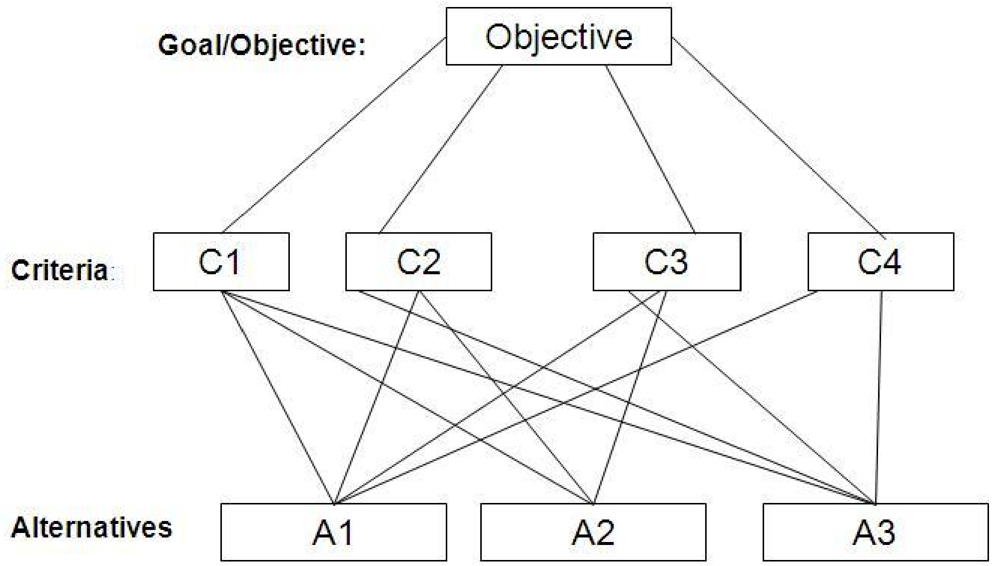

Figure 3.1: MCDM Hierarchy

sented in a matrix format. Let A = $\{a_1,...,a_m\}$ be a set of decision alternatives and C = $\{c_1,...,c_n\}$ a set of criteria according to which desirability of an alternative is measured. A decision matrix D is an m x n matrix, in which element d_{ij} indicates the score or utility of alternative a_i, evaluated against the decision criterion c_j. It is often assumed that the decision maker has determined the weights of relative importance of the decision criteria, W = $\{w_1,...,w_n\}$ [59]. The total score for each alternative is obtained by the following formula:

$$S_i = \sum_j w_j d_{ij}$$

When using the principle of maximum utility, the alternative that has the highest score becomes the preferred choice.

An introduction to BBNs and Influence Diagrams (IDs) is presented section 3.1. Section 3.2 describes the BBN based MCDM method for handoff decision. Section 3.3 describes how personalisation is provided by the use of user profiles. BBNs present a challenge of large table sizes when dealing with complicated domains and criteria in MCDM problem increases. Section 3.4 describes how the BBN model presented in this thesis was reduced for scalability. Section 3.5 presents an alternative MCDM called Analytical Hierarchy Process (AHP) which has been used extensively for handoff decision in literature. A summary of the chapter is included in section 3.6

3.1 Bayesian Belief Networks

The BBN method has been adopted in this thesis because of its ability of knowledge representation, reasoning under uncertainty, reasoning with conflicting criteria and modelling interdependent criteria [59]. BBNs are a form of a directed acyclic graph (DAG) in which nodes correspond to the random variables of interest and the directed arcs represent causal or influential probabilistic relationship between the nodes [27]. It is the existence of such independence represented by arcs and the small set of parents for each node that makes it possible to specify the conditional probabilities and to perform inference [31].

For example, in recent handoff support solutions, handoff can be triggered by many factors, including users moving out of range of the network coverage or QoS dropping below a certain threshold unacceptable to the application running on the mobile terminal. It may be important to know what context change caused the handoff trigger. This scenario can be modelled with a BBN (see Figure 3.2). The BBN consists of three nodes: OutOfRange, QoSDropped, and HandoffTriggered which can all be in one of two states: OutOfRange can be either ”outOfRange” or ”not” - QoSDropped can be either ”dropped” or ”not” - and HandoffTriggered can be either ”yes” or ”no”. The BBN models that there is a causal dependency between OutofRange and HandoffTriggered. This implies that when OutOfRange is in a certain state i.e. ”outofrange” this has an impact on the state of HandoffTriggered. The corresponding conditional probability tables for the BBN nodes are also shown in Figure 3.2. Using Bayes’ rule defined below, the probability of each cause of handoff can be inferred.

$$P(X|Y=y) = \frac{P(Y=y|X)P(X)}{P(Y=y)},$$

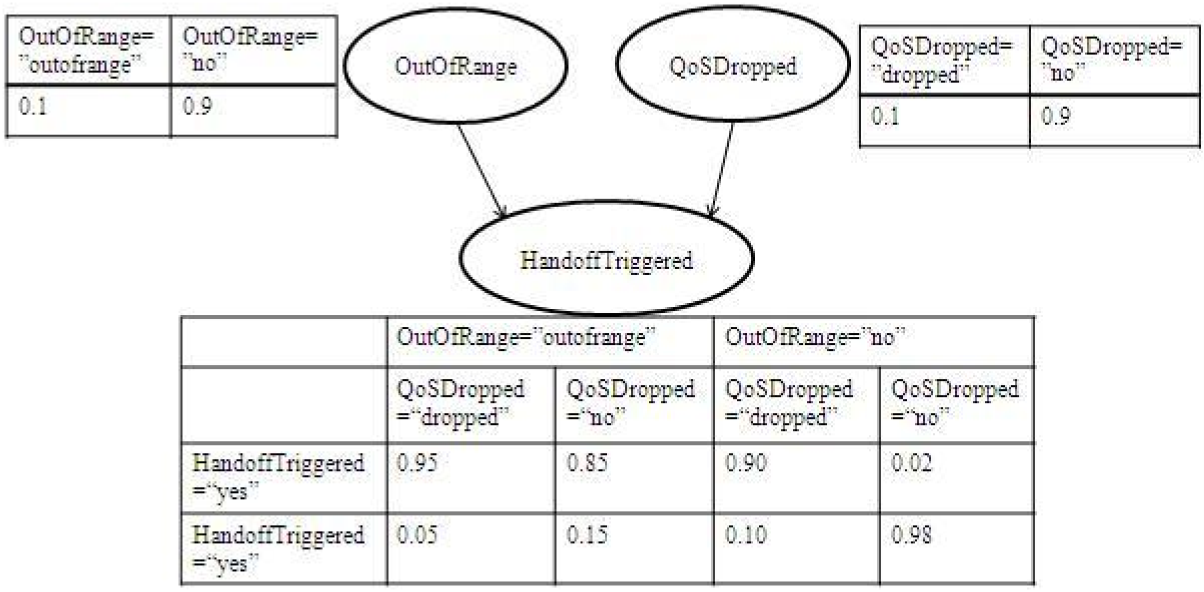

Figure 3.2: BBN for handoff trigger

where P(Y = y) = $\sum_x P(Y = y|X = x)P(X = x)$. X is a cause of Y, where Y often takes the role of an observable effect of X [31].

3.1.1 Influence Diagrams

Using BBN for decision making results into Influence Diagrams (ID). Influence diagrams provide a modelling framework for representation and analysis of decision making under uncertainty [31]. An influence diagram is a BBN that is augmented with decision nodes, representing decision options, and utility or value nodes, representing preferences, that may depend on both random (or chance) variables and decision variables. A decision node defines the action alternatives considered by the user. The primary task of an influence diagram is the determination of the decision alternatives that maximise expected values [17]. There are different types of influence diagrams [31], namely: Discrete influence diagrams, Conditional linear-quadratic Gaussian (CLQG) influence diagrams and Limited-memory influence diagrams (LIMIDs). The work presented here deals with Limited-Memory Influence Diagrams. Limited Memory means that the standard "no-forgetting" rule usually employed in influence diagrams (that is, values of observed variables and decisions that have been taken are remembered at all later times) is relaxed [31, 33, 18].

Using the BBN handoff trigger model in Figure 3.2, a decision can be taken as a result of the handoff trigger. The decision alternatives can be: i) Perform a handoff to a new network if the user is out of coverage or ii) Remain in the current network until the QoS improves. These decisions are based on the preferences of the user and can be altered as desired. Two additional nodes ("handoff" decision node and "cost" utility node) are added as shown in Figure 3.3. The handoff decision node has the states "handoff" and "not". The utility node "cost"

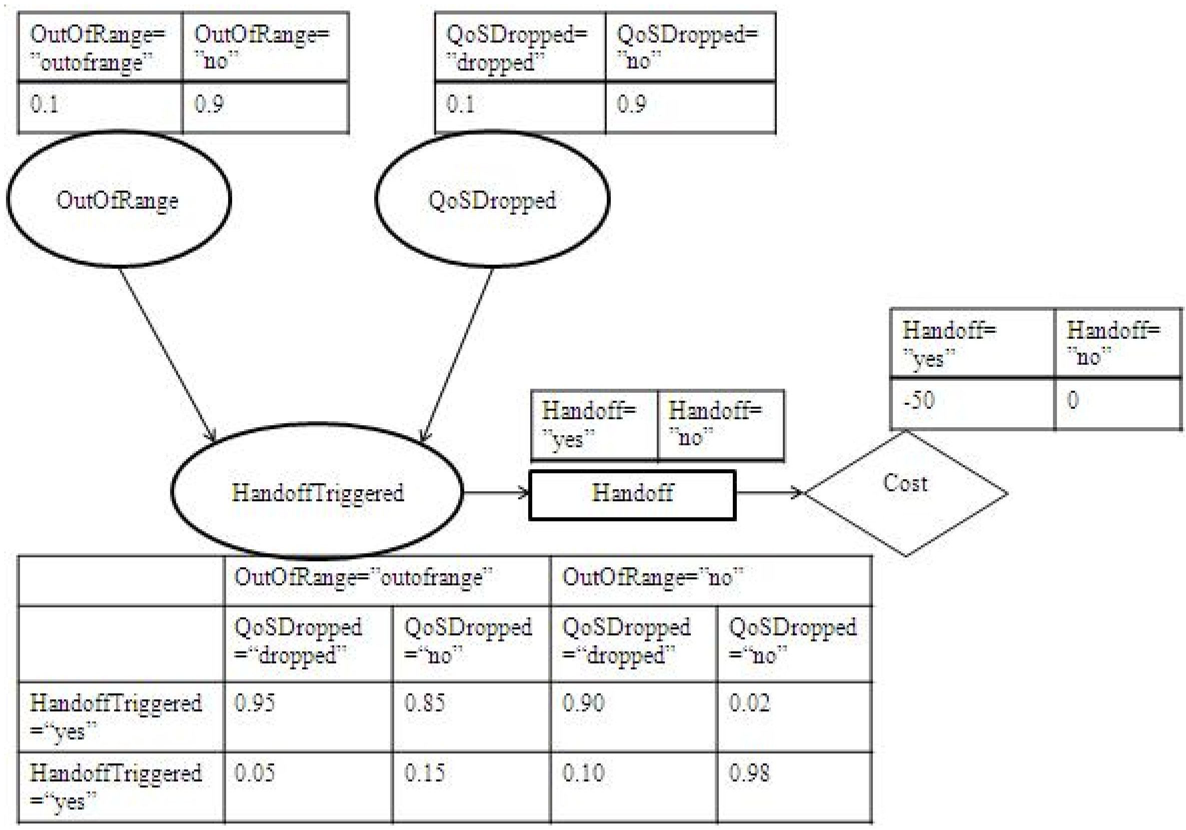

Figure 3.3: Influence Diagram for handoff trigger

gathers information about the cost of the handoff to the new network. To complete the influence diagram, conditional probability table (CPT) for each chance node and the utility table for each utility node should be constructed. A decision node does not have any table. The CPTs will be the same as in Figure 3.2. The utility table for the utility node is shown in Figure 3.3. It denotes that the cost of handoff to a new network is 50cents.

3.1.2 Using Bayesian Belief Networks as a MCDM Tool

When using a BBN as a MCDM tool, the involved variables need to be identified when applying standard MCDM procedure. These include a set of possible actions (set of alternatives), the set of criteria and a set of constraints (properties of the criteria) and other factors that affect them. As a result, from the influence diagram's set of nodes, the MCDM framework will contain three types of nodes: chance node, utility node and decision node. The decision node represents the set of decisions or actions, the utility node represents the set of objectives and the chance nodes represent the criteria and sub-criteria [59]. Figure 3.4 shows how the handoff trigger LIMID model in Figure 3.3 can be represented as a MCDM model. In this example, the goal can be to minimise costs in a user's session, by deciding whether to handoff or not depending on the user moving out of the network coverage or the QoS dropping.

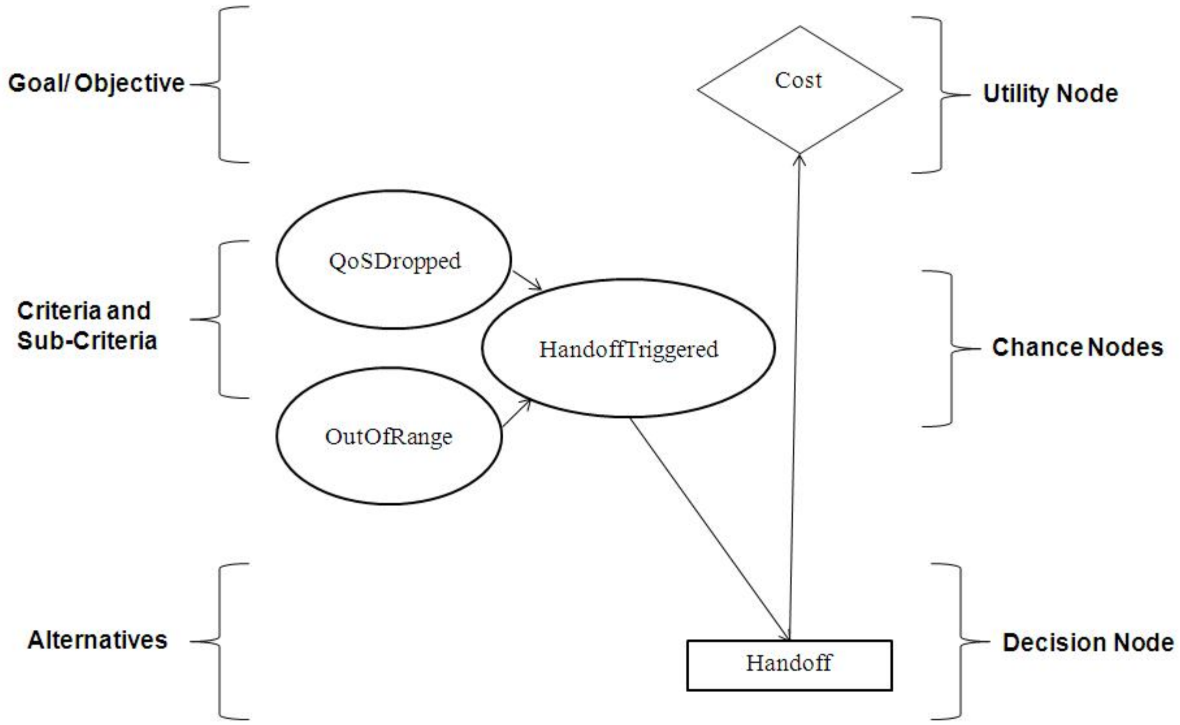

Figure 3.4: Handoff LIMID as a MCDM

3.2 Handoff Decision using BBN

Handoff decision problem presented in this thesis involves making a decision about which target network to handoff to. Context information is used to decide if handoff is required, the target wireless network, handoff management strategy and any additional adaptation methods required after the handoff process. Handoff decision is driven by user preferences (mainly transmission cost and wireless interface power consumption), wireless environment constraints (access network availability and properties and client communication capabilities), and SLS requirements of the application [9].

3.2.1 Building the Network

Building the network involves identifying the variables and the interdependence between them. All the nodes can be classified into objectives, actions, criteria, and constraints as shown in Table 3.1. The identified criteria (User preferences, Network QoS, SLS/Application require-

Table 3.1: MCDM variables in the handoff decision problem

	Handoff Decision Problem
Objectives	To perform a handoff to the target network
Decision Problem	To decide which network to handoff to based on multiple criteria.
Set of Possible Actions	Target Network(Wi-Fi,3G,Bluetooth)
Criteria	User preferences (cost, interface power consumption), Network QoS (bandwidth, delay, jitter), SLS/Application requirements (Application Tolerable delay, jitter, data loss)

ments) are not easily defined. These types of criteria are defined as synthetic criteria. Synthetic criteria are decomposed into lower level attributes that are assumed to be well defined [20]. In BBNs, this type of decomposition is referred to as definitional/synthetic idioms. In definitional/synthetic idioms, a synthetic variable is defined in terms of other variables. A synthetic variable representing a definitional relation could be specified as a deterministic function where the relationship between the concepts is clear, otherwise a mechanism to state the degree to which some combination of parent variables combine to define some synthetic variable is used [41]. Using the latter, all the synthetic criteria were decomposed. For instance, user preferences is decomposed into cost and interface power consumption (IPC), network QoS into bandwidth, jitter and delay, SLS/Application requirements into application tolerable delay, jitter and data loss.

Since all the candidate networks have to be evaluated before a decision is made, then the influence diagram will contain information links from the criteria to the decision node. The resulting LIMID is shown in Figure 3.5. The thesis proposes a four step handoff management process (Figure 3.6): Handoff, Target Network, Handoff Management Strategy and QoS mapping. The ”Handoff”, ”Handoff Management Strategy” and ”QoS mapping” decisions are ruled based hence are not part of the handoff LIMID. The rules for each step are specified informally as follows:

- For ”Handoff” i.e. issue handoff trigger when QoS of the current network drops below an acceptable level to the application.

- For ”Handoff Management Strategy” i.e. employ hard handoff strategy for horizontal handoff and soft handoff strategy for vertical handoff.
- ”QoS mapping” i.e. perform QoS mapping when moving from a high bandwidth network to a low bandwidth network.

The ”Target Network” decision is the only decision considered in the LIMID as its decision is influenced by multiple criteria.

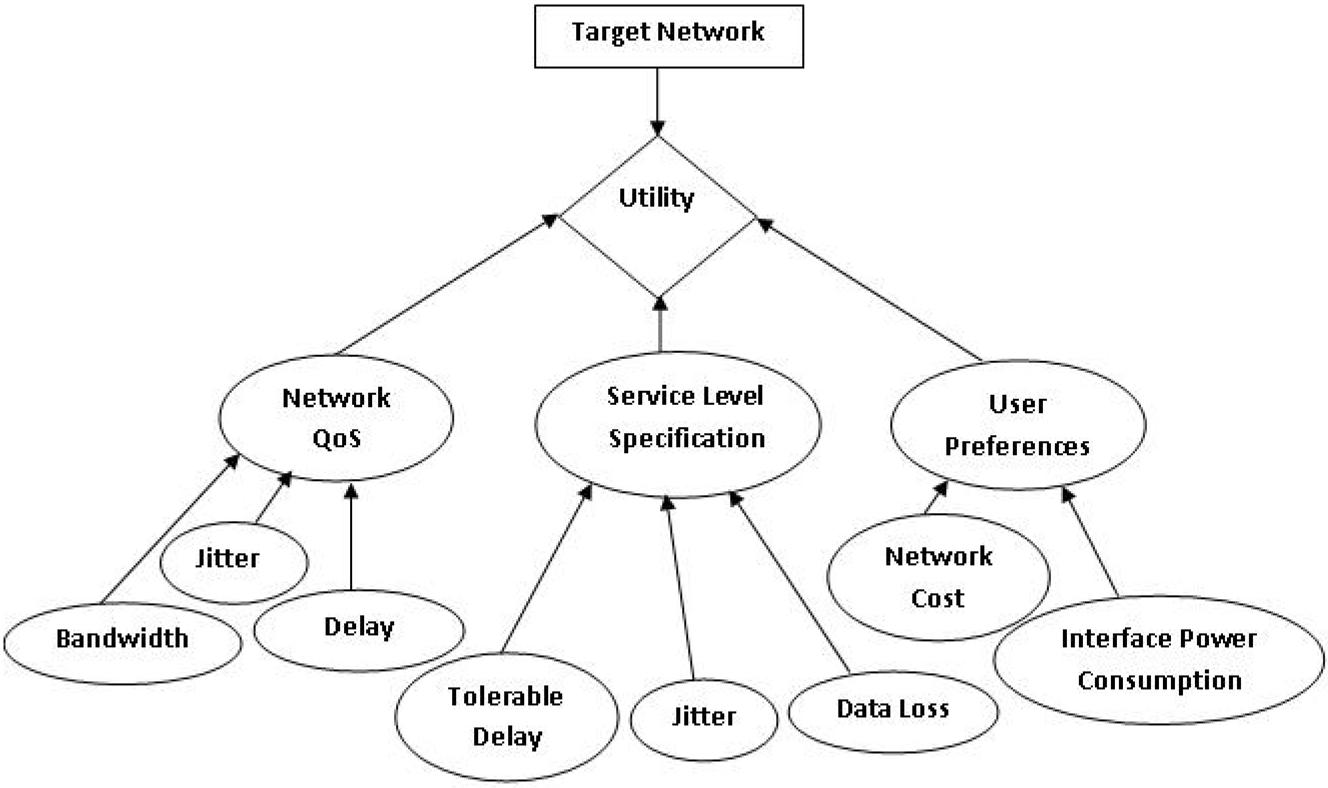

Figure 3.5: Handoff Decision LIMID

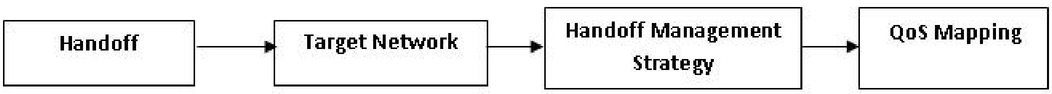

Figure 3.6: Handoff Sequence

3.2.2 Generating Conditional Probability Tables

The contents of a Conditional Probability Tables (CPTs) depend on the problem domain. In some cases, the knowledge can be specified as a mathematical formula relating configurations of parent nodes to states of child nodes in which case the *table generator* may be a useful tool.

In other cases, information on the relationship is available as a database of cases, in which case parameter estimation is used. Parameter estimation incorporates the concept of uncertainty. Finally, if the knowledge is not available as a mathematical formula or data, then subjective expert estimate is used. Subjective estimate also inherits uncertainty. One powerful feature of BBNs is the ability to combine different sources of knowledge such as formulas, data and subjective estimates.

When implementing the BBN LIMID (see Figure 3.7), the two synthetic criteria (network QoS and user preferences) from Figure 3.5 were further decomposed into specific nodes representing all the alternative networks (Wi-Fi,3G,Bluetooth) under consideration.

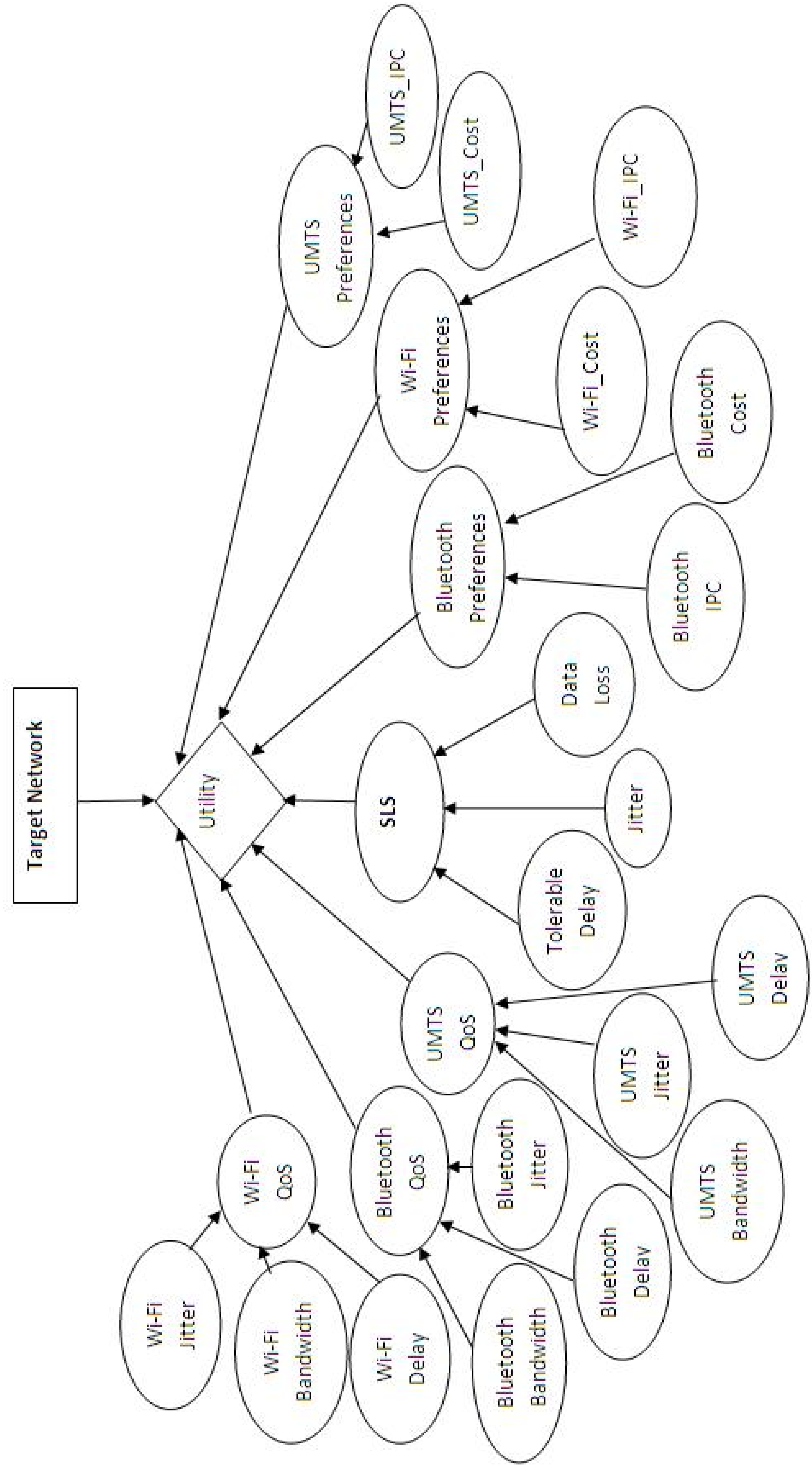

Figure 3.7: Implemented Handoff Decision LIMID

For instance, network QoS was decomposed into (Wi-Fi QoS, Bluetooth QoS and UMTS QoS). User preferences was decomposed into (Wi-Fi User Preferences, Bluetooth User Preferences and UMTS User Preferences). This decomposition was done in order to allow for easy evidence entering and propagation. The SLS criteria was left intact as it is independent and does not relate to any network. The probability tables in this thesis have been constructed using subjective estimate for emulation purposes. In practice, some of the criteria can be measured, hence can be available as a database of cases. For instance, the network cost and QoS parameters can be obtained from the network provider, while the interface power consumption can be obtained from the mobile device. Tables 3.2, 3.3 and 3.4 show the conditional probability tables for Wi-Fi_IPC, Wi-Fi_Cost and Wi-Fi_User_Preferences respectively.

Table 3.2: CPT of node Wi-Fi_Interface Power Consumption

Wi-Fi_Interface Power Consumption	
high	0.75
low	0.25

Table 3.3: CPT of node Wi-Fi Cost

Wi-Fi_Cost	
cheap	0.55
expensive	0.45

Table 3.4: CPT of node Wi-Fi_User Preferences

	Wi-Fi_User Preferences		
Cost	Wi-Fi_Interface Power Consumption	low	high
cheap	low	0.05	0.95
cheap	high	0.6	0.4
expensive	low	0.4	0.6
expensive	high	0.05	0.95

3.2.3 Differentiated Profiling on Synthetic Criteria

The thesis uses a differentiated profiling technique for the states of the synthetic criteria. All the synthetic criteria assume two values (low and high). For network QoS and SLS, a simple technique for mapping user perceived QoS to QoS parameters is used as shown in Table 3.5 as illustrated by [32]. For network cost and IPC, the differentiated levels are shown in Table 3.6.

All synthetic values are deterministic from the combination of lower level attributes. For instance, the user preferences are set to high if both the cost and the interface power consumption are high.

Table 3.5: Service Differentiation for QoS and SLS

Parameter / Service Level	Bandwidth	Delay	Jitter	Application Type
Low	64Kbps - 512Kbps	200ms - 400ms	200ms - 400ms	WWW, Telnet, SMTP
High	> 512Kbps	< 50ms	< 50ms	Video, Audio, Health

Table 3.6: Level Differentiation for Cost and IPC

Parameter / Level	Cost	IPC
Low	Cheap: Low: $< R2/Mb$	$< 40w$
High	Expensive: High: $> R2/Mb$	$> 40w$

3.2.4 Utility Function

The utility node encodes the preferences of the decision maker. As a result, utilities are associated with the state configuration of the network. Since the interest of the decision maker is to make the best decision, the objective decision analysis is to identify the decision options that produce the highest expected utility. In this thesis, the weighting for each state configuration of the network was also subjectively estimated in the utility table. This is another advantage of using BBNs, as the preferences of the decision maker can be directly specified in the utility table. A sample utility table is shown in Table 3.7. When making decisions, the configurations of the network are influenced. Therefore the principle of the maximum expected utility is used. Since there are three alternatives [Wi-Fi, UMTS and Bluetooth] for the target network selection, the utility for each has to be computed. The network that exhibits a higher expected utility is the preferred choice for handoff [40].

Table 3.7: Part of a sample utility table

data = (
45 % **UMTS_Preferences=low** UMTS_QoS=low **Wi-Fi_Preferences=low** Bluetooth_QoS=low **Bluetooth_Preferences=low SLS=low** Wifi_QoS=high Target_Network=**Wi-Fi**
30 % **UMTS_Preferences=low** UMTS_QoS=high **Wi-Fi_Preferences=low** Bluetooth_QoS=low **Bluetooth_Preferences=low SLS=low** Wifi_QoS=low Target_Network=**UMTS**
85 % **UMTS_Preferences**=low UMTS_QoS=low **Wi-Fi_Preferences=low** Bluetooth_QoS=high **Bluetooth_Preferences=low** SLS=low Wifi_QoS=low Target_Network=**Bluetooth**
75 % UMTS_Preferences=low **UMTS_QoS=high** Wi-Fi_Preferences=low **Bluetooth_QoS=high** Bluetooth_Preferences=low **SLS=high** Wifi_QoS=high Target_Network=**Wi-Fi**
55 % UMTS_Preferences=low **UMTS_QoS=high** Wi-Fi_Preferences=low **Bluetooth_QoS=high** Bluetooth_Preferences=low **SLS=high** Wifi_QoS=low Target_Network=**UMTS**
15 % UMTS_Preferences=high UMTS_QoS=low Wi-Fi_Preferences=low **Bluetooth_QoS=low** Bluetooth_Preferences=low **SLS=high Wi-Fi_QoS=high** Target_Network=**Bluetooth**
)

3.2.5 Handoff Decision Algorithm based on BBN

The handoff decision algorithm proposed in this thesis in based on BBNs described above. The algorithm is divided into four stages (refer to Figure 3.8): Pre-configuration, Context/Evidence Entering, Real-time calculation, Decision. The working principle of the algorithm assumes that the BBN LIMID model for the handoff decision has already been built.

Step 1: Pre-configuration

In the first stage, the decision maker specifies: The CPTs for all the criteria, the alternatives for the decision and the utility values/weightings for the preferences. Different methods for specifying the values have already been described in subsection 3.2.2. For the utility values, the decision maker can specify their new values or choose from different user profiles (this will be described in more detail in section 3.3. All these configuration information is fed into the BBN model.

Step 2: Context/Evidence Entering

Once the configuration information has been entered, the BBN model can be used for inference and decision making. The next step is to enter evidence or context information that signifies the execution environment into the model. Context information can be static or dynamic and can be gathered from different entities participating in the handoff scenario. This step can be repeated as desired. This step can also present ”no evidence”, whereby no evidence is entered. ”No evidence” models the concept of uncertainty, whereby the nodes in the BBN are characterised by their posterior probabilities.

Step 3: Real-time calculation

Once the context information has been entered, the BBN model calculates the expected utilities for all the alternative networks and ranks the networks based on the principle of maximum utility.

Step 4: Decision

This last step decides on the target wireless network based on network ranking from step 3.

3.3 Creating User Profiles

One important attribute of a handoff decision solution is the ability incorporate user preferences. One advent of 4G networks and services is to provide personalisation. Deep personalisation is the method used to provide tailored services that are built on the individual preferences of users in a given context, automatically reflecting user needs in a specific situation [10]. In handoff scenario, this means a user can specify a network with certain properties they wish to handoff to based on their own preferences in certain or dynamic context. In order to provide such personalisation, three test user profiles are created using the BBN model to serve three types of users. The profiles are described as follows:

1. Profile A (QoS Oriented) - This profile considers QoS to be the most important criteria for decision making. As a result the target network that has the highest QoS will invariably be the preferred network. This profile is applicable for users who always run critical time-sensitive and high QoS applications such as live multimedia, health and clinical applications. This profile implements a utility function that rewards state configurations that have networks with high QoS. The sample utility table for this profile is shown in Table 3.8.

2. Profile B (Cost Oriented) - This profile on the contrary considers User Preferences (cost and IPC) to be the important criteria for decision making. This means the target network that is low in terms of cost and IPC is a preferred choice. This profile is suited for users who mostly run delay tolerant applications. This profile therefore implements

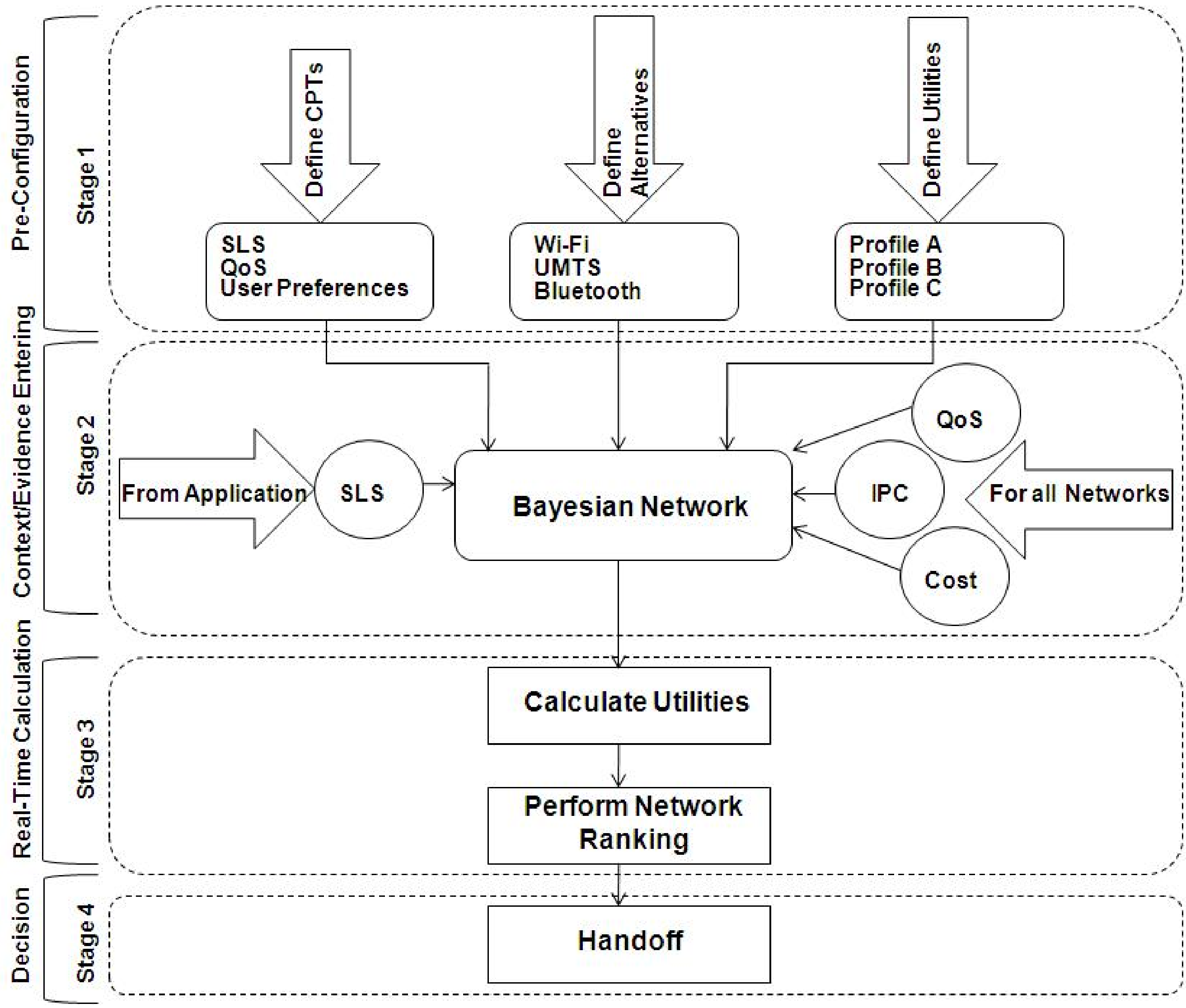

Figure 3.8: Handoff decision algorithm based on BBN

a utility function that rewards state configurations that have networks with low ”User Preferences”. The sample utility table for this profile is shown in Table 3.9.

3. Profile C (SLS controlled) - This profile combines the concepts from profile A and B by introducing constraint criteria, SLS. This profile is suitable for a common user that runs different types of applications that differ in terms of time sensitivity and criticality. This user prefers a high QoS network when they are running a critical and time sensitive application and low cost network when they are running non-critical application. Profile C therefore implements a utility function that rewards state configurations that have networks with high QoS if the SLS is declared to be ”high” else rewards state configurations that have networks with low user preferences if the SLS is declared to be ”low”. The

Table 3.8: Part of the utility table for Profile A

data = (
95 % UMTS_Preferences=low **UMTS_QoS=low** Wi-Fi_Preferences=low **Bluetooth_QoS=low** Bluetooth_Preferences=low SLS=low **Wi-Fi_QoS=high** Target_Network=**Wi-Fi**
80 % UMTS_Preferences=low **UMTS_QoS=high** Wi-Fi_Preferences=low **Bluetooth_QoS=low** Bluetooth_Preferences=low SLS=low Wi-Fi_QoS=low Target_Network=**UMTS**
75 % UMTS_Preferences=low **UMTS_QoS=low** Wi-Fi_Preferences=low **Bluetooth_QoS=high** Bluetooth_Preferences=low SLS=high **Wi-Fi_QoS=low** Target_Network=**Bluetooth**
65 % UMTS_Preferences=low **UMTS_QoS=high** Wi-Fi_Preferences=low **Bluetooth_QoS=high** Bluetooth_Preferences=low SLS=high **Wi-Fi_QoS=high** Target_Network=**Wi-Fi**
55 % UMTS_Preferences=low **UMTS_QoS=high** Wi-Fi_Preferences=low **Bluetooth_QoS=high** Bluetooth_Preferences=low SLS=high **Wi-Fi_QoS=low** Target_Network=**UMTS**
15 % UMTS_Preferences=high **UMTS_QoS=low** Wi-Fi_Preferences=low **Bluetooth_QoS=low** Bluetooth_Preferences=low SLS=high **Wi-Fi_QoS=high** Target_Network=**Bluetooth**
)

Table 3.9: Part of the utility table for Profile B

data = (
55 % **UMTS_Preferences=low** UMTS_QoS=low **Wi-Fi_Preferences=low** Bluetooth_QoS=low **Bluetooth_Preferences=low SLS=low** Wi-Fi_QoS=high Target_**Network=Wi-Fi**
10 % **UMTS_Preferences=high** UMTS_QoS=high **Wi-Fi_Preferences=low** Bluetooth_QoS=low **Bluetooth_Preferences=low** SLS=low Wi-Fi_QoS=low Target_Network=**UMTS**
95 % **UMTS_Preferences=low** UMTS_QoS=low **Wi-Fi_Preferences=low** Bluetooth_QoS=high **Bluetooth_Preferences=low** SLS=high Wi-Fi_QoS=low Target_Network=**Bluetooth**
5 % **UMTS_Preferences=low** UMTS_QoS=high **Wi-Fi_Preferences=high** Bluetooth_QoS=high **Bluetooth_Preferences=low** SLS=high Wi-Fi_QoS=high Target_Network=**Wi-Fi**
45 % **UMTS_Preferences=low** UMTS_QoS=high **Wi-Fi_Preferences=low** Bluetooth_QoS=high **Bluetooth_Preferences=low** SLS=high Wi-Fi_QoS=low Target_Network=**UMTS**
95 % **UMTS_Preferences=high** UMTS_QoS=low **Wi-Fi_Preferences=low** Bluetooth_QoS=low **Bluetooth_Preferences=low** SLS=high Wi-Fi_QoS=high Target_Network=**Bluetooth**
)

sample utility table for this profile is shown in Table 3.10.

3.4 Dealing with table size for BBN model

When dealing with large and complex problems, constructing the conditional and utility tables for the BBN model becomes difficult. This issue is particularly complex when the conditional probabilities and utility weights/values are obtained using subjective estimate. The utility table size grows exponentially with the increase in the number of criteria. The general BBN model presented in this thesis (see Figure 3.5) has three criteria (Network QoS, Service Level Specifi-

Table 3.10: Part of the utility table for Profile C

data = (
45 % **UMTS_Preferences=low** UMTS_QoS=low **Wi-Fi_Preferences=low** Bluetooth_QoS=low **Bluetooth_Preferences=low SLS=low** Wi-Fi_QoS=high Target_Network=**Wi-Fi**
30 % **UMTS_Preferences=low** UMTS_QoS=high **Wi-Fi_Preferences=low** Bluetooth_QoS=low **Bluetooth_Preferences=low SLS=low** Wi-Fi_QoS=low Target_Network=**UMTS**
85 % **UMTS_Preferences**=low UMTS_QoS=low **Wi-Fi_Preferences=low** Bluetooth_QoS=high **Bluetooth_Preferences=low** SLS=low Wi-Fi_QoS=low Target_Network=**Bluetooth**
75 % UMTS_Preferences=low **UMTS_QoS=high** Wi-Fi_Preferences=low **Bluetooth_QoS=high** Bluetooth_Preferences=low **SLS=high** Wi-Fi_QoS=high Target_Network=**Wi-Fi**
55 % UMTS_Preferences=low **UMTS_QoS=high** Wi-Fi_Preferences=low **Bluetooth_QoS=high** Bluetooth_Preferences=low **SLS=high** Wi-Fi_QoS=low Target_Network=**UMTS**
15 % UMTS_Preferences=high UMTS_QoS=low Wi-Fi_Preferences=low **Bluetooth_QoS=low** Bluetooth_Preferences=low **SLS=high Wi-Fi_QoS=high** Target_Network=**Bluetooth**
)

cation, User Preferences). The implemented model (see Figure 3.7) results into 7 criteria which yields a utility table of 256 entries. Handpicking 256 values for the table is cumbersome and highly prone to errors and inconsistencies. There are techniques for reducing the table size such as decomposing tables with introduction of hidden nodes, and hybrid approach that combines deterministic, functional, and probabilistic means [59]. In MCDM models, the number of criteria can be reduced depending on their importance to the overall goal.

3.4.1 Reducing the BBN Model for Scalability

Two of the three test user profiles described in section 3.3 were used reduce the table size since they are specific to one criteria. Profile A only considers networks with high QoS and profile B only considers networks with low ”user preferences”. In both profiles, the unutilised criteria can be removed from the model. The reduced implemented models for the profiles have 3 criteria for the focused profile criteria. Figure 3.8 shows the reduced implemented model for profile A whereby network QoS is the criteria of interest. Figure 3.9 on the other hand shows the reduced implemented model for profile B whereby ”user preferences” is the criteria of interest. The utility table for reduced models contained 24 entries which is manageable and easy to populate by handpicking subjectively estimated utility values. The sample utility tables for both reduced profiles is shown in Table 3.11 and 3.12. Since profile C does not focus on a single criterion, but all the criteria are under consideration, then the profile cannot be reduced for scalability using this technique.

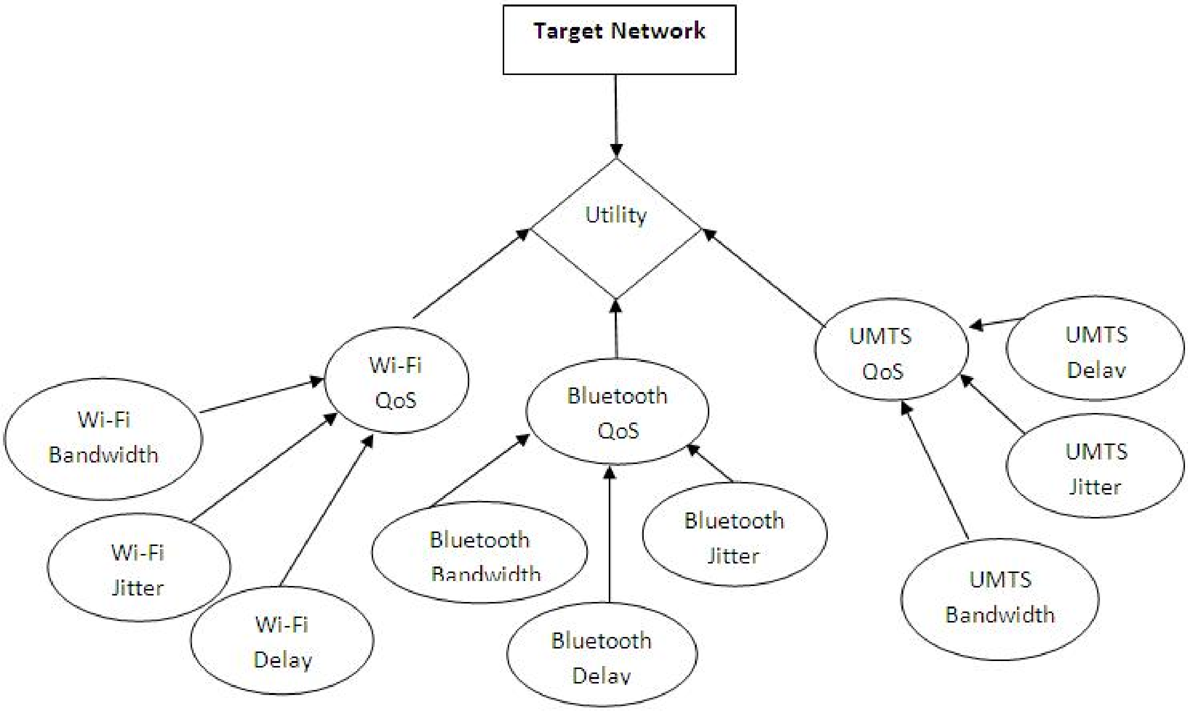

Figure 3.9: Reduced handoff decision LIMID for Profile A

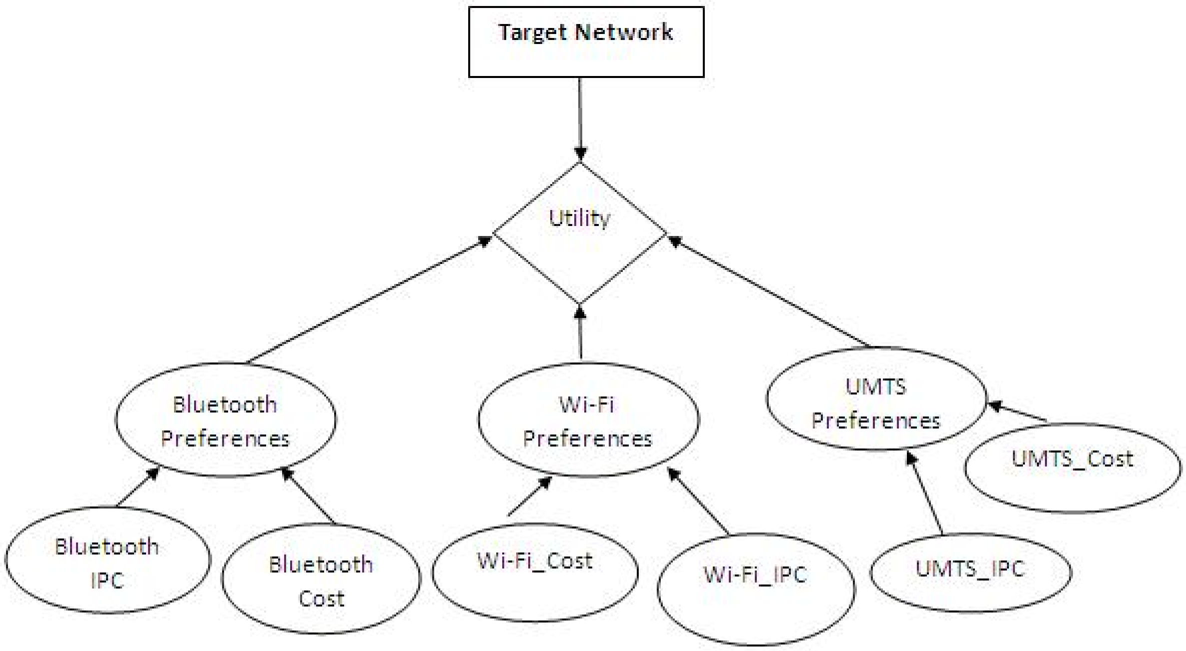

Figure 3.10: Reduced handoff decision LIMID for Profile B

Table 3.11: Part of the utility table for reduced Profile A

data = (
95 % **UMTS_QoS=low Bluetooth_QoS=low Wi-Fi_QoS=high** Target_Network=**Wi-Fi**
80 % **UMTS_QoS=high Bluetooth_QoS=low Wi-Fi_QoS=low** Target_Network=**UMTS**
55 %**UMTS_QoS=high Bluetooth_QoS=high Wi-Fi_QoS=low** Target_Network=**UMTS**
15 % **UMTS_QoS=low Bluetooth_QoS=low** T **Wi-Fi_QoS=high** Target_Network=**Bluetooth**
)

Table 3.12: Part of the utility table for reduced Profile B

data = (
10 % **UMTS_Preferences=high Wi-Fi_Preferences=low Bluetooth_Preferences=low** Target_Network=**UMTS**
5 % **UMTS_Preferences=low Wi-Fi_Preferences=high Bluetooth_Preferences=low** Target_Network=**Wi-Fi**
45 % **UMTS_Preferences=low Wi-Fi_Preferences=low Bluetooth_Preferences=low** Target_Network=**UMTS**
95 % **UMTS_Preferences=high Wi-Fi_Preferences=low Bluetooth_Preferences=low** Target_Network=**Bluetooth**
)

3.5 Analytical Hierarchy Process

The Analytical Hierarchy Process (AHP) is a decision theory developed by [49, 50]. The theory was built on the notion of a human mind. When making a decision, a person groups elements of the problem by certain characteristics which are comparable. AHP follows a MCDM hierarchy structure in Figure 3.1. The AHP hierarchy depends on the problem at hand and varies depending on the judgements, knowledge and needs of the decision maker. In the handoff decision problem for instance, the objective is to decide on the target wireless network depending on cost, coverage, QoS and handoff latency. The wireless networks (Wi-Fi, UMTS and Bluetooth) represent the decision alternatives (Figure 3.11).

3.5.1 Building Criteria Priorities

Once the hierarchy has been constructed, then priorities are built. This is done through the use of pairwise comparison. AHP uses pairwise comparison of components for a decision to produce

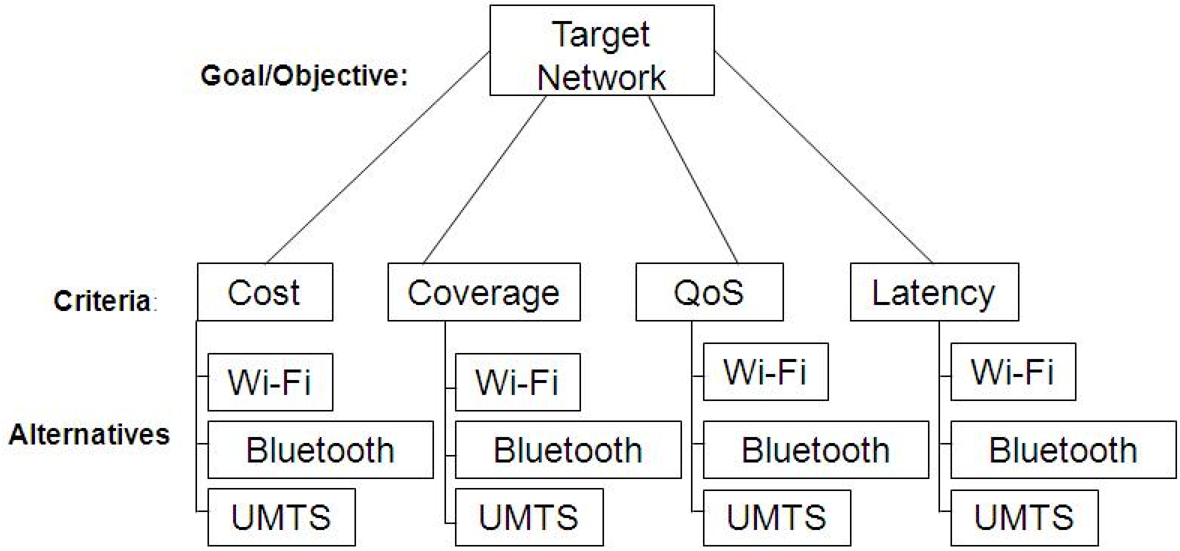

Figure 3.11: AHP hierarchy for handoff decision

preference measures. As a level of the decision hierarchy is addressed through a sequence of pairwise comparisons, an overall preference emerges for each level [19]. Priority is reflected in the weight assigned to factors of a decision or to alternatives of a course of action [19, 50]. Pairwise comparison can be done by creating a priority matrix that represents the alternatives (Wi-Fi, Bluetooth, UMTS) for all the criteria (Cost, Coverage, QoS and Latency). Table 3.13 shows an example of the cost priority matrix. The values of the matrix convey preference information about one network versus another with respect to the cost. AHP uses a comparison scale (Table 3.14) to translate the preference expressions into qualitative values [50]. The same process is repeated for the remainder of the criteria (Coverage, QoS and Latency).

Table 3.13: Cost Priority Matrix

Cost	Wi-Fi	Bluetooth	UMTS
Wi-Fi	1	0.33333	3
Bluetooth	3	1	5
UMTS	0.33333	0.2	1

3.5.2 Building Goal Priorities

Furthermore, a goal priority matrix is created to rank the importance of the criteria relative to the overall goal. The next step is to get the priority ranking from the pairwise comparison matrices. This can be achieved in different ways. One way is to compute a priority vector for each network to criteria priority matrix [19]. The process begins by computing the sum of each column. Each cell is then divided by its column sum and the rows added. The vector

Table 3.14: Pairwise Comparison Scale

Intensity of importance	Definition	Explanation
1	Equal Importance	Two activities contribute equally to the objective
3	Moderate importance	Experience and judgement slightly favour one activity over another
5	Strong importance	Experience and judgement strongly favour one activity over another
7	Very strong importance	An activity is favoured very strongly over another
9	Extreme importance	The evidence favouring one activity over another is of the highest possible order of affirmation

resulting from summing the rows is called the eigenvector (EV), see Table 3.15. The priority vector (PV) is the normalized eigenvector; normalized by dividing each eigenvector element by the elements sum. The same process is repeated for the remaining criteria (Coverage, QoS

Table 3.15: Cost Priority Vector Determination

Cost	Wi-Fi	Bluetooth	UMTS	EV	PV
Wi-Fi	0.125	0.1	0.143	0.368	0.123
Bluetooth	0.375	0.3	0.286	0.961	0.320
UMTS	0.5	0.6	0.571	1.671	0.557

and Latency). This yields the overall criteria priority matrix (Table 3.16) Then finally a simple

Table 3.16: Overall Criteria Priority Matrix

	Cost	Coverage	QoS	Latency
Wi-Fi	0.123	0.087	0.593	0.265
Bluetooth	0.320	0.274	0.341	0.655
UMTS	0.557	0.639	0.065	0.080

matrix multiplication is done between the overall criteria priority vector with goal priority vector to produce the decision vector (Table 3.17). In this case Bluetooth exhibits the highest value. Therefore, Bluetooth is the preferred target network. AHP main steps are summarised in Figure 3.12.

Table 3.17: Decision vector

	PV
Wi-Fi	0.265
Bluetooth	0.655
UMTS	0.08

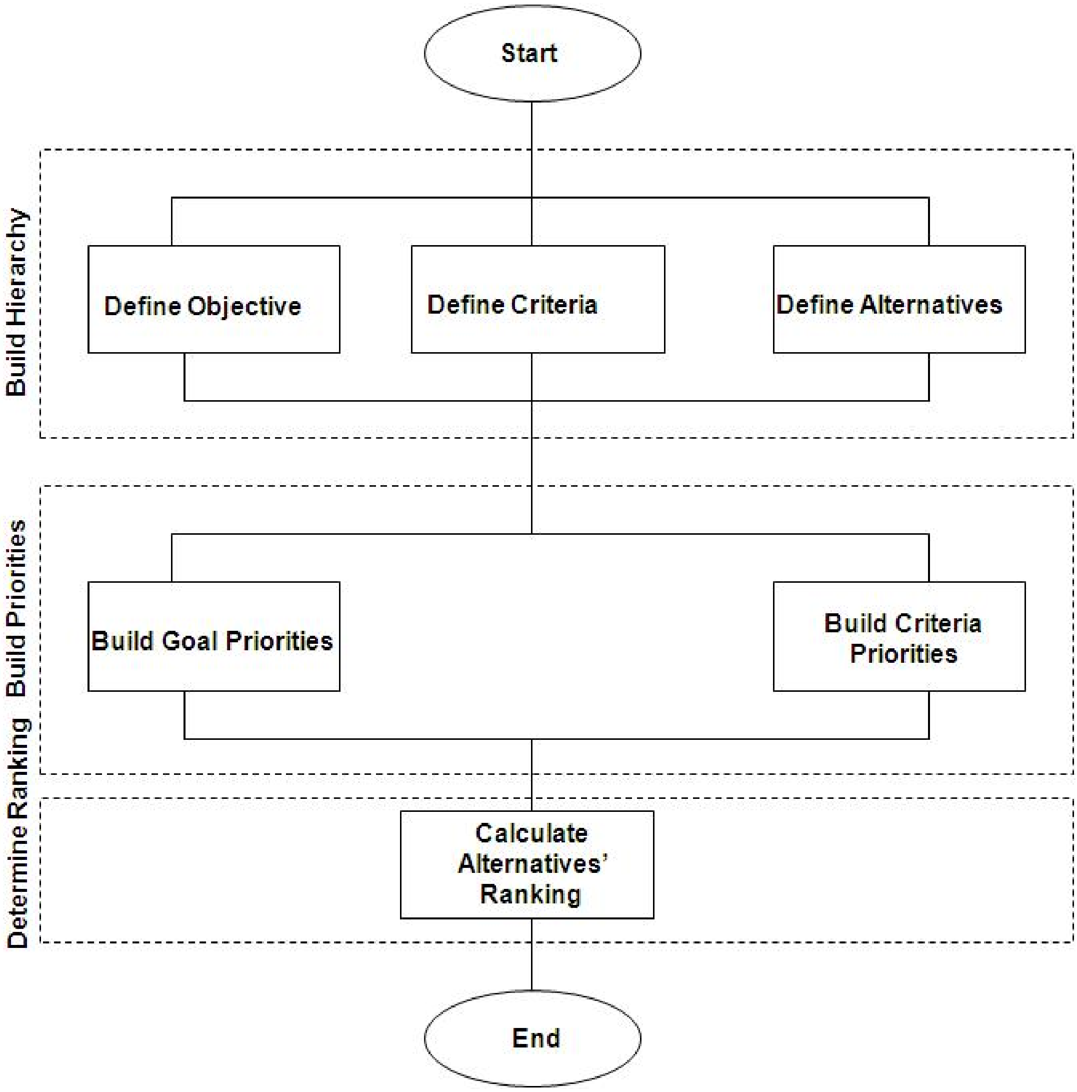

Figure 3.12: AHP main steps

3.6 Summary

The introduction of context information which characterises the deployment scenario in handoff decision has provided more flexibility than the traditional techniques that only considered RSS. Context information is used to decide if handoff is required and the target wireless network. The decision is based on different criteria such as operator's cost, network QoS, client application requirements etc. As the criteria increases, the decision process becomes complicated, hence, requires a MCDM approach. MCDM involves making decisions in the presence of multiple and often conflicting criteria. The thesis proposes a new MCDM algorithm based on BBN. BBN have been used extensively for their ability of knowledge representation, reasoning under uncertainty, reasoning with conflicting criteria and modelling interdependent criteria. BBN can be used to perform inference and for decision making. An influence diagram is a BBN augmented with decision and utility nodes. An influence diagram can easily be used as a MCDM method by matching the IDs nodes into MCDM elements. The handoff decision problem addressed in this thesis has three criteria namely: network QoS, SLS and user preferences. The CPTs for the nodes have been subjectively estimated. The utility node encodes the preferences of the decision maker. Using the different state configurations of the utility table, three user profiles were created to offer personalisation for users. Even though BBN offer many benefits, it presents a challenge when the number of criteria increases. This renders the table size huge and unmanageable. The other commonly used MCDM method is AHP. AHP groups elements of the problem by certain characteristics which are comparable. Then employs pairwise comparison between various alternatives relative to the criteria in the hierarchy. The handoff decision algorithm presented in this chapter is a component of the handoff support system. The next chapter presents the design consideration for the experiment handoff support system and how it was implemented.

4 Design and Implementation

The intricacy of the wireless internet domain especially when dealing with handoff support calls for a solution that can address the heterogeneity of the mobile devices and the wireless technologies involved. The solution should also be flexible enough for it to be extensible. For this reason, the handoff support presented here is realised at the application layer. This level of abstraction offers the needed flexibility and expressiveness and is able to comply with user and application service requirements. [9, 8].

The implementation of the system has been based on re-developing a component of an existing software prototype called Mobile agent based Ubiquitous multimedia Middleware (MUM) [36]. The implementation included re-developing the MUM handoff management component to include:

1. 3G to the list of wireless cards being emulated,
2. Decision engine based on BBN MCDM for deciding the target network and the handoff management strategy (soft or hard handoff),
3. Handoff trigger emulator for handoff initiation and context changing, and
4. Context repository for storing and managing context information.

The experimental handoff support system was implemented in Java and uses some Java technologies: SUN Java Media Framework (JMF) for RTP-based video streaming [38]. The decision engine exploits the Hugin Java API [4]. This chapter presents the design and implementation of the developed experimental handoff support system. Section 4.1 presents the handoff support architecture while the system components, the server, the proxy and the client are presented in sections 4.2, 4.3 and 4.4 respectively. Section 4.5 describes the emulator related to the handoff support while the implementation is presented in section 4.6. A summary of the chapter is included in section 4.7.

4.1 Architecture

4.1.1 Distributed Systems Architecture

The experimental Handoff support system presented here is a distributed middleware infrastructure with components running on fixed hosts in the wired network, *proxy*, as well as on the user's mobile device, client *stub*. The Proxy is deployed on client-to-server distribution path and coordinates with the client stub for handoff decision and execution. The high level components of the system are shown in Figure 4.1.

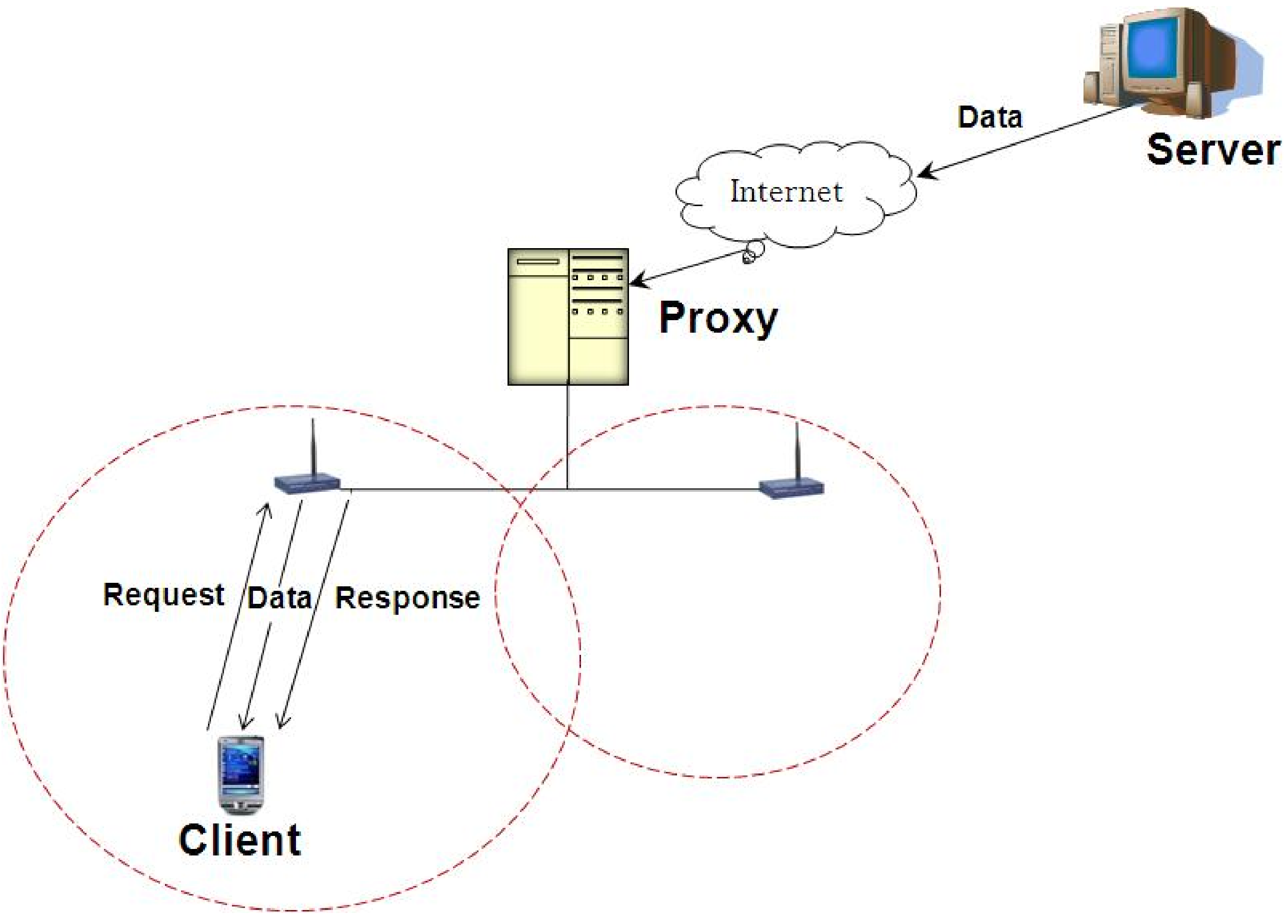

Figure 4.1: High Level Middleware Components

4.1.2 Proxy-Based Architecture

The experimental Handoff support system adopts a proxy-based architecture. This approach can be classified as the mobile-assisted handover (MAHO) solution, whereby the network, in this case the proxy which resides at the wired network collects information and other measurements from the mobile device in order to perform handoff decision and execution procedure.

The introduction of a middle component (proxy) has provided significant benefits in many distributed systems. The current internet is already populated with different kinds of proxies which are used for authentication. caching, re-directing etc. Some of the benefits of proxy-based architectures are outlined below [8]:

- Proxy-based architectures provide an effective alternative to designing and implementing fat clients/servers.
- Proxy-based architectures can act as the middleware glue to extend client/server capabilities with new facilities.
- Proxy-based middleware solutions can enhance the traditional client-to-server interaction model by decoupling the two endpoints with the insertion of an active service path.
- Proxy-based solutions can also reduce client-to-server signalling when handoffs occur.

The architecture has three main entities: Server, Proxy and Client. Each entity in the system has an important role as described below.

4.2 System Components

4.2.1 Server

The server represents any server in the internet domain. This ranges from webs servers, streaming servers, file servers etc. Since the experimental handoff support presented here focuses on multimedia applications, the server herein is a streaming server. The server stores a series of multimedia video and audio files which are transmitted to the client on request. Before the transmission of the multimedia files, the client sends a request to the server through the proxy. The request usually contains the identity of the client (i.e. IP address) and the name of the requested file. Once the request is approved, the server sends the requested media in a continuous manner. Since the server does not form part of the main functionality of the handoff support system, its implementation is not presented.

4.2.2 Proxy

The proxy represents the most important component of the handoff support as it performs much of the management operations. The proxy is logically situated between the client and the server. The proxy manages the client throughout the duration of the client's session. The proxy's functionality can be divided as follows: 1) Management of client's session and delivering content to the client and 2) Handoff decision and execution.

During the client's initialisation, the client sends login request to the proxy. The proxy assigns each client a session ID which is valid for the duration of the session. The proxy then uses the clients ID to identify the client in subsequent request from the client. When the proxy has authenticated the client, it forwards the video request from the client to the server. Before handoff, the client sends a handoff trigger to the proxy, which determines the target network for the client and sends the response back to the client. The proxy can be subdivided in two main parts: A general proxy and a proxy specific for every client (refer to Figure 4.2).

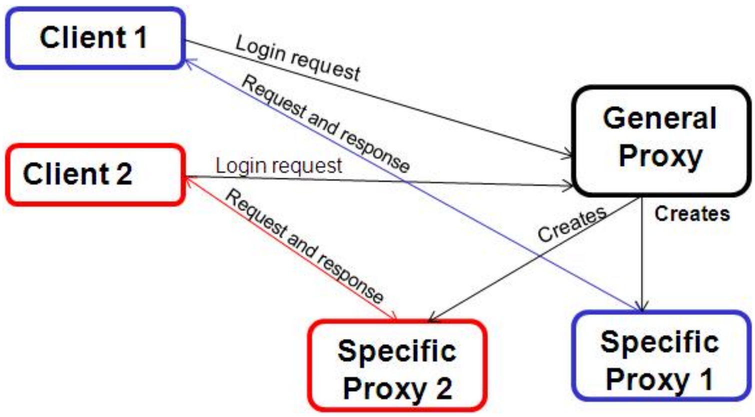

Figure 4.2: Proxy and client interaction

The general proxy creates and manages a specific proxy for every client that requests a service. When a client first logs in, the general proxy creates a specific proxy for it and subsequent responses to requests from that client are delegated to the specific proxy. The specific proxy's ID and clientID are paired and stored in a table. Therefore each request from a client is bundled with the clientID, which is used to search for its specific proxy if it exists and delegates the request to it or creates a new specific proxy for the client if it does not exist. The general proxy is implemented as "MainProxy" class and the specific proxy as "ProxyThread" class.

ProxyThread

As mentioned, the ProxyThread provides the implementation for the specific proxy. The ProxyThread receives a stream request from the client. This request is then forwarded to the server to initiate a streaming session. If the specified stream is available, then the server begins transmitting the stream to the client through the proxy (refer to Figure 4.3).

During the client's session, the ProxyThread receives handoff requests from the client. The ProxyThread then gathers the context information to decide on the target network and the handoff management strategy. The ProxyThread implements a BBN-based decision method dis-

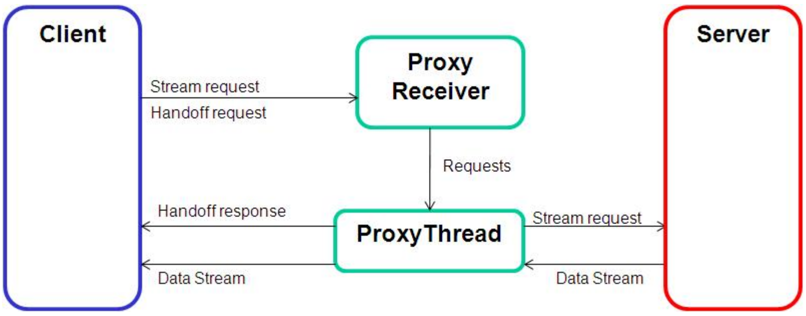

Figure 4.3: Flow of requests from the client, proxy and server

cussed in section 3.1. Furthermore, the ProxyThread implements an internal buffer, for buffering data from the server during the handoff process. The data is then forwarded to the client after the handoff is complete.

Handoff Initiation, Decision and Execution

During the handoff process, the system first has to recognise that handoff is needed, and then evaluate which of the available networks will provide the required QoS (i.e. in the case of QoS sensitive applications). In general, the process is triggered when a user leaves or enters a network coverage or when the QoS of the current network drops below an acceptable level to the application. This behaviour is implemented by the handoff trigger emulator. The handoff initiation is rule-based [6]. The rules determine whether handoff is necessary. The rules are informally described as thus:

- Rule 1: When the user moves out of the current network coverage, determine the target network and handoff to it.
- Rule 2: When the user enters transition zone of a new network/s, determine the target network and handoff to it. This is useful when the user has preferred networks in terms of QoS or network costs.
- Rule 3: When the current network QoS drops below a certain threshold for the application, determine the target network and handoff to it. This is focused on QoS oriented applications and does not apply to delay tolerant applications.

4.2.3 Handoff Components

The components are implemented as functionality of the proxy as described below (refer to Figure 4.1):

- Handoff Decision - This is the decision engine. This process uses the context information to decide if handoff is required, to which AP, and handoff management strategies required. This is implemented in the ProxyThread class (see Figure 4.7).
- Handoff Management Strategy - This process is responsible for executing the handoff management strategy. In the case of hard handoff: a buffering technique is used during the client's disconnection. Hard handoff is implemented by the CircularBuffer class. For soft handoff, multiple network interfaces are activated at the same time and data duplicated and send to them during handoff execution. Soft handoff is implemented by the MultiThreadProxy class.
- QoS mapping - This last procedure ensures that the content is still in suitable format for the new deployment scenario applying content adaptation techniques if necessary. QoS mapping is implemented by MultiThreadProxy class.
- Context repository - Which gathers, manages and evaluates context information. All the components above subscribe to the context repository. This is implemented as a case file (see sub-section 4.4.2)

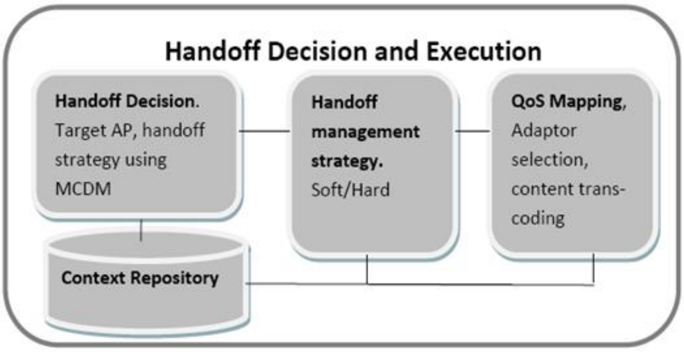

Figure 4.4: Handoff decision and execution components

4.2.4 Client

The main task of the client is to request services from the server through the proxy. Since the experimental handoff support system presented here follows a middleware approach, the client has a component participating in the middleware. The client is divided into two components: The application component and a stub. The application component is responsible for receiving and rendering the content. The stub is responsible for communicating with the proxy and performing handoff operations. These operations include:

- Sending context information to the proxy during the initialisation and subsequently during the client's session.
- Sending handoff request to the proxy and activating/de-activating interfaces on the client during handoffs.

This implies that the client must be aware of its own execution environment in order to perform these management operations.

4.3 Emulation

To thoroughly test the effectiveness of the system, an appropriate infrastructure is needed that has several APs from several wireless networks: Wi-Fi, UMTS/GPRS, and Bluetooth which can be categorised into several domains and subnets. There is currently only one wireless network technology, Wi-Fi, within the Department of Computer Science with 3 APs. As many other research proposals in this area, emulation is a natural option.

4.3.1 Emulated Environment

The Emulator's main task is to produce the information that a normal wireless card would receive if it was in real wireless context. The communication between AP and wireless cards includes an exchange of messages that follow various protocols, for example the initial associations, or the signalling messages (calls beacon) from the AP (refer to section 2.2). The emulator does not necessarily simulate the entire wireless context as this would be complex but limited to handoff signalling and related context transfer.

4.3.2 Wireless Card Emulation

This entity runs on the client's mobile device and emulates the behaviour of wireless network card interface under consideration. It provides an interface similar to the real wireless card and introduces the same behaviour at the login level. In particular the emulator simulates handoff and therefore the loss of connection that results. The wireless cards emulated (Wi-Fi, UMTS, Bluetooth) can be interrogated to supply their information such their IP addresses and the MAC of the AP to which they are currently associated with. The three networks were chosen as they represent today's widely deployed networks.

4.3.3 Handoff Trigger Emulation

The HandoffTrigger class implements the handoff triggering emulator that periodically sends handoff trigger messages to the client by emulating context changes. HandoffTrigger also implements a contextChanger() method. This method periodically writes context changes represented by different state configurations of the utility table into the context repository.

4.4 Implementation

The following section gives in detail the implementation of the experimental handoff support system. It is important at this stage to reiterate that the main focus of this thesis is on the handoff decision and execution, hence only the components that relate to handoff decision will be presented.

4.4.1 Client

This section provides the implementation of the client. Figure 4.5 shows the class interaction for the client. The MainClient operates as the control on the client and coordinates with the proxy for handoff procedure. It gets context changes from the ManagerEventListener. The MainClient uses AckController class to send periodic keep-alive messages to the proxy to indicate to the proxy that it is still alive.

MainClient

MainClient is the main class of the client. The constructor for the class is as follows:

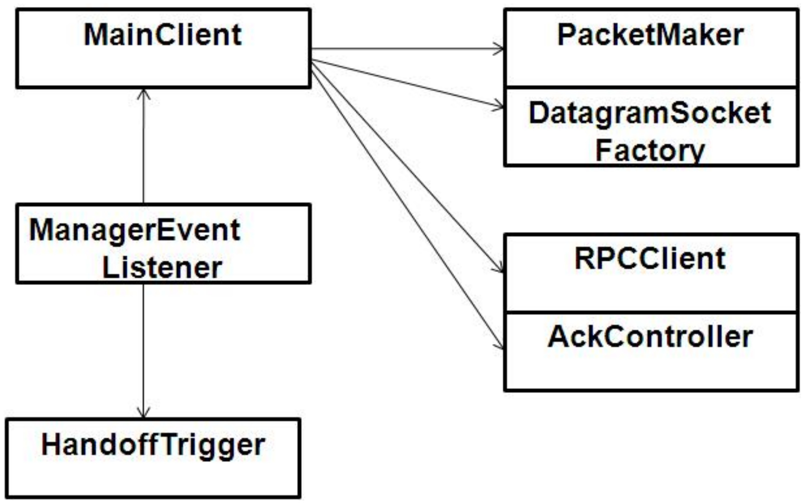

Figure 4.5: Client class interaction

- public MainClient(int clientID, int startInterface, InetAddress proxyAddress, int proxyPort, String wifiInterfaceName, int wifiPort, String btInterfaceName, int btPort, String umtsInterfaceName, int umtsPort,int slsLevel)

The parameters in the constructor are explained below:

- clientID: Identification of the client to the proxy.
- startInterface: This defines the initial interface when the client initialises.
- proxyAddress: The IP address of the proxy.
- proxyPort: The UDP port for the proxy.
- wifiInterfaceName: The name of the Wi-Fi interface on the client.
- wifiPort: The port assigned to the Wi-Fi traffic.
- btInterfaceName: The name of the Bluetooth interface on the client.
- btPort: The port assigned to the Bluetooth traffic.
- umtsInterfaceName: The name of the UMTS interface on the client.
- umtsPort: The port assigned to the UMTS traffic
- slsLevel: The service Level Specification level (high or low) for the application.

The clientID is a parameter used to identify the client to the proxy. Since the general proxy instantiates a specific proxy for every client, the clientID is used to direct communication to the correct client. The constructor obtains the addresses and ports associated with all the wireless network interfaces (Wi-Fi, Bluetooth and UMTS) available at the client. The constructor also initialises the SLS level for the application running. If the SLS level is high, it means that application running has strict QoS requirements. On the contrary, if the SLS level is low, then QoS requirements are not strict. The SLS level informs the proxy of the type of profile the client is currently running for handoff decision purposes.

PacketMaker

The PacketMaker is a class that provides methods for creating various Datagram Packets according to the request of the client to the proxy. The PacketMaker is part of the utilities package. The following methods apply to the handoff process:

- loginRequest(clientID,serviceClass, 10, CLIENT_BUFFER_CAPACITY, slsLevel, proxyAddress,proxyPort);
- handoffTrigger(clientID,interfaceBeforeHandoff,proxyAddress, proxyPort);
- HardHandoffRequest(clientID, interfaceAfterHandoff, interfaceBeforeHandoff, clientNewAddress, clientNewPort, clientNewRTPPort, proxyAddress, proxyPort, TIME_BEFORE_OLD_INTERFAC TIME_BEFORE_HARD_VERTICAL_HANDOFF);
- duplicateStream(clientID, interfaceAfterHandoff, clientNewAddress, clientNewPort, clientNewRTPPort, proxyAddress, proxyPort);

DatagramSocketFactory

The datagram packets carrying requests from the client to the proxy are delivered using a DatagramSocket.

RPCClient

RPCClient is implemented as a listener of events that notify the arrival of data from the RTP protocol and uses the data for rendering. RPCClient has to know the client and proxy addresses and ports in order to transfer the data.

AckController

The AckController sends periodic keep-alive messages to the proxy to notify the proxy that it is still active. AckController is also part of the utilities package.

Handoff Procedure

The execution of a handoff event (horizontal or vertical) is triggered by various context changes. This procedure comprises two main parts: the first part denotes the notification to the proxy of the possibility of handoff. The second involves receiving the response (target wireless network and handoff strategy) from the proxy in order to execute the handoff procedure.

The handoff procedure (refer to Figure 4.6) is begun by the ManagerEventListener which listens to the context changes from the handoffTrigger then it invokes the handoffEvent() method from the MainClient. The method sends a trigger request to the ProxyThread requesting a handoff. The ProxyThread calls the handoffDecision() method to determine the target wireless network after considering the context and sends it back to the client. The client then requests the handoff execution to the proxy. This then alerts the proxy to increase its buffer for hard

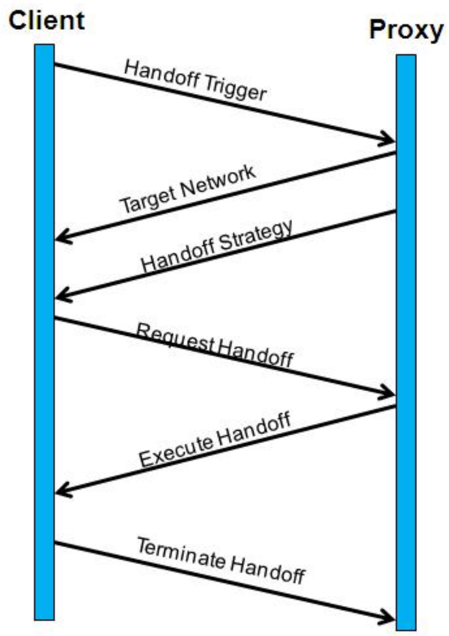

Figure 4.6: Handoff decision and execution message flow

handoff execution by calling the CircularBuffer class or prepare for content duplication for soft handoff by calling the MultiThreadProxy class. This request further updates the proxy with the following information:

- id: Identification of the client
- requestType: Number that identifies the type of request
- clientAddress: New IP address of the client
- clientPort: UDP port for the new DatagramSocket of the MainClient
- clientRTPport: New RTP port of the RPCClient
- interfaceBeforeHandoff: Identification of the running technology on the client

In order to execute the entire handoff procedure on the DatagramSocketFactory on the client creates a temporary DatagramSocket on the interface after handoff. Then it updates the client addresses (UDP and RTP) and sends the request to the proxy in order to start duplicating the stream to the new interface(in the case of soft handoff) or forwarding the buffered stream to the new interface(in the case of hard handoff) . Consequently, the client activates the new interface and creates an RPCClient for it. Table 4.2 considers the case where Bluetooth is the new interface.

Table 4.1: sendHandoffTrigger Method

```
private void sendHandoffTrigger() {

InetAddress proxyAddress = getProxyAddress();
int proxyPort = getProxyPort();

DatagramPacket request =
PacketMaker.handoffTrigger(clientID,interfaceBeforeHandoff,proxyAddress, proxyPort);
try {
sockUDP.send(request);
} catch (IOException ioe) {
System.out.println(sendHandoffTrigger(): error sending message HANDOFF_TRIGGER);
}
receivePacket(); }
```

Table 4.2: softHandoff Method

```
clientNewAddress = btInterface.getAddress();
clientNewPort = btInterface.getPort();
clientNewRTPPort = btInterface.getRTPPort();
btInterface.isActive()
btReceiver = new RPCclient(proxyBt.getAddress(), proxyBt.getRTPPort(),
proxyBt.getPort(), btInterface.getRTPPort(), circularBuffer, timeToWait);
btReceiver.setTimeToWait(timeToWait);
btReceiver.start();
```

4.4.2 Proxy

This section provides the implementation of the Proxy. Figure 4.7 shows the class interaction for the proxy. The MainProxy implements the generic proxy for the system. Its task is to authenticate the client at login and then associate it to the same proxy specific for all subsequent requests.

MainProxy

MainProxy is the main class of the proxy. The constructor for the class is as follows:

- public MainProxy(String wifiInterfaceName, int wifiPort, String btInterfaceName, int btPort, String umtsInterfaceName, int umtsPort, int timeToWait)

The parameters in the constructor are explained below:

- wifiInterfaceName: The name of the Wi-Fi interface associated with the proxy.

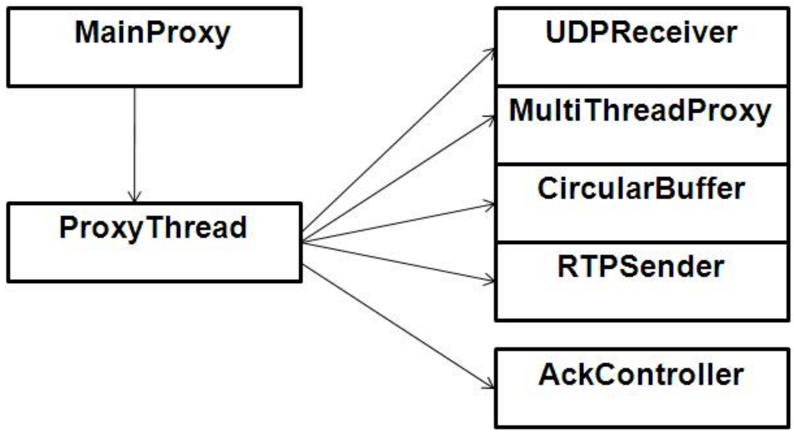

Figure 4.7: Proxy class interaction

- wifiPort: The port assigned to the Wi-Fi traffic.
- btInterfaceName: The name of the Bluetooth interface associated with the proxy.
- btPort: The port assigned to the Bluetooth traffic.
- umtsInterfaceName: The name of the UMTS interface associated with the proxy.
- umtsPort: The port assigned to the UMTS traffic
- timeToWait:

ProxyThread

The ProxyThread implements the proxy specific for each client. It is created for a specific client by the MainProxy when the client first logs on. Its task is to manage all the requests from the client and to perform all handoff decision related operations. UDPReceiver handles the arrival of requests from the client, by listening on a socket. Upon arrival of requests, its forwards them to the ProxyThread which in turn manages them. ProxyThread uses three UDPReceiver instances for the three technologies (Wi-Fi, UMTS and Bluetooth) under consideration.

- ProxyThread(long id, MemoryBroker broker, DatagramPacket packet, int timeToWait, Endpoint proxyWiFi, Endpoint proxyBt, Endpoint proxyUMTS, int firstTechnology)

- id: Identification of the specific proxy.
- broker: Reference to the Proxy Table

- packet: Stream input packet received from the client.
- timeToWait: Maximum time in milliseconds to wait for client request/response before the proxy times out.
- proxyWiFi: The port assigned to the Wi-Fi traffic
- proxyBt: The port assigned to the Bluetooth traffic
- proxyUMTS: The port assigned to the UMTS traffic
- firstTechnology: The initial interface used by the client.

Handoff Decision Engine

Once the proxy receives a handoff trigger from the client, it starts preparing the handoff process. Firstly, the ProxyThread determines the target network by calling a HandoffDecision() method which implements the BBN based handoff decision algorithm. HandoffDecision() reads the context information from the context repository. Then computes the expected utilities for each alternative network and performs the network ranking. Then it sends the response back to the client. To complete the handoff procedure, the ProxyThread calls the method addDestination(clientNewAddress, clientNewRTPPort) from MultiThreadProxy to activate a new stream transmission to the new client interface (in the case of soft handoff) or uses the CircularBuffer for buffering incoming flow then forwards it to the new client address.

The ProxyThread decision engine is built using the BBN. The BBN model and its subsequent processing were implemented by utilizing the widely deployed Hugin API [4].The BBN model presented here contains three criteria that influence the decision. The two criteria (QoS and network preferences) have been expanded to represent the three networks under consideration (Wi-Fi, Bluetooth and UMTS) in order to allow easy evidence entering (see Figure 3.7). The model contains decision nodes, chance nodes and utility nodes. A node is specified as follows:

Secondly, each node has a corresponding conditional probability table. conditional probability tables have been constructed using subjective estimate. Table 4.4 shows an example of UMTS_Cost CPT. This implies that, the probability of the UMTS network being either cheap or expensive is 0.3 and 0.7 respectively. The next step is to create the utility node and its utility table. The utility table defines the preference of the decision maker expressed as utilities which are associated with the state configuration of the network. The resulting utility table contains 256 entries. Part of the utility table configurations are shown in Table 4.5.

Table 4.3: Node specification

```
node UMTS_Cost
{
label = ””;
position = (876 194);
states = (”cheap” ”expensive”);
HR_LinkMode = ”[UMTS_Preferences:0][UMTS_Preference:0]”;
HR_Group = ”0”;
HR_Desc = ””;
HR_State_1 = ””;
HR_State_0 = ””;
}
```

Table 4.4: CPT for node UMTS_Cost specification

```
potential (UMTS_Cost)
{
data = ( 0.3 0.7 );
}
```

Table 4.5: Part of the utility table

data = (95 % UMTS_Preferences=low UMTS_QoS=low Wi-Fi_Preferences=low Bluetooth_QoS=low Target_Network=Wi-Fi Bluetooth_Preferences=low SLS=low Wi-Fi_QoS=low
5 % UMTS_Preferences=low UMTS_QoS=low Wi-Fi_Preferences=low Bluetooth_QoS=low Target_Network=Wi-Fi Bluetooth_Preferences=low SLS=low Wi-Fi_QoS=high
95 % UMTS_Preferences=low UMTS_QoS=low Wi-Fi_Preferences=low Bluetooth_QoS=low Target_Network=Wi-Fi Bluetooth_Preferences=low SLS=high Wi-Fi_QoS=low
95 % UMTS_Preferences=low UMTS_QoS=low Wi-Fi_Preferences=low Bluetooth_QoS=low Target_Network=Wi-Fi Bluetooth_Preferences=low SLS=high Wi-Fi_QoS=low
95 % UMTS_Preferences=low UMTS_QoS=low Wi-Fi_Preferences=low Bluetooth_QoS=low Target_Network=Wi-Fi Bluetooth_Preferences=low SLS=high Wi-Fi_QoS=low
95 % UMTS_Preferences=low UMTS_QoS=low Wi-Fi_Preferences=low Bluetooth_QoS=low Target_Network=Wi-Fi Bluetooth_Preferences=low SLS=high Wi-Fi_QoS=low
95 % UMTS_Preferences=low UMTS_QoS=low Wi-Fi_Preferences=low Bluetooth_QoS=low Target_Network=Wi-Fi Bluetooth_Preferences=low SLS=high Wi-Fi_QoS=low
95 % UMTS_Preferences=low UMTS_QoS=low Wi-Fi_Preferences=low Bluetooth_QoS=low Target_Network=Wi-Fi Bluetooth_Preferences=low SLS=high Wi-Fi_QoS=low)

After the BBN model has been built and the conditional probability and utility values entered, it can be used to make inference and decisions. First it has to be compiled. The compilation process is a follows as outlined in [5]:

1. The model is converted into its moral graph: The parents of each node are "married" (i.e., links are added between them), and the directions of the links are dropped.

2. The graph is triangulated.

3. The cliques (maximal complete sets) of the triangulated graph are identified, and the collection of cliques is organized as a tree (with the cliques forming the nodes of the tree). Such a tree is called a junction tree (also known as a join tree and a Markov tree [31]). If the original network is disconnected, there will be a tree for each connected component.

4. Finally, potentials are associated with the cliques and the links (the separators) of each junction tree. These potentials are initialized from the evidence and the conditional probability tables (and the policies and the utility tables in the case of LIMIDs), using a sum-propagation

Table 4.6: Compiling the model using Hugin API

```
ParseListener parseListener = new DefaultClassParseListener();
Domain domain = new Domain (domainName + ".net", parseListener);
domain.triangulate (Domain.H_TM_FILL_IN_WEIGHT);
domain.compile();
```

After the compilation process is complete the next step of the inference process is to make the inference engine aware of the evidence. This is called entering the evidence, and it can be done before or after the compilation step. Evidence which is synonymous to context in the handoff support system presented here is gathered from different location. Firstly, during the client's initialisation, the client declares the SLS level of the application. This evidence remains constant throughout the duration of the client's session. Secondly, during the client's active session, evidence is periodically gathered and propagated throughout the network as context changes.

Entering the evidence into the decision engine can be done in two ways:

1. Enter evidence for a given set of nodes (one node at a time).

2. Load the evidence for all nodes at once, when the evidence is stored in a case file. A case file provides an implementation of the context repository. Given the size of the handoff decision BBN and the dynamic nature of the wireless internet, using a case file to enter evidence is preferred. The context changes is simulated by contextChanger() method

from the HandoffTrigger class which periodically writes the state configurations from the utility table into a case file. These state configurations represent the context changes. This evidence is then propagated throughout the BBN to cause the other nodes to update their beliefs.

Table 4.7: Propagating the evidence

domain.propagate (Domain.H_EQUILIBRIUM_SUM, Domain.H_EVIDENCE_MODE_NORMAL);

Case file

As already mentioned, a case file provides an implementation of the context repository. A case file is a text file. The format (i.e., syntax) of a case file can be described by the following grammar [5]:

(Case file) $->$ (Node finding)*
(Node finding) $->$ (Node name):(Value)
(Value) $->$ (State index)|(Likelihood)|(Label)|(Real number)|true|false
where:

- (State index) is a valid specification for any discrete node,
- (Likelihood) is also a valid specification for all discrete nodes,
- (Real number) is a valid specification for Conditional Gaussian (CG), numbered, and interval nodes,
- (Label) is a valid specification for the labelled nodes,
- true and false are valid specifications for boolean nodes.

An example of a case file for a single case is shown in Table 4.8

For the handoff decision, once the evidence is entered and propagated, expected utilities associated with the states of a node (usually a decision node) can be retrieved. The handoff decision network contains the target network node. This node provides three alternative networks (Wi-Fi, Bluetooth and 3G). The utility values for each candidate network are computed. The principle of the maximum expected utility is used. The networks are then ranked. The network that exhibits the highest expected utility is the preferred target network. The function below can be used to retrieve expected utilities. When no observations are made, the nodes are characterised by their posterior probabilities.

Table 4.8: Example of case file entries

UMTS_Cost: ”expensive”
UMTS_IPC: ”low”
UMTS_DataRate: ”1Mbps”
UMTS_Bandwidth: ”1Mbps”
Wi-Fi_IPC: ”high”
Wi-Fi_Cost: ”cheap”
Bluetooth_DataRate: ”1Mbps”
Bluetooth_Bandwidth: ”1Mbps”
Bluetooth_IPC: ”low”
Bluetooth_Cost: ”cheap”
Data_Loss: ”1Mbps”
Jitter: ”10ms”
Tolerable_Delay: ”10ms”
Wi-Fi_DataRate: ”10Mbps”
Wi-Fi_Bandwidth: ”10Mbps”

Table 4.9: Retrieving utilities from decision nodes

network_Utility = DecisionNode.getExpectedUtility(stateIndex);

MultiThreadProxy

The MultiThreadProxy class is responsible for transmitting the multimedia data from the server to the client. MultiThreadProxy implements the soft handoff for duplicating media content to the client. Lastly, the MultiThreadProxy implements the QoS mapping procedure. This is in charge of compression, e.g., reduction of frame size/rate, and format transcoding, e.g., from AVI to MPEG, on VoD contents to tune the provided QoS level to suit the new network bandwidth. MultiThreadProxy exploits the JMF API.

RTPSender

In terms of transmitting the multimedia from the server to the client, the MultiThreadProxy delegates this task to the RTPSender class.

UDPReceiver

The UDPReceiver class serves as the receiver for the MainProxy, receiving requests from the client on all wireless network interfaces. Once a request arrives, the client’s ID is extracted and matched against the proxy-client pairs. If the specific proxy exists for the client, the ProxyThread calls the parseDatagram() to process the request.

CircularBuffer

The CircularBuffer class's main functionality is storage and retransmission of packets. In general it can be seen as a normal circular buffer in which the elements are stored until they are overwritten by subsequent writes. The internal buffer is an array of pointers to objects "Frame". The buffer is initially of a certain size, which corresponds to the array of pointers. Initially the read and write pointer point to the same block. For each frame the write pointer increments and the read pointer increments as well whenever the multiplexer reads a frame.

4.5 Summary

The experimental handoff support system was implemented as a proxy-based middleware architecture, with components running on fixed hosts in the wired network, *proxies*, as well as on the mobile device, client *stubs*. The Proxy is deployed on client-to-server distribution path and coordinate with their client stubs for handoff decision and execution. The system utilises emulator, emulating the different wireless networks/cards (Wi-Fi, UMTS, Bluetooth) under consideration. environment. The handoff engine based on BBN MCDM, which is the main contribution of the thesis is implemented as a proxy functionality to select the target network. The BBN handoff algorithm uses the Hugin API. Handoff execution is implemented as two strategies: Hard handoff and Soft handoff. Hard handoff utilises a buffering technique while soft handoff uses a double casting technique. All media processing and transmission is done by exploiting the JMF API. The QoS mapping is performed by the proxy by employing QoS adapters in JMF. The context repository is implemented as a Hugin case file for storing and processing context changes. The next chapter discusses the evaluation of the experimental handoff support system in an emulated environment.

5 EVALUATION

This thesis re-visits the problem of handoff support in wireless access networks. The research focuses on handoff decision and support which go beyond traditional techniques that only considered signal strength. These new handoff mechanisms have to now incorporate context information into decision making and execution which exemplifies deployment scenarios in future generation networks. Furthermore, the advent of pervasive computing especially when dealing with continuous services is calling for novel solutions tailor designed to provide seamless service continuity by minimizing flow interruptions as clients roam around different wireless networks [9]. The experimental handoff support system developed in this thesis is evaluated on how well it appropriately decides and executes handoff decisions based on dynamic environment and context in an emulated environment. The main result is; given varying context information, whether the system can effectively decide on the target AP for handoff (session transfer), and then apply the correct handoff management strategy to execute the handoff procedure. This procedure should be executed in a timely manner that minimises flow interruptions. Lastly, perform QoS mapping to fit the new deployment environment.

A list of the measurables for the system are summarised below:

- Handoff decision using context information.
- Decide handoff management strategy.
- Execute strategy by redirecting flow to new interfaces.
- Buffer incoming flow.
- Perform QoS mapping techniques where applicable.

In order to evaluate the performance of the experimental handoff support system in terms of the measurables above, a series of experiments were conducted. These experiments are divided into two categories: 1) Performance of the handoff decision algorithm. 2) Performance of the handoff management strategies.

The handoff decision is the main contribution of this research; therefore, more focus lies on this section. The experiments conducted to evaluate the performance of the handoff decision algorithm are:

1. Handoff decision correctness. This experiment aimed to evaluate the correctness of the handoff decision under different execution environments and context changes,

2. Handoff decision correctness under constraint criteria. This experiment aimed to evaluate the correctness of the handoff decision under different execution environments and context changes after introducing the SLS as a constraint criterion,

3. Handoff decision correctness for reduced BBN models. This experiment aimed to determine the correctness of the reduced BBN models for the two specialised profiles (see section 3.4) and comparing them with the full BBN models,

4. Time required for handoff decision completion. This experiment aimed to determine the average execution time of the handoff algorithm. Handoff decision execution time is important as it relates to the handoff latency, and

5. Time required for handoff decision completion for reduced models. This experiment aimed to determine the impact of reducing the BBN model on the execution time. The average execution time for the reduced BBN models for the two specialised profiles was compared to the full BBN models.

Handoff latency is an important factor in the handoff execution process. Minimizing handoff latency is particularly crucial when dealing with time sensitive applications. The experiments conducted to evaluate the handoff execution process under different handoff management strategies are:

1. Time required for handoff completion for soft handoff strategy. This experiment aimed to determine the total handoff procedure completion time when employing soft handoff strategy,

2. Time required for handoff completion for hard handoff strategy. This experiment aimed to determine the total handoff procedure completion time when employing hard handoff strategy,

3. Time required for vertical handoff completion between Bluetooth and Wi-Fi which represent today's widespread wireless technologies, and

4. Buffer utilisation. In the case of hard handoff management strategy, the proxy buffers incoming flow and forwards it to the client upon client re-connection. Measuring how much buffering is performed indicates how the system utilises the resources.

5.1 Resources

The following computer systems were used in the experiments:

- One machine referred to as the Proxy machine. This machine is equipped with 2.50 GHz Intel Core 2 Duo processor, 2 GB of RAM. The machine runs an ubuntu 9.10 operating system. This machine also acts as the RTP server.
- Two laptop machine referred to as a client machine. This machine is equipped with 1.66 GHz Intel Core 2 Duo processor, 1 GB of RAM. The machine runs an ubuntu 8.04 operating system.

5.2 Performance of the handoff decision algorithm

The handoff decision uses a MCDM method for target wireless network selection. The thesis proposes a MCDM method based on BBN. For the sake of comparison, AHP, another MCDM method which has been used extensively in literature was implemented and compared with BBN. Before the performance of these two methods is compared, it is worth discussing challenges and limitations encountered upon implementation due to their structures.

5.2.1 Comparison between BBN and AHP structures

For evaluation, the adopted hierarchies for AHP and BBN are shown in Figures 5.1 and 5.2 respectively. The comparison focused on how the two methods represent a decision problem and how information for each MCDM elements is entered into the models.

Issues with AHP

1. Inability to model constraints: AHP involves building pairwise comparison matrices representing the options/alternatives relative to each other based on different criteria. In some decision problems, the criteria might be a constraint and does not relate to the alternatives. This criteria cannot be represented in AHP as it does not relate to the alternatives. Furthermore, the constraint criteria cannot be ranked in importance to the other criteria relative to the overall goal. This limitation is shown in Figure 5.1 which reveals that though being an important criteria in the handoff decision problem, the SLS is not included in the AHP hierarchy. The SLS encodes the application requirements for the user and does not relate to the alternative networks. The SLS can be regarded as a constraint criterion in this handoff problem as the user can require different decisions based on the values represented by the SLS, and

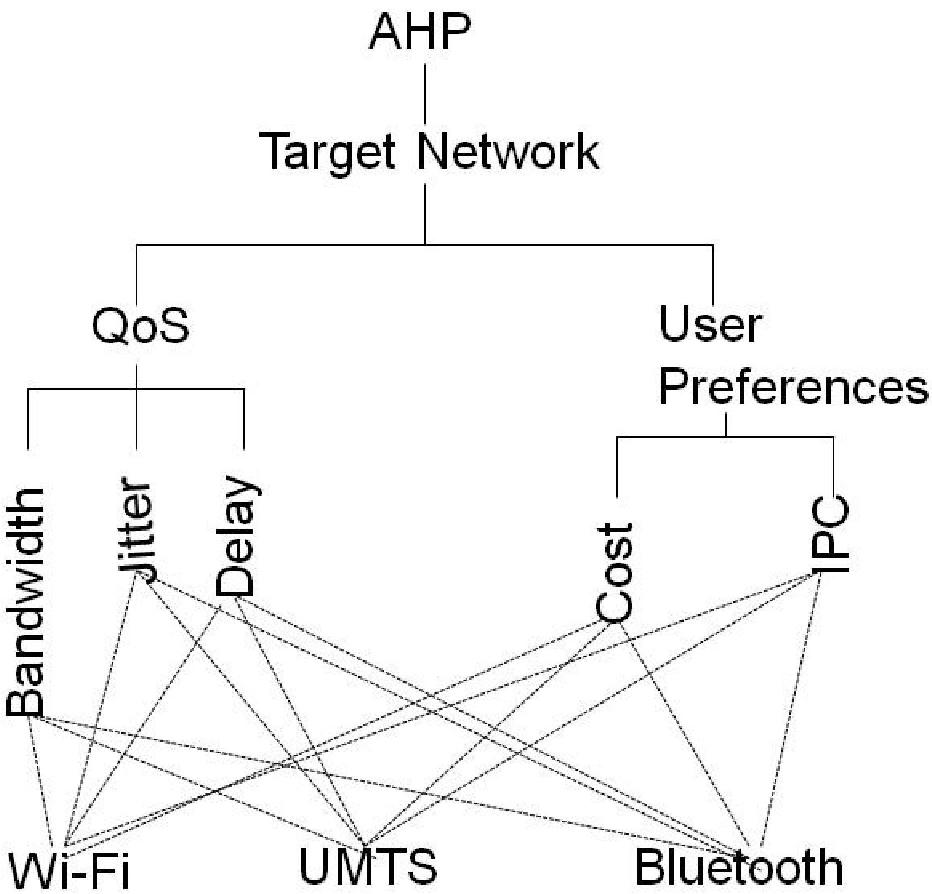

Figure 5.1: AHP hierarchy

On the contrary, the BBN framework can represent constraints criteria as shown in Figure 5.2, since it does not require pairwise comparison of alternatives against criteria. This limitation with AHP marks a significant shortcoming as it does not model all the relevant context information needed in the handoff decision and execution.

2. Inability to model diverse context changes: The structure of the pairwise comparison scale is limited in specifying a diversity of context differences. This further adds a complication to an ordinary user in terms of specifying the relative importance of alternatives against a criteria. For instance when, comparing the three networks (Wi-Fi, UMTS and Bluetooth) relative to QoS, a priority matrix can be specified such as in Table 5.1.

Table 5.1 implies that, in terms of QoS, Wi-Fi is moderately preferred to UMTS and Bluetooth. Furthermore, UMTS and Bluetooth are equally preferred. If the user wants to explore more context representation for comparison, AHP provides options shown by table 5.2 and 5.3. Table 5.2 conveys that Wi-Fi is strongly preferred to UMTS and Bluetooth, while UMTS and Bluetooth are equally preferred. Table 5.3 furthermore conveys that Wi-Fi is extremely preferred to UMTS

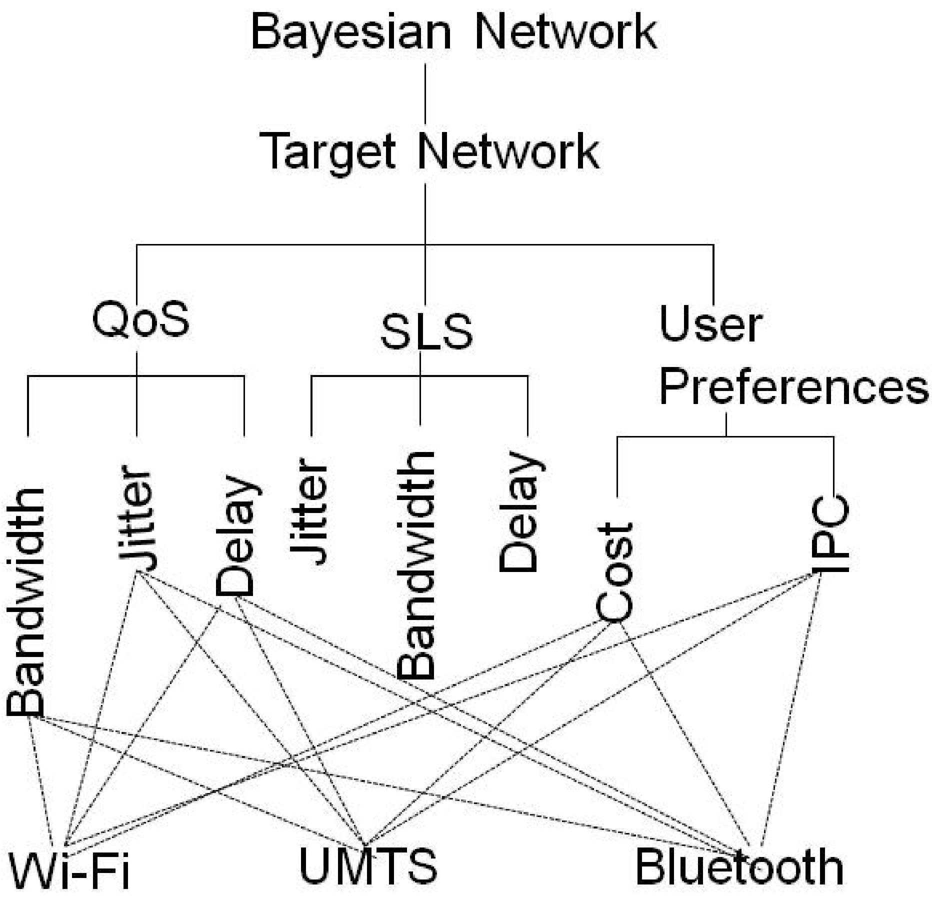

Figure 5.2: BBN hierarchy

Table 5.1: QoS Priority Matrix 1

QoS	Wi-Fi	Bluetooth	UMTS
Wi-Fi	1	3	3
Bluetooth	0.3333	1	1
UMTS	0.33333	1	1

and Bluetooth, while UMTS and Bluetooth are equally preferred. These three tables all mean that Wi-Fi is a preferred choice in terms of QoS regardless of the level of importance (moderately, strongly, extremely preferred). This can be classified as a limitation of AHP in terms of representing diverse context information as these table represent but one context scenario.

This can be furthermore demonstrated by trying to model a scenario whereby all the networks have equally poor QoS, which would require an additional adaptation i.e. downscale the content in the case of multimedia or equally good QoS, which does not require an additional adaptation. These two scenarios can only be modelled by the same table, table 5.4 whilst they represent dif-

Table 5.2: QoS Priority Matrix 2

QoS	Wi-Fi	Bluetooth	UMTS
Wi-Fi	1	5	5
Bluetooth	0.2	1	1
UMTS	0.2	1	1

Table 5.3: QoS Priority Matrix 3

QoS	Wi-Fi	Bluetooth	UMTS
Wi-Fi	1	9	9
Bluetooth	0.142	1	1
UMTS	0.142	1	1

ferent environments. Hence the model cannot inform whether additional adaptation is required or not. BBN on the other hand provides evidence entering for each alternative independently. Hence the decision maker knows the state of each alternative relative to a given criteria. As shown in utility tables 3.8, 3.9 and 3.10, this method provides clear context representation and easy evidence entering.

Table 5.4: QoS Priority Matrix 4

QoS	Wi-Fi	Bluetooth	UMTS
Wi-Fi	1	1	1
Bluetooth	1	1	1
UMTS	1	1	1

Issues with BBN

1) Large table sizes: As already discussed in sub-section 3.4, when dealing with large and complex problems, constructing the conditional and utility tables for the BBN model becomes difficult. This issue is particularly complex when the conditional probabilities and utilities are obtained from subjective estimate. The utility table size grows exponentially with the increase in number of criteria as shown by Figure 5.3. This figure considers only when the criteria under consideration have two states. AHP on the contrary provides a simpler method of pairwise comparison. When the number of criteria increases, only the number of pairwise comparison matrices increase but the number of entries in the matrices remain constant. The following experiments were conducted to determine the correctness of the BBN decision algorithm in terms of network ranking.

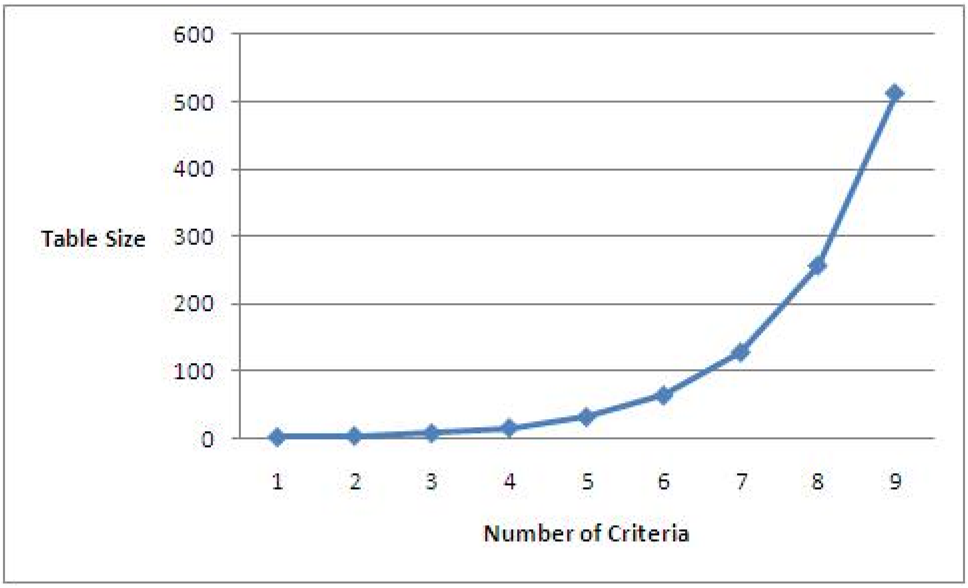

Figure 5.3: Utility table growth vs. Increase in criteria

5.2.2 Handoff decision correctness in terms of network ranking

The correctness of the algorithm is crucial in the handoff decision process. Traditional solutions only considered the RSS which can take a long time to discover and does not represent the execution environment of the NGN. These NGN consist of heterogeneous wireless networks with different QoS, cost and coverage. Furthermore, the user takes part in the decision process by specifying their preferences. Therefore, decision process takes into account different criteria, as a result, the correctness of the decision algorithm based on varying context is important.

Aim

The aim of this experiment was to determine the correctness (in terms of network ranking) of the BBN based decision algorithm in deciding the target wireless access network under varying context. In order to determine the network ranking correctness, the method was compared with AHP.

Methodology and Results

Two test profiles described in section 3.3 were created for BBN based model. Two profiles (A1 and B1) were further created for AHP. For AHP, a goal priority matrix was created for each profile. Table 5.5 shows the goal priority matrix for profile A1. This table denotes that relative to the overall goal (deciding best target network), the QoS criteria is "Extremely preferred" over user preferences (cost and interface power consumption). Furthermore, different weightings for the alternative networks were used in the QoS feature priority matrix. The second profile B1 on the other hand, considers ”user preferences” criteria to be ”Extremely preferred” over QoS as shown in table 5.6.

To re-iterate, for BBN, profile A implements a utility function that rewards state configuration that has networks with high QoS. The sample utility table is shown in table 3.8. Profile B on the contrary implements a utility function that rewards state configuration that has networks with low user preferences, Table 3.9. Therefore, in essence, profiles A and A1 for BBN and AHP respectively focus solely on QoS criteria. Profiles B and B1 for BBN and AHP on the contrary focus mainly on ”user preferences” criteria.

Table 5.5: Goal Priority Matrix for Profile A1

	QoS	User Preferences
QoS	1	9
User Preferences	0.111	1

Table 5.6: Goal Priority Matrix for Profile B1

	QoS	User Preferences
QoS	1	0.111
User Preferences	9	1

The two methods (AHP and BBN) differ significantly as outlined in section 5.2. Therefore, out of 30 runs for the BBN and 30 runs for AHP, only 4 runs were directly comparable in each profile category.

When there is no context observed (see Table 5.7), the expected utilities for BBN are: 0.5127, 0.4843 and 0.4728 for Wi-Fi, UMTS and Bluetooth respectively. AHP does not model uncertainty or ”no-context observation”, hence does not compare with BBN in that respect. With AHP (see Table 5.7), when all networks are equally preferred based on QoS, the expected utilities are: 0.333, 0.333 and 0.333 for Wi-Fi, UMTS and Bluetooth respectively. With BBN, when all the networks have ”high” QoS (see table 5.7), the expected utilities are: 0.5208, 0.5 and 0.4999 for Wi-Fi, UMTS and Bluetooth respectively.

For AHP, when Wi-Fi is preferred over UMTS and Bluetooth, the expected utilities are: 0.5670, 0.2164 and 0.2164 for Wi-Fi, UMTS and Bluetooth respectively. This context is similar to BBN where Wi-Fi has ”high” QoS and both UMTS and Bluetooth have ”low” QoS. The expected utilities for BBN are: 0.5697, 0.0 and 0.0 for Wi-Fi, UMTS and Bluetooth respectively. Both methods produce the same network ranking, with Wi-Fi as the preferred network.

For AHP, when Bluetooth is preferred over UMTS and Wi-Fi, the expected utilities are: 0.2164, 0.2164 and 0.5670 for Wi-Fi, UMTS and Bluetooth respectively. This context is similar to BBN where Bluetooth has ”high” QoS and both UMTS and Wi-Fi have ”low” QoS. The expected utilities for BBN are: 0.0, 0.0 and 0.5697 for Wi-Fi, UMTS and Bluetooth respectively. Both methods produce the same network ranking, with Bluetooth as the preferred network.

For AHP, when UMTS is preferred over Bluetooth and Wi-Fi, the expected utilities are: 0.2164, 0.5670 and 0.2164 for Wi-Fi, UMTS and Bluetooth respectively. This context is similar to BBN where UMTS has ”high” QoS and both Bluetooth and Wi-Fi have ”low” QoS. The expected utilities for BBN are: 0.0, 0.5697 and 0.0 for Wi-Fi, UMTS and Bluetooth respectively. Both methods produce the same network ranking, with UMTS as the preferred network.

For profile B, when no observation is made for BBN (see Table 5.8), the expected utilities: 0.5015, 0.4998 and 0.5146 for Wi-Fi, UMTS and Bluetooth respectively.

With AH, when all networks are equally preferred based on ”user preferences”, the expected utilities are: 0.333, 0.333 and 0.333 for Wi-Fi, UMTS and Bluetooth respectively. With BBN, when all the networks have ”low” ”user preferences” (meaning low cost and low IPC), the expected utilities are: 0.5001, 0.4999 and 0.5116 for Wi-Fi, UMTS and Bluetooth respectively. Tables 5.7 and 5.8 and Figures 5.4 and 5.5 show the rest of the results.

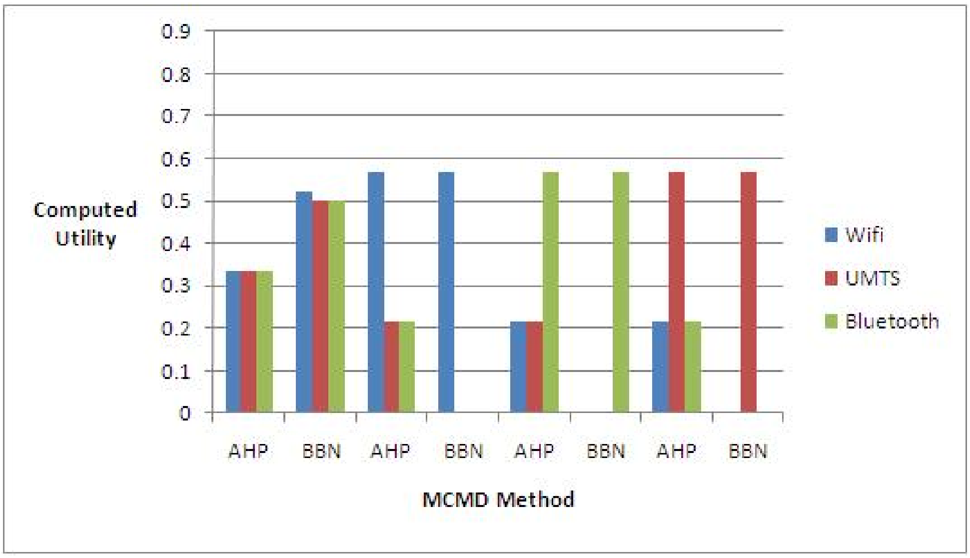

Figure 5.4: Expected utilities for BBN(Profile A) and AHP(Profile A1)

Discussion

When no observation is made for BBN (Table 5.7), the nodes are characterised by their posterior probabilities. Wi-Fi has the highest utility. This means, by default, the model yields Wi-Fi as a preferred target network. Deducing from the conditional probability and utility tables, Wi-Fi QoS is high as a result, a high QoS network, is preferred in this case. AHP on the contrary requires all evidence to be entered before making a decision. This is a significant shortcoming as not all context information in wireless networks is available at the time of the decision making.

Table 5.7: Utilities from BBN(Profile A) and AHP(Profile A1)

Method	Context	Wi-Fi	UMTS	Bluetooth
BBN	No Context	0.5127	0.4843	0.4728
AHP	(Wi-Fi, Bluetooth):1 (Wi-Fi, UMTS):1 (UMTS, Bluetooth):1	0.3333	0.3333	0.3333
BBN	Wi-Fi_QoS = "high" UMTS_QoS = "high" Bluetooth_QoS = "high"	0.5208	0.5	0.4999
AHP	(Wi-Fi, Bluetooth):3 (Wi-Fi, UMTS):3 (UMTS, Bluetooth):1	0.5670	0.2164	0.2164
BBN	Wi-Fi_QoS = "high" UMTS_QoS = "low" Bluetooth_QoS = "low"	0.5697	0.0	0.0
AHP	(Bluetooth, Wi-Fi):3 (Bluetooth, UMTS):3 (UMTS, Wi-Fi):1	0.2164	0.2164	0.5670
BBN	Wi-Fi_QoS = "low" UMTS_QoS = "low" Bluetooth_QoS = "high"	0.0	0.0	0.5697
AHP	(UMTS, Wi-Fi):3 (UMTS, Bluetooth):3 (Bluetooth, Wi-Fi):1	0.2164	0.5670	0.2164
BBN	Wi-Fi_QoS = "low" UMTS_QoS = "high" Bluetooth_QoS = "low"	0.0	0.5697	0.0

Hence, the ability to make decisions under uncertainty is important.

With AHP (see Table 5.7), when all networks are equally preferred based on QoS, all networks have equal expected utilities. This implies that all networks have a equal chance of being chosen. The decision maker still has to decide on which network to pick. This can be by default or random. This context is similar to BBN, when all the networks have "high" QoS. Wi-Fi has the highest expected utility. This feature gives BBN more advantage over AHP with regards to specifying preferred target network. The user can still specify their preferred network when all networks are equal based on QoS.

For both profiles classes (see Tables 5.7 and 5.8), AHP yields the same expected utilities. This is because AHP is deterministic is terms of computing utilities. The same values from the comparison scale are used in the pairwise comparison matrices. On the contrary, BBN yields different expected utilities for different profiles. This is due to the fact that the decision maker can assign different weightings on different state configurations based on their own interests.

Table 5.8: Utilities from BBN(Profile B) and AHP (Profile B1)

Method	Context	Wi-Fi	UMTS	Bluetooth
BBN	No Context	0.5015	0.4998	0.5146
AHP	(Wi-Fi, Bluetooth):1 (Wi-Fi, UMTS):1 (UMTS, Bluetooth):1	0.3333	0.3333	0.3333
BBN	Wi-Fi_userPreferences = ”low” UMTS_userPreferences = ”low” Bluetooth_userPreferences = ”low”	0.5001	0.4999	0.5116
AHP	(Wi-Fi, Bluetooth):3 (Wi-Fi, UMTS):3 (UMTS, Bluetooth):1	0.5670	0.2164	0.2164
BBN	Wi-Fi_userPreferences = ”high” UMTS_userPreferences = ”low” Bluetooth_userPreferences = ”low”	0.5697	0.0	0.0
AHP	(Bluetooth, Wi-Fi):3 (Bluetooth, UMTS):3 (UMTS, Wi-Fi):1	0.2164	0.2164	0.5670
BBN	Wi-Fi_userPreferences = ”high” UMTS_userPreferences = ”high” Bluetooth_userPreferences = ”low”	0.0	0.0	0.5480
AHP	(UMTS, Wi-Fi):3 (UMTS, Bluetooth):3 (Bluetooth, Wi-Fi):1	0.2164	0.5670	0.2164
BBN	Wi-Fi_userPreferences = ”high” UMTS_userPreferences = ”low” Bluetooth_userPreferences = ”high”	0.0	0.7692	0.0

The remainder of the results is consistent for both methods according to the given context, by producing the same network ranking. Therefore, both methods proved correctness under changing context, however, BBN provides more expressiveness to the context information, while AHP is rigid.

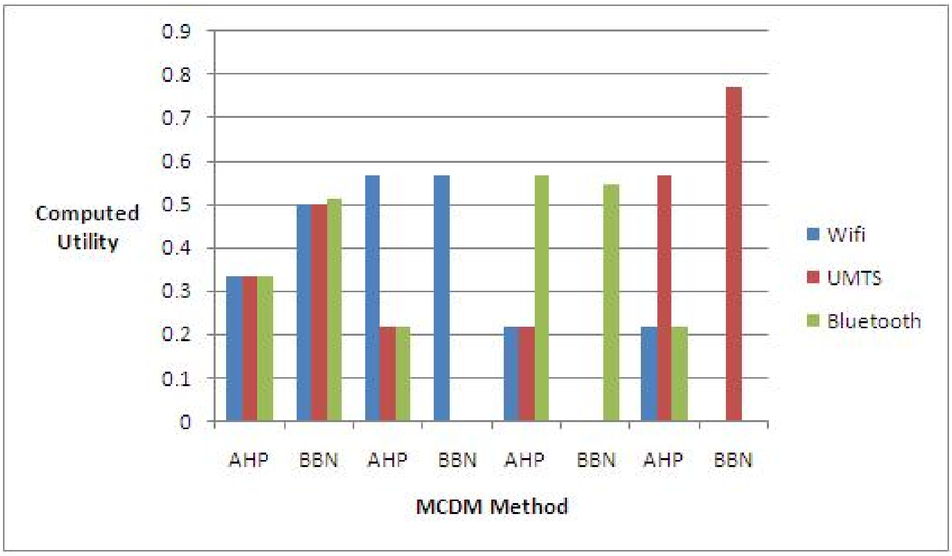

Figure 5.5: Expected utilities for BBN(Profile B) and AHP(Profile B1)

5.2.3 Handoff decision correctness in terms of network ranking under constraint criteria

Aim

The aim of this experiment was to determine the correctness of the decision algorithm based on constraint criteria.

Methodology and Results

Only BBN based decision algorithm was evaluated in this experiment as AHP does not provide the modelling of constraint criteria (refer to sub-section 5.2.1). A third profile C was created from the BBN based model. This profile combines the concepts from profile A1 and B1 whereby a constraint criteria, SLS, is introduced. Profile C implements a utility function that rewards state configuration that has networks with high QoS if the SLS is declared to be ”high” else rewards state configuration that has networks with low user preferences if the SLS is declared to be ”low”. If the SLS is high and all the networks can provide the required QoS, the expected utilities are: 0.5466, 0.5 and 0.4999 for Wi-Fi, UMTS and Bluetooth respectively. When all networks have low QoS, the expected utilities are: 0.2101, 0.0 and 0.0 for Wi-Fi, UMTS and Bluetooth respectively. If only Bluetooth has a high QoS, the expected utilities are: 0.0, 0.0 and 0.5262 for Wi-Fi, UMTS and Bluetooth respectively. If the SLS is low and all networks have low ”user preferences”, the expected utilities are: 0.5099, 0.4647 and 0.5177 for Wi-Fi, UMTS and Bluetooth respectively. The rest of the results are shown in Table 5.9.

Table 5.9: Utilities for BBN (profile C)

Method	Context	Wi-Fi	UMTS	Bluetooth
BBN	No Context	0.5068	0.4593	0.4819
BBN	Wi-Fi_QoS = "high" UMTS_QoS = "high" Bluetooth_QoS = "high" SLS = "high"	0.5466	0.5	0.4999
BBN	Wi-Fi_QoS = "low" UMTS_QoS = "low" Bluetooth_QoS = "low" SLS = "high"	0.2101	0.0	0.0
BBN	Wi-Fi_QoS = "low" UMTS_QoS = "low" Bluetooth_QoS = "high" SLS = "high"	0.0	0.0	0.5262
BBN	Wi-Fi_QoS = "low" UMTS_QoS = "high" Bluetooth_QoS = "low" SLS = "high"	0.0	0.5394	0.0
BBN	Wi-Fi_userPreferences = "low" UMTS_userPreferences = "low" Bluetooth_userPreferences = "low" SLS = "low"	0.5099	0.4647	0.5177
BBN	Wi-Fi_userPreferences = "high" UMTS_userPreferences = "low" Bluetooth_userPreferences = "low" SLS = "low"	0.0	0.4720	0.4831
BBN	Wi-Fi_userPreferences = "high" UMTS_userPreferences = "high" Bluetooth_userPreferences = "low" SLS = "low"	0.0	0.0	0.6745
BBN	Wi-Fi_userPreferences = "high" UMTS_userPreferences = "low" Bluetooth_userPreferences = "high" SLS = "low"	0.0	0.5412	0.0

Discussion

The SLS constraint criteria added controls which criteria is important in the decision process. The SLS can either be high or low depending on the application requirement for the user. This implies that profile C is focused at users who want a network that provides high QoS if they are running a critical application else prefer a network with lower user preferences when their application in not critical. From the results, the network ranking favours a high QoS network when SLS is high. If two or more networks have high QoS, the decision maker can specify their preferred network. For instance, when Wi-Fi, Bluetooth and UMTS all have high QoS, Wi-Fi has the highest expected utility of 0.5466 because the decision maker prefers Wi-Fi in terms of QoS. Furthermore, the network ranking favours a low user preference network when SLS is low. The SLS is being used as a constraint whereby the network ranking is controlled by the current value of the SLS. This ability to model constraints provides the decision maker i.e. a service provider in creating different user profiles for their customers specific to their needs.

5.2.4 Handoff decision correctness in terms of network ranking for reduced BBN models

BBN provides more expressiveness in representing the context information as outlined in subsection 5.2.1. However, when the number of criteria increases, the corresponding utility table grows exponentially, rendering it difficult to manage. Section 3.4 described how this challenge was partly addressed by reducing the BBN model for specific profiles.

Aim

The aim of this experiment was to determine if reducing the number of criteria for models that focused on a single criterion would still produce the correct network ranking.

Methodology and Results

Profile A and B described in section 3.3 were reduced to one criteria each. Profile A and B were reduced to network QoS and ”user preferences” respectively as the only criteria. After implementation, the resulting models resulted into 3 criteria each with utility table of 24 entries. The models for the reduced profiles (A' and B') were ran under the matching state configurations (same context changes) of utility tables with their full models' counterparts and the computed expected utilities compared.

When there is no evidence (see Figure 5.6), expected utilities for profile A are: 0.5127, 0.4843 and 0.4728 for Wi-Fi, UMTS and Bluetooth respectively, for profile A': 0.804, 0.7 and 0.5 for Wi-Fi, UMTS and Bluetooth respectively. Both models for profile A and A' yield Wi-Fi to have the highest expected utility.

When UMTS has ”high” QoS and both Bluetooth and Wi-Fi have ”low” QoS., expected utilities for profile A are: 0.5, 0.5421 and 0.4999 for Wi-Fi, UMTS and Bluetooth respectively, for profile A': 0.6, 0.75 and 0.55 for Wi-Fi, UMTS and Bluetooth respectively. Both models for profile A and A' yield UMTS to have the highest expected utility.

When there is no evidence (see Figure 5.7), expected utilities for profile B: 0.4998, 0.5115 and 0.5146 for Wi-Fi, UMTS and Bluetooth respectively, and for profile B': 0.6666, 0.6118 and 0.75354 for Wi-Fi, UMTS and Bluetooth respectively. Both models for profile B and B' yield Bluetooth to have the highest expected utility.

When Bluetooth has "low" "user preferences" and both Wi-Fi and UMTS have "high" "user preferences", expected utilities for profile B are: 0.4999, 0.5 and 0.5178 for Wi-Fi, UMTS and Bluetooth respectively, for profile B': 0.0, 0.0 and 0.6118 for Wi-Fi, UMTS and Bluetooth respectively. Both models for profile B and B' yield Bluetooth to have the highest expected utility.

Figures 5.6 and 5.7 show the results.

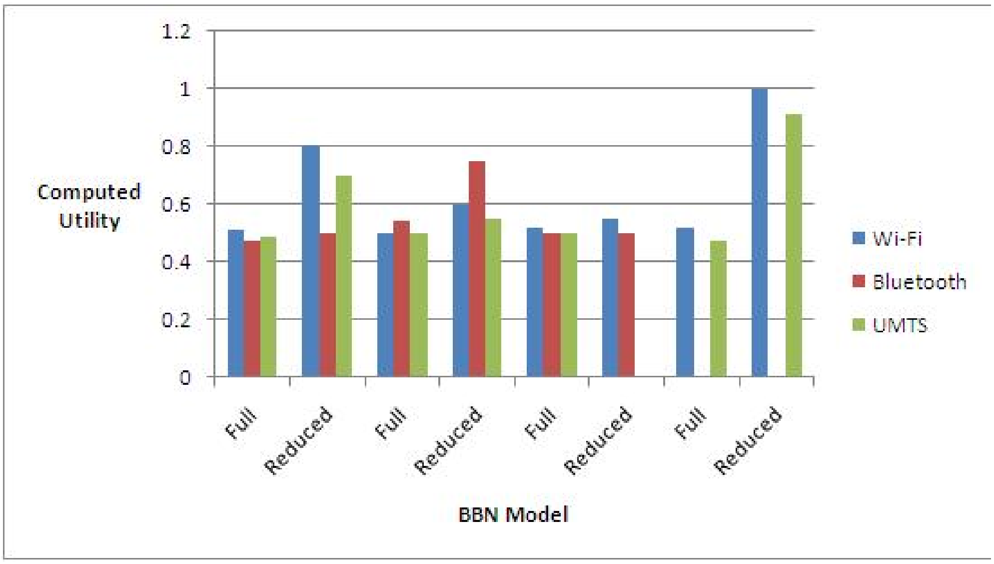

Figure 5.6: Expected utilities for BBN(Profile A) and BBN(Profile A')

Discussion

The analysis of the results show that both models (Old BBN and Reduced BBN) on both criteria (QoS and User Preferences) yield the same network ranking, under the same state configurations of the utility tables. The computed utility values are not the same, but the ranking of the networks is that same. This implies removing the criteria which is not considered in a utility table does not affect the correctness of the computation. These results probe a further investigation on whether removing the unused criteria has an impact on the overall execution time of the BBN model.

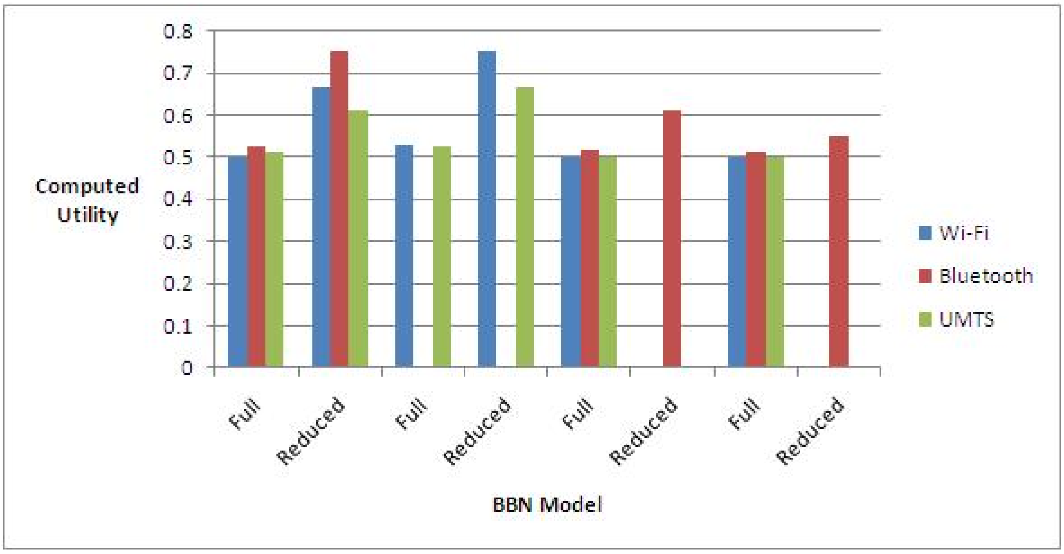

Figure 5.7: Expected utilities for BBN(Profile B) and BBN(Profile B')

5.2.5 Time required for handoff decision completion

The handoff decision process is responsible for deciding on the target wireless network when a client is amid different wireless networks. Introducing the handoff decision incurs a time overhead into the whole handoff support process. Since the aim of handoff support systems is to minimise handoff latency, then the handoff decision execution time should be minimal.

Aim

The aim of this experiment was to determine the time it takes for the handoff decision algorithm to execute for both the BBN and AHP models. This time includes reading the context information and computing the corresponding utilities for all alternative networks.

Methodology and Results

To determine the handoff decision execution time, the two MCDM methods (BBN and AHP) were employed and compared. Each method was executed under different context changes. For the AHP, 30 different pairwise comparisons of the three alternative networks were fed into the method. For the BBN, 30 different state configurations of the handoff decision model were fed into the method. The average time for the AHP method to perform the decision was found to be 5ms. On the other hand, the BBN method took 13ms on average to perform the decision. The results of the execution time are shown in the Figure 5.8.

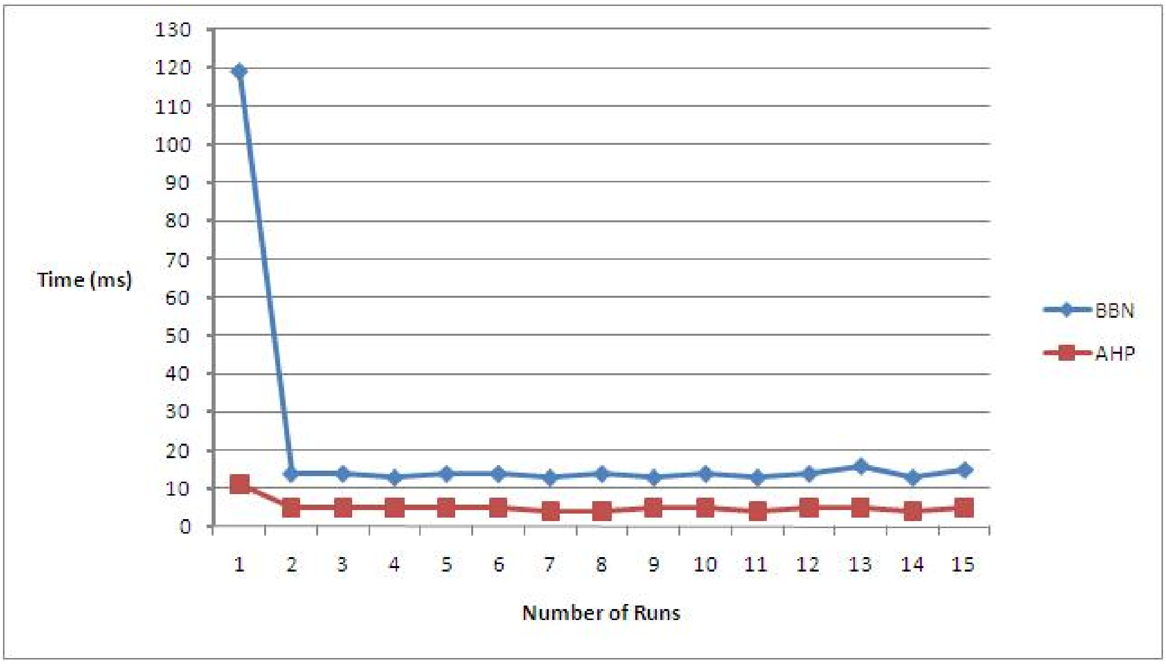

Figure 5.8: Execution time for both AHP and BBN

Discussion

According to Figure 5.8, both methods take a significantly longer time in their first execution than the subsequent executions. This is due to the initial memory operations for loading the method data structures. Since the BBN forms a tree-like data structure, it takes longer to load and process. AHP uses pairwise comparison matrices. These can be represented as lists in memory, hence faster to process. Furthermore, the difference of 8ms in execution time between the two methods is not very significant. The reason is, the handoff latency of most network technologies in between 100-500ms depending on the speed. Therefore both methods do not add excess overhead on the handoff latency. AHP however is much faster than BBN in terms of execution.

5.2.6 Time required for handoff decision completion for reduced BBN models

Aim

The aim of this experiment was to determine if the reducing the number of criteria on the BBN models has an impact on the time it takes for the handoff decision algorithm to execute. This time includes reading the context information and computing the corresponding utilities for all alternative networks.

Methodology and Results

To determine the handoff decision execution time for this experiment, the reduced BBN models were used and compared with their full model counterparts. 30 different state configurations of the handoff decision model were fed into the method. The average time it took to perform the handoff decision dropped from 15ms for the full model to 6ms for the reduced models. The results of the execution times are shown in the Figure 5.9.

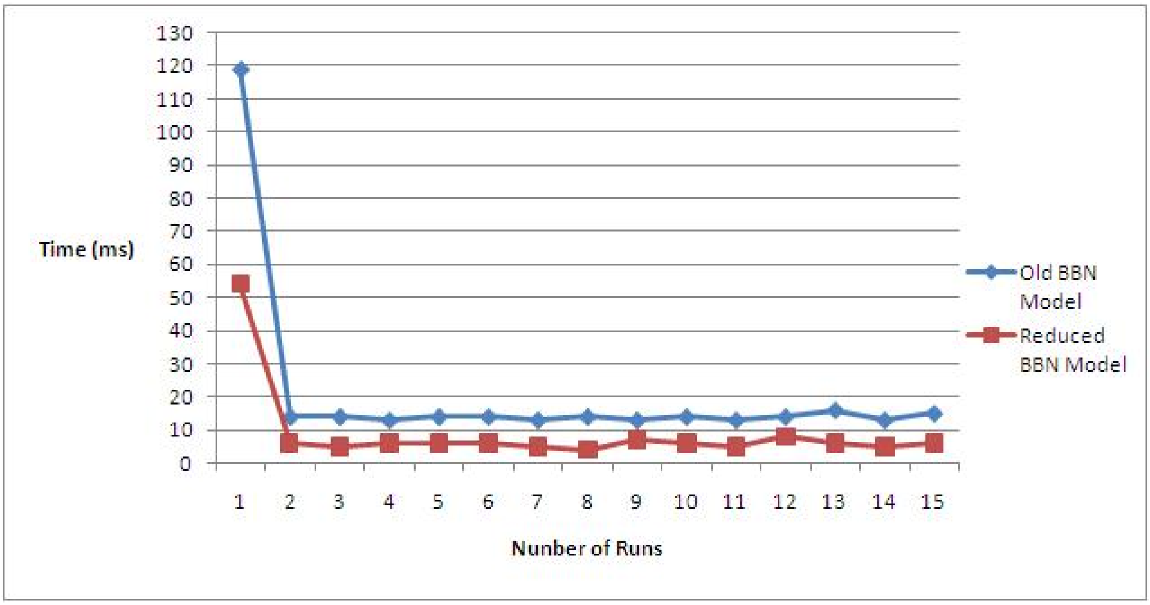

Figure 5.9: Execution times for full BBN model vs. reduced BBN model

Discussion

From the results, the average execution time for the BBN models decreases with decreasing criteria. Since BBN forms a tree-like data structure, reducing the criteria reduces the time taken to traverse through the tree. This implies that, for profiles that focus on a certain criteria, i.e. QoS, the reduced models can be used as they execute faster that the full models.

5.3 Performance of the handoff management strategies

The experimental handoff support system presented in this thesis implements two handoff management strategies; hard and soft handoff. The performance of the overall system relates to: 1) The time it takes to execute a handoff process i.e. target network selection and handoff strategy execution. 2) The utilisation of system resources upon handoff execution.

5.3.1 Time required for handoff completion for hard handoff strategy

Aim

The aim of this experiment was to determine the time it takes for the overall handoff process to execute when employing hard handoff strategy. This process includes the 1) target network decision time for both AHP and BBN and 2) the hard handoff strategy execution time.

Methodology and Results

To determine the overall handoff execution time for hard handoff, the two MCDM methods (BBN and AHP) running on the proxy machine were employed. The decision result from the MCDM is used to decide the handoff strategy. The proxy then coordinates with the client stub to execute the appropriate handoff strategy. If it's a horizontal handoff, a hard handoff strategy is employed, else if it is vertical handoff, soft handoff strategy is employed. In this experiment, only horizontal handoff occurrences were considered. Therefore, the overall time is divided into: Target network selection, hard handoff strategy preparation and hard handoff strategy execution. Both MCDM methods and subsequent hard handoff strategies were timed under different context changes of 30 runs. According to the results, the average time for hard handoff procedure was 3050ms The results of the execution time are shown in Figures 5.10.

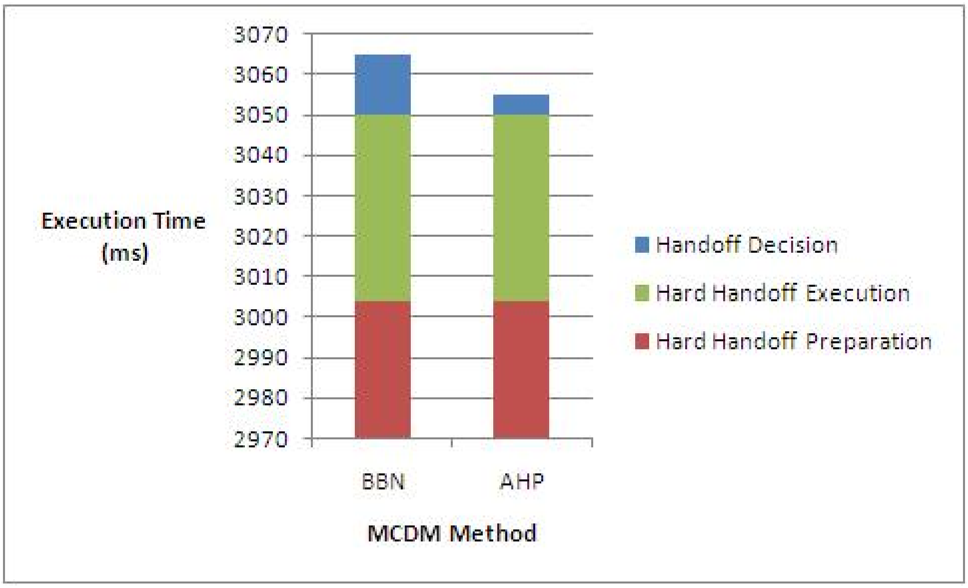

Figure 5.10: Overall Hard Handoff Execution Time

Discussion

The average time for a hard handoff procedure is 3050ms. Hard handoff management strategy uses one wireless interface on the client during the handoff process. In order to address packet losses upon client disconnection, it buffers incoming flow. Then forwards the buffered data to the client upon re-connection. This phase constitutes the longest period of 3004ms on average. The last phase of interface activation and flow completion averages 46ms. The sequence of steps in the hard handoff take a significantly long time which exceed the handoff latency of many wireless networks, hence can undermine service continuity. Hard handoff is suitable for applications that are more sensitive to data loss and it provides a mechanism for data buffering.

5.3.2 Time required for handoff completion for soft handoff strategy

Aim

The aim of this experiment was to determine the time it takes for the overall handoff process to execute when employing soft handoff strategy. This process includes the 1) target network decision time for both AHP and BBN and 2) the soft handoff strategy execution time.

Methodology and Results

This experiment used the same methodology as the one described above but focusing on the vertical handoff occurrences. In this case the overall time is divided into: Target network selection, soft handoff strategy preparation and soft handoff strategy execution. According to the results, the average time for soft handoff was procedure was 42ms. The results of the execution time are shown in Figure 5.11.

Discussion

Soft handoff takes an average of 42ms execution time. Upon handoff triggering, the proxy coordinates with the client stub for activation of the second interface on the client and duplication of the stream on the proxy to the client through the two interfaces. This phase averages 5ms. The last phase of soft handoff takes 37ms to complete flow duplication and old interface de-activation when the new interface has been fully activated. Soft handoff execution time of 42ms is lower than the handoff latency of many wireless networks, hence soft handoff is more preferred over hard handoff for critical time sensitive applications. However, theoretically, soft handoff is more expensive in terms of interface power consumption as two interfaces are active at the same time during the duration of the handoff process. This may be undesirable when the battery power is significantly low. The average time of 42ms in this emulated environment

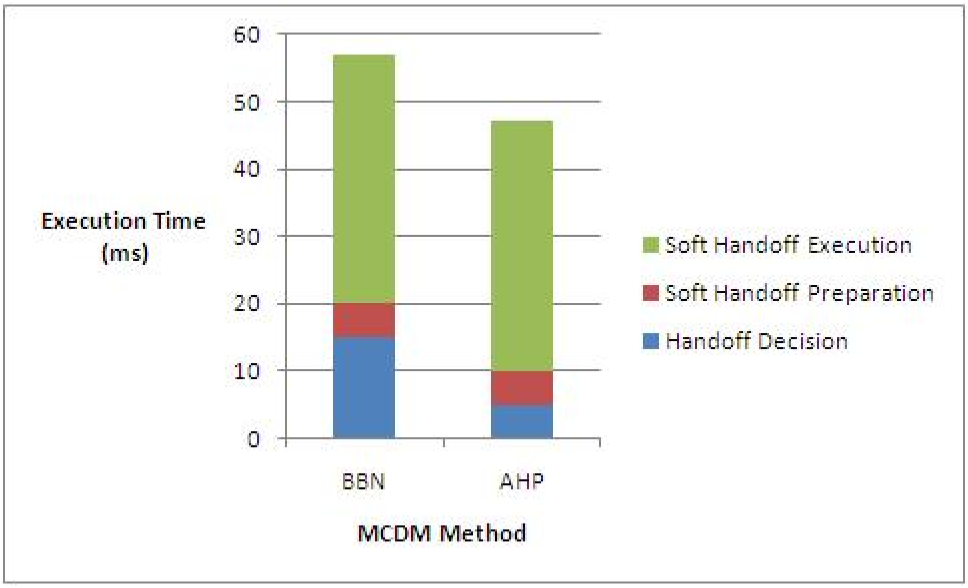

Figure 5.11: Overall Soft Handoff Execution Time

is significantly lower than expected in the real environment hence does not presented a real wireless behaviour. This experiment also proved that the handoff decision time using the two MCDM (BBN and AHP) is constant and does not have a huge impact on the overall handoff process under the two handoff strategies.

5.3.3 Time required for vertical handoff completion between Bluetooth and Wi-Fi

Aim

The aim of this experiment was to measure the average time required for vertical handoff completion in four different situations: from Bluetooth to Wi-Fi and vice versa, by employing both BBN and AHP decision methods.

Methodology and Results

This experiment employed a similar methodology to the previous one. The only difference is the focus is on soft handoff between Bluetooth and Wi-Fi which represent today's widespread wireless technologies. The overall time is divided into: Target network selection, handoff strategy preparation and handoff strategy execution. however, for handoff procedure from Wi-Fi to Bluetooth, there is an additional overhead of Bluetooth pre-inquiry. Figure 5.12 shows the results from different context changes of 30 runs.

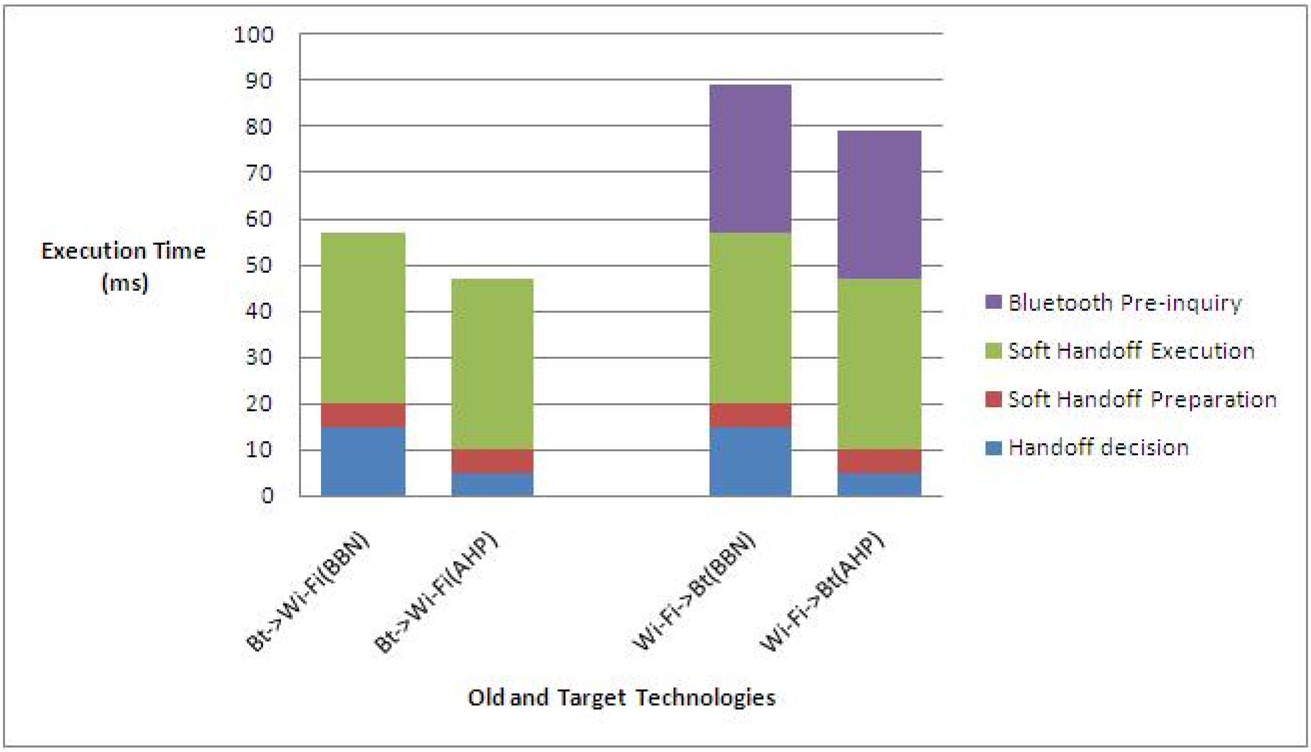

Figure 5.12: Overall Soft Handoff Execution Time between Bluetooth and Wi-Fi

Discussion

The first phase of soft handoff, target wireless network selection remains invariant with 5ms and 15ms for AHP and BBN respectively regardless of the direction of the handoff i.e. from Bluetooth to Wi-Fi and vice versa. The last phase of soft handoff execution takes an average of 37ms from Bluetooth to Wi-Fi and 59ms from Wi-Fi to Bluetooth. This implies that the activation of a new link on the Bluetooth incurs an additional 32ms attributed to Bluetooth pre-inquiry. Pre-inquiry is a procedure whereby the device scans for other Bluetooth devices in its vicinity. This additional overhead time slows down the handoff process and can undermine service continuity.

5.3.4 Buffer utilisation

Aim

The aim of this experiment was to determine the usage of the proxy buffer during a client's session for hard handoff occurrences.

Methodology and Results

For hard handoff, the proxy buffers the incoming data flow from the server destined to the client during client's disconnection. The experiment consists of a RTP server transmitting a H263 encoded video to the client. For this experiment, the RTP server and proxy run on the same machine. In reality, these are two separate machines in different network domains. The focus is on examining the proxy buffer usage during hard handoff occurrence. During a hard handoff usage, the buffer level rises to 150 as shown in Figure 5.13. This level is maintained for the duration of the handoff (3005ms) and then drops to a normal level.

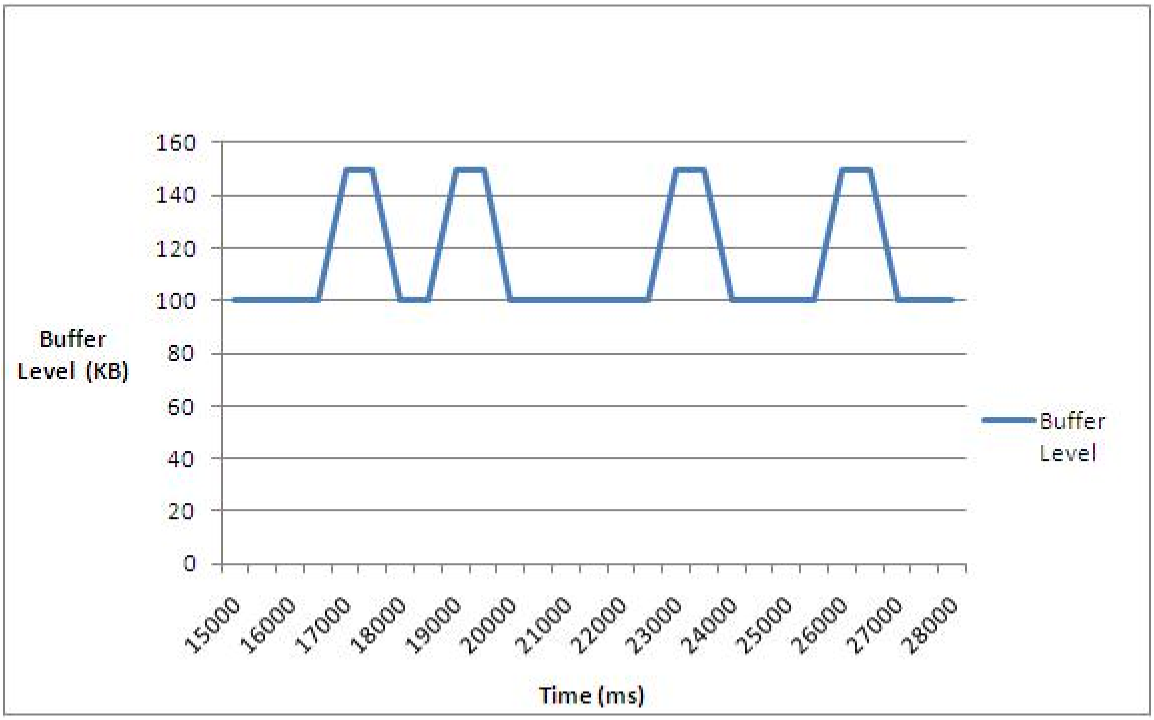

Figure 5.13: Proxy buffer usage during hard handoff

Discussion

The vertical lines indicate the arrival of hard handoff request. The level of the buffer is almost constant along the course of normal work and grows vertically to the arrival of a request: this is because, for hard handoff, the proxy makes use of new frames that had already consumed with the consequent rise in the level. Soon after, however, the proxy returns the buffer to the normal level.

5.4 Performance of the overall handoff support system

This section illustrates the overall functionality of the handoff support system. The illustration focuses on a user running a multimedia application requiring high QoS in a university campus and shows how the components on the system coordinate to support the user during their session.

Aim

The aim of this experiment was to illustrate the functionality of the overall handoff support system.

Results and Discussion

The user starts their session on profile A, which focuses on networks with high QoS. The user's initial interface is Bluetooth connecting via Bluetooth dongles in a laboratory only served with Bluetooth. The QoS mapping on the proxy downscales the content to fit the Bluetooth capacity. The monitored QoS parameters for each network are shown in Figures 5.14, 5.15 and 5.16. The user moves out of laboratory into a hall served with Bluetooth and Wi-Fi. The client stub sends a handoff trigger (Rule 2) to the proxy. The proxy in turn calculates the utilities for each network: 0.5161, 0.4999 and 0.5 for Wi-Fi, UMTS and Bluetooth respectively. The proxy responds with Wi-Fi as the target network and commands the stub to activate the new interface. The proxy then starts double casting the media on the new Wi-Fi path to the client. The double casting lasts 79ms and after that the stub de-activates the Bluetooth interface. The proxy then re-adjusts the video frames to original state that fits Wi-Fi capacity.

At time 8, The user passes nearby a library area served with Wi-Fi and UMTS. The Wi-Fi signal degrades because of many Wi-Fi users sitting in the library. The client stub sends a handoff trigger (Rule 3) to the proxy. The proxy in turn calculates the utilities for each network: 0.5154, 0.5001 and 0.5 for Wi-Fi, UMTS and Bluetooth respectively. The proxy responds with Wi-Fi as the target network and commands the stub to connect to a different Wi-Fi AP. The proxy starts buffering the media until the client has fully connected to the target AP, then starts forwarding the buffered content. This process takes 4013ms to complete.

The user then gets into the university bus which transports people home from university. As the bus start moving out of coverage of the Wi-Fi network, the client stub sends a handoff trigger (Rule 1) to the proxy. The proxy in turn calculates the utilities for each network: 0.0, 0.5509 and 0.0 for Wi-Fi, UMTS and Bluetooth respectively. The proxy responds with UMTS as the target network and commands the stub to activate the new interface. The proxy then starts double casting the media on the new UMTS path to the client. The double casting lasts 95ms and the stub de-activates the Wi-Fi interface. he proxy downscales the content to fit the UMTS capacity. The user then terminates their session.

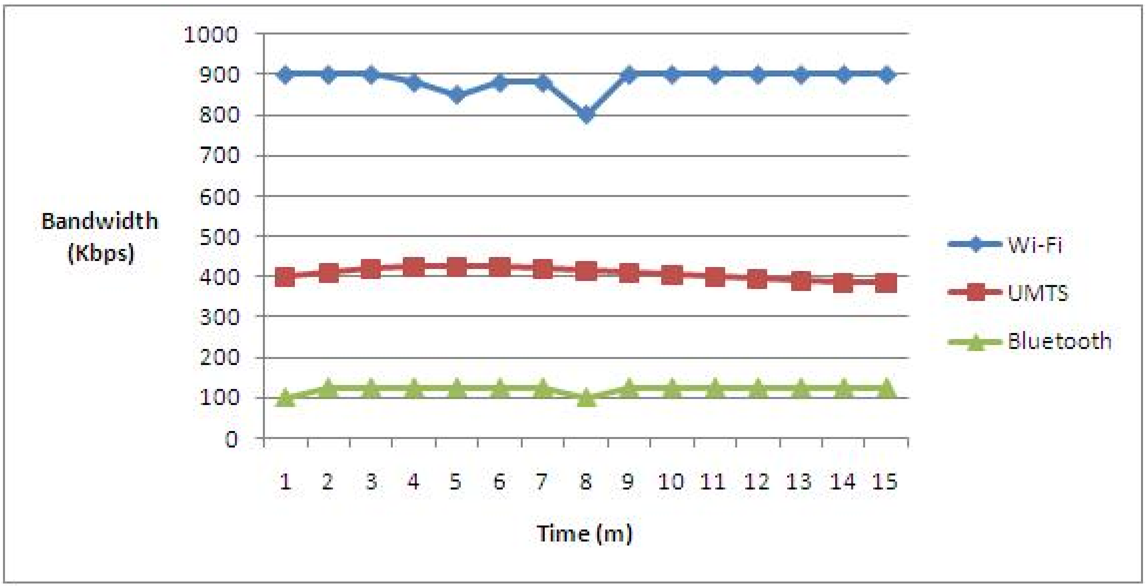

Figure 5.14: Monitored bandwidth for all networks.

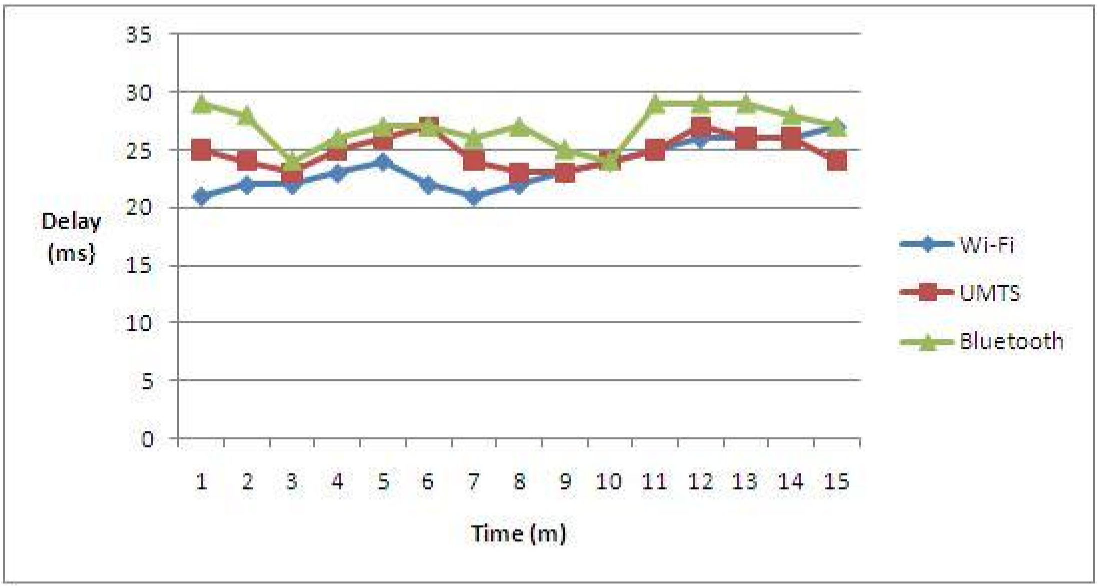

Figure 5.15: Monitored delay for all networks.

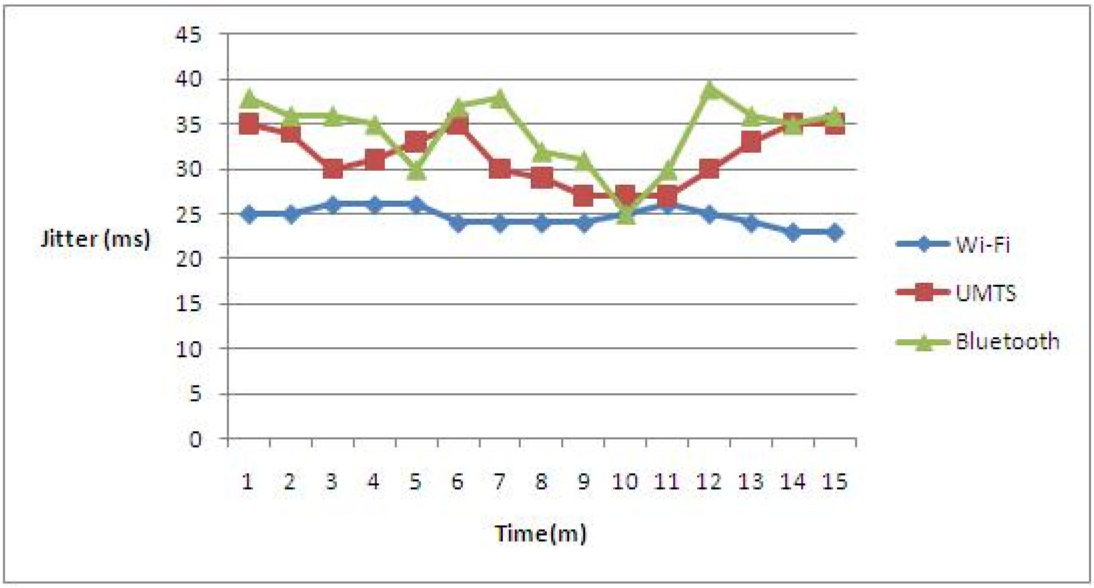

Figure 5.16: Monitored jitter for all networks.

5.5 Summary

The results of the experiments can be summarised as follows.

1. Performance of the handoff decision algorithm. The two MCDM methods (BBN and AHP) employed differ significantly on how they represent the decision problem. AHP is deterministic in nature and models the human decision making approach. It is easier to use, however it is rigid in representing the context information. BBN on the contrary uses probabilistic concepts. As the decision problem becomes complex, BBN become difficult to use. However, BBN is able to model full context information.

   Comparing the correctness in terms of network ranking under varying context changes between the two MCDM methods is not straightforward. However, the two methods produced the same network ranking which proved correctness.

   BBN provide more expressiveness in representing the context information with their ability to model constraints criteria. This is useful in creating more user profiles for customers based on their different needs.

   For BBN, the model can be reduced to one criteria for specialised user profiles. The resulting reduced models produce the same networking ranking as the full models. The reduced models however execute faster than the full models.

AHP uses matrix data structures to represent the MCDM elements, hence faster in terms of execution time of 5ms. BBN which forms a tree-like data structure takes a longer average of 15ms to compute.

2. Performance of the handoff management strategies. Two handoff management strategies (hard and soft) employed differ significantly in execution time. Hard handoff takes an average of 3005ms which is more than most networks handoff latencies', hence undermines service continuity. Furthermore, hard handoff utilises the proxy buffer to a level of 150KB during hard handoff procedure. Soft handoff on the contrary takes an average of 57ms to complete. Soft handoff is the preferred implementation for vertical handoff especially when dealing with applications that are time sensitive. However, soft handoff utilises more battery power on the mobile device as more than one network interface is active during the vertical handoff process. The average time for soft handoff is expected to be higher in real wireless environment.

6 Conclusion

Mobile networks are more and more widespread in our daily life, therefore offering better support for wireless services becomes an important issue [60]. This research focuses on handoff support for wireless access networks, placing more emphasis on the handoff decision process, which go beyond traditional techniques that only considered signal strength. These handoff mechanisms presented in this work incorporate context information into decision making and execution which exemplifies deployment scenarios in future generation networks. The context aware handoff support system proposed in this thesis employs a proxy-based architecture with components running on the wired network, *proxies* and on the mobile terminal, *stubs*. This is in agreement with modern proxy-based architectures which are slowly emerging as a preferred choice to support resource limited clients during handoffs to avoid packet losses and perform handoff decision and execution operations on behalf of the client. The current internet is already populated with different kinds of proxies which are used for authentication. caching, re-directing. Proxy-based solutions can also reduce client-to-server signalling when handoffs occur.

For handoff decision, the thesis proposes Multi-Criteria Decision Making (MCDM) method based on Bayesian Belief Networks (BBN). BBNs have been used extensively in many expert and artificial intelligence systems because of their ability of knowledge representation and reasoning under uncertainty. The implementation of a handoff decision engine based on BBN MCDM was compared to an engine based on Analytical Hierarchy Process (AHP). The implementation of a BBN based handoff decision engine, and its deployment as a proxy functionality has provided the following benefits compared to AHP:

1. Easy modelling of the handoff decision problem and the full representation of the context information.
2. Ability to make decision under uncertainty. In wireless networking, not all information is available during the time of decision making. The ability to make a decision under uncertainty is vital.

3. Ability to model constraints and interdependent criteria.

4. Support for user preferences. The use of utility tables provides the decision maker with the ability to specify their preferences in the decision making.

5. Profile-based handoff decision. The use of utility tables also provides network operators to create differentiated profiles for different users based on their preferences.

However, when considering the execution time, which includes reading the context information and computing the corresponding utilities for all alternative networks, AHP is faster than BBNs by an average of 8ms. The difference of 8ms in execution time between the two methods is not very significant. This is due to the reason that handoff latency of most network technologies in between 100-500ms depending on the speed.

The implemented handoff support system provided two handoff management strategies: Soft handoff and Hard handoff. Soft handoff provides support for vertical handoff scenarios by utilising two active wireless interfaces on the client during the handoff event. The proxy co-ordinates with the client stub to duplicate content destined to the client via the proxy through two service paths. This solution minimises packet losses. However, soft handoff is expensive in terms of battery power as more than one interface is active at the client. Hard handoff provides a support for horizontal handoff scenarios. During a handoff event, hard handoff uses buffering technique at the proxy to buffer the incoming flow from the server to the client. Upon client reconnection, the proxy forwards the buffered flow to the client. Hard handoff takes a significant time of 3000ms to complete and hence can undermine service continuity for time sensitive applications. Hard handoff can also be used to support vertical handoffs, whereby the client reconnects with a new interface.

The handoff support system can be useful to many academic institutions, large companies and telecommunication companies whereby universal access is important. All network providers (public and private) can take advantage of the handoff support system capabilities and the handoff proxies in their private network and configure them to interconnect with other external handoff proxies deployed within other neighbour wireless domains outside their own administration domain, so to form a service proxy federation. The BBN based decision engine can be useful particularly to large service providers in creating different user profiles/classes depending on the different needs of the customers.

Limitation

The MCDM method based on BBN has a scalability problem. When dealing with large and complex problems, constructing the conditional and utility tables for the BBN model becomes difficult. This issue is particularly complex when the conditional probabilities and utilities are obtained from subjective estimate. The utility table size grows exponentially with the increase in number of criteria. In the handoff decision problem, more criteria can be added. This however, renders the tables unmanageable, incurring high error rates and inconsistencies of decisions.

Future Work

The rapidly increase processing capabilities of mobile devices has created incentives for using these devices as computing engines. With such evolution, having the decision engine moved to the mobile device to create a mobile controlled handover is an avenue for future research work involving the implementation of the BBN based decision model on the resource restricted mobile devices. While being investigated in this thesis, the profile-based scalability model is only a preliminary step which reveals the relevance of using such model to implement scalable Bayesian Belief Networks. Further steps and experiments need to be conducted to evaluate the robustness of the approach. This has also been reserved for future work.

Bibliography

[1] T. Ahmed, K. Kyamakya, and M. Ludwig. A context-aware vertical handover decision algorithm for multimode mobile terminals and its performance. In *Proceedings of the IEEE/ACM Euro American Conference on Telematics and Information Systems (EATIS 2006)*, page 1928. ACM, 2006.

[2] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci. Wireless sensor networks: a survey. *Computer Networks 38*, pages 393–422, 2002.

[3] I.F. Akyildiz, J. Xie, and S. Mohanty. Survey of mobility management in next-generation all-ip-based wireless systems. *IEEE Wireless Communications, special issue on Mobility and Resource Management*, pages 16–28, 2004.

[4] Hugin Expert A/S. Hugin. Website, 2009. http://www.hugin.com/.

[5] Hugin Expert A/S. Hugin api manual. Website, 2009. http://download.hugin.com/webdocs/manuals/api-manual.pdf.

[6] S. Balasubramaniam and J. Indulska. Vertical handover supporting pervasive computing in future wireless networks. *Believer Computer Communication 27*, pages 708–719, 2004.

[7] P. Bellavista, M. Cinque, D. Cotroneo, and L. Foschini. Integrated support for handoff management and context awareness in heterogeneous wireless networks. *Proceedings of the 3rd international workshop on Middleware for pervasive and ad-hoc computing*, pages 1–8, 2005.

[8] P. Bellavista, A. Corradi, and L. Foschini. Application-level middleware to proactively manage handoff in wireless internet multimedia. In *Proceedings of the IEEE International Conference On Management of Multimedia Networks and Services*, 2006.

[9] P. Bellavista, A. Corradi, and L. Foschini. Context-aware handoff middleware for transparent service continuity in wireless networks. *Believer Pervasive and Mobile Computing 3*, pages 439–466, 2007.

[10] H. Berndt. *Towards 4G Technologies Services with Initiatives*, chapter Introducing the 4G Mobile Adventure. John Wiley and Sons, Ltd, 2008.

[11] J. Bray and C. Sturman. *Bluetooth: Connect Without Cables*. Prentice Hall, 2000.

[12] Broadcom. The new mainstream wireless lan standard. Website, 2003. http://www.dell.com/downloads/global/shared/broadcom_802_11_g.pdf.

[13] J.B. Carruthers. Wireless infrared communications. *Wiley Encyclopedia Of Telecommunications*, 2002.

[14] C. Castelluccia. Hmipv6: A hierarchical mobile ipv6 proposal. *ACM SIGMOBILE Mobile Computing and Communications Review*, pages 48–59, 2000.

[15] L. Chen, T. Sun, B. Chen, V. Rajendran, and M. Gerla. A smart decision model for vertical handoff. In *In Proceedings 4th ANWIRE International Workshop on Wireless Internet and Reconfigurability*, 2004.

[16] D. Chow. Cellular wireless network handoff protocols. Website, 2005. http://www.articlesnatch.com/Article/Cellular-Wireless-Network-Handoff-Protocols/81941.

[17] G.F. Cooper. A method for using belief networks as influence diagrams. In *In Proceedings of the Twelfth Conference on Uncertainty in Artificial Intelligence*, pages 55–63, 1988.

[18] C. de Campos and Q. Ji. Strategy selection in influence diagrams using imprecise probabilities. In *Proceedings of the Twenty-Fourth Annual Conference on Uncertainty in Artificial Intelligence (UAI-08)*, pages 121–128. AUAI Press, 2008.

[19] B.E. Fad. Analytical hierarchy process (ahp) approach to size. estimation. PRICE Systems L.L.C., 2005.

[20] N. Fenton and M. Neil. Making decisions: Using bayesian nets and mcda. *Knowledge-Based Systems*, pages 307–325, 2001.

[21] L. Foschini. *Support Infrastructures for Multimedia Services with Guaranteed Continuity and QoS*. PhD thesis, University of Bologna, 2007.

[22] V.K. Garg. *Wireless Network Evolution: 2G to 3G*. Prentice Hall PTR, 2001.

[23] M. Gast. *802.11 Wireless Networks: The Definitive Guide*. O'Reilly, 2002.

[24] D. Ghosh. Mobile ip. Website, 2000. http://www.acm.org/crossroads/xrds7-2/mobileip.html.

[25] M.K. Goff. The scope of network distributed computing. website, 2004.

[26] T. Halonen, J. Romero, and J. Melero. *GSM, GPRS and EDGE performance: evolution towards 3G/UMTS*. John Wiley and Sons, 2003.

[27] D. Heckerman. A tutorial on learning with bayesian networks. Technical report, Microsoft Research, 1995.

[28] R. Hsieh, Z.G. Zhou, and A. Seneviratne. S-mip: A seamless handoff architecture for mobile ip. *in Proceedings of INFOCOM*, pages 1774–1784, 2003.

[29] S. Hussain, Z. Hamid, and N.S. Khattak. Mobility management challenges and issues in 4g heterogeneous networks. In *Proceedings of the First International Conference on Integrated Internet Ad hoc and Sensor Networks*, 2006.

[30] M. Kassar, B. Kervella, and G. Pujolle. An overview of vertical handover decision strategies in heterogeneous wireless networks. *Computer Communications*, pages 2607–2620, 2008.

[31] U. Kieerulff and A.Madsen. *Bayesian Networks and Influence Diagrams*. Springer, 2007.

[32] H. Kung and F. Kuo. Dynamic user-oriented qos specification and mapping for multimedia differentiation services. *International Journal of Computer Science and Network Security (IJCSNS)*, pages 116–122, 2006.

[33] S.L. Lauritzen and D. Nilsson. Representing and solving decision problems with limited information. *Management Science*, 47:1235–1251, 2001.

[34] A. Ezil Sam Leni and S.K. Srivatsa. A handoff technique to improve tcp performance in next generation wireless networks. *Information Technology Journal*, pages 504–509, 2008.

[35] C. Li, Y. Tseng, and H. Chao. A neighbor caching mechanism for handoff in ieee 802.11 wireless networks. In *Proceedings of the 2007 International Conference on Multimedia and Ubiquitous Engineering*, 2007.

[36] LIA. Mobile agent based ubiquitous multimedia middleware (mum). Website, 2009. http://lia.deis.unibo.it/Research/MUM/.

[37] A. McCrory. Ubiquitous, pervasive, sorry, they don't compute. website, 2000.

[38] Sun Microsystems. Java media framework api (jmf). Website, 2009. http://java.sun.com/javase/technologies/desktop/media/jmf/.

[39] A. Mishra, M. Shin, and W. Arbaugh. An empirical analysis of the ieee 802.11 mac layer handoff process. Technical report, ACM SIGCOMM Computer Communication Review, 2003.

[40] L. Mokhesi and A. Bagula. Context-aware handoff decision for wireless access networks using bayesian networks. In *Proceedings of Southern African Institute for Computer Scientists and Information Technologists (SAICSIT)*, pages 104–111. ACM International Series, 2009.

[41] M. Neil, N. Fenton, and L. Nielsen. Building large-scale bayesian networks. *The Knowledge Engineering Review*, pages 257–284, 2000.

[42] Y. Nkansah-Gyekye and J. I. Agbinya. Vertical handoff decision algorithm for umts-wlan. In *Proc. of the 2nd International Conference on Wireless Broadband and Ultra Wideband Communications*, 2007.

[43] P. Pawar, K. Wac, and B. van Beijnum. Context-aware middleware architecture for vertical handover support to multi-homed nomadic mobile services. In *Proceedings of the 2008 ACM symposium on Applied computing*, pages 481–488. ACM, 2008.

[44] C. Perkins, S. Alpert, and B. Woolf. *Mobile IP: Designs Principles and Practices*. Addison-Wesley Longman Publishing Co., Inc, 1997.

[45] M. Pirdashti, A. Ghadi, M. Mohammadi, and G. Shojatalab. Multi-criteria decision-making selection model with application to chemical engineering management decisions. *World Academy of Science, Engineering and Technology 49*, pages 54–59, 2009.

[46] C. Prehofer. *Towards 4G Technologies Services with Initiatives*, chapter Context-aware Mobility Management. John Wiley and Sons, Ltd, 2008.

[47] C. Prehofer, C. Bettstetter, and J. Widmer. *Towards 4G Technologies Services with Initiatives*, chapter Mobile Communication Networks. John Wiley and Sons, Ltd, 2008.

[48] D. Ronzani. The battle of concepts: Ubiquitous computing, pervasive computing and ambient intelligence. *Ubiquitous Computing and Communication Journal*, pages 100–111, 2007.

[49] T. Saaty. *Multicriteria Decision Making, The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation.* RWS Publications, 1990.

[50] T. Saaty. Decision making with the analytic hierarchy process. *International Journal of Services Sciences*, pages 83–98, 2008.

[51] D. Saha, A. Mukherjee, I.S. Misra, and M. Chakraborty. Mobility support in ip: A survey of related protocols. Technical report, IEEE Network, 2004.

[52] B.N. Schilit, N. Adams, and R. Want. Context-aware computing applications. *Proceedings of the 1994 First Workshop on Mobile Computing Systems and Applications*, pages 85–90, 1994.

[53] Searchmobilecomputing. Access point. Website, 2003. http://searchmobilecomputing.techtarget.com/definition/access-point.

[54] N. Shenoy and S. Mishra. *Heterogeneous Wireless Access Networks*, chapter Vertical Handoff and Mobility Management for Seamless Integration of Heterogeneous Wireless Access Technologies. Springer US, 2008.

[55] E. Shim, H. Wei, Y. Chang, and R.D. Gitlin. Low latency handoff for wireless ip qos with neighborcasting. In *In Proceedings of IEEE International Conference on Communications*, pages 3245–3249, 2002.

[56] C. Townsend and S. Arms. *Sensor Technology HandBOOK*, chapter Wireless Sensor Networks: Principles and Applications. Believer Inc, 2005.

[57] S. Vijayarangam and S. Ganesan. An architecture for mpls implementation in wireless networks. *Proceedings of Electro/Information Technology Conference (EIT)*, 2001.

[58] J. Wang and L. Bao. Mobile context handoff in distributed ieee 802.11 systems. *International Conference on Wireless Networks, Communication, and Mobile Computing*, pages 13–16, 2005.

[59] W. Watthayu and Y. Peng. A bayesian network based framework for multi-criteria decision making. In *Proceedings of the 17th International Conference on Multiple Criteria Decision Analysis*, 2004.

[60] Q. Wei, K. Farkas, C. Prehofer, and P. Mendes. Context-aware handover using active network technology. *Computer Networks: The International Journal of Computer and Telecommunications Networking*, pages 2855 – 2872, 2006.

[61] M. Weiser. Some computer science issues in ubiquitous computing. *Communications of the ACM 36*, pages 75–84, 1993.

[62] Wisegeek. What is hsdpa? Website, 2010. http://www.wisegeek.com/what-is-hsdpa.htm.

[63] Q. Zeng and D. Agrawal. *HandBOOK of Wireless Networks and Mobile Computing*, chapter Handoff in Wireless Mobile Networks. John Wiley and Sons, Ltd, 2002.

[64] W. Zhang. Handover decision using fuzzy madm in heterogeneous networks. In *In proceedings: IEEE Wireless Communications and Networking Conference*, 2004.

[65] F. Zhu and J. McNair. Multiservice vertical handoff decision algorithms. *EURASIP Journal on Wireless Communications and Networking*, pages 1–13, 2006.