Design & Simulation of Intelligent Multi-Criteria Vertical Handover Model for Wireless NGN

تصميم ومحاكاة نموذج ذكي ومقصود معايير المناولة الرأسية في شبكات الجيل القادم اللاسلكية

Report submitted in partial fulfillment of the requirement for the degree

of

M.SC.

In

TELECOMMUNICATION ENGINEERING

By

SAYID GUDB IBRAHIM HASSAN

Under the Supervision of

Dr. ASHRAF GASIM EL-SEED ABD-ALLAH

To

Department of Electronic Engineering

Sudan University

of Science and Technology

June 2009

بسم الله الرحمن الرحيـــــــــــــم

(قُلْ إِنَّ صَلَاتِي وَنُسُكِي وَمَحْيَايَ وَمَمَاتِي لِلَّهِ رَبِّ الْعَالَمِينَ)

الأنعام - 162

Dedication

- *To whom the paradise under her feet, spring of love...*

  Mother.

- *To who gave me and taught me the meaning of life…*

  Father.

- *To my Brothers, Sisters, Family and Colleagues…*

- *To anyone taught me a letter…*

- *To all my friends and lovers all over the world…*

ACKNOWLEDGEMENT

*Firstly, all praise and my great thanks be to my Glory God, **ALLAH**, for everything in my life, and for giving me the power and assisting me to complete this work.*

*I am profoundly grateful to **Dr. Ashraf Gasim El-seed Abd-allah,** (my supervisor) whose detailed criticisms of earlier drafts, valuable suggestions and wise counsel, have proved most invaluable to me. It has been a most rewarding learning experience to work under her guidance.*

I thank all individuals who sacrificed their time to speak to me during the work. I wish to acknowledge with great appreciation all the encouragement from my friends; your advice were soul searching.

List of Figures

List of Tables

Abbreviations

1G	First Generation
2G	Second Generation
3G	Third Generation
3GPP	Third Generation Partnership Project
4G	Fourth Generation
AAA	Authentication, Authorization, And Accounting
AAL2	ATM Adaptation Layer Type 2
AMPS	Advanced Mobile Phone Service
ANFIS	Adaptive Neuro-Fuzzy Inference System
ANN	Adaptive Neural Network
ATM	Asynchronous Transfer Mode
CDMA	Code Division Multiple Access
CS	Circuit Switching
EDGE	Enhanced Data GSM Environment
FDMA	Frequency Division Multiple Access
FL	Fuzzy Logic
FLS	Fuzzy Logic System
GPRS	General Packet Radio Service
GSM	Global System For Mobile Communications
HBP	Handoff Blocking Probability
HDE	Handover Decision Engine
HO	Hand Over / Hand Off
IMT-2000	International Mobile Telecommunications 2000
IP	Internet Protocol
ITU	International Telecommunication Union
MAC	Media Access Control
MF	Membership Function
MIMO	Multiple-Input Multiple-Output
OFDM	Orthogonal Frequency Division Multiplexing
PDC	Personal Digital Cellular
PLMN	Public Land Mobile Network
PS	Packet Switching
PSTN	Public Switched Telecommunications Network
QoS	Quality Of Service
SIM	Subscriber Identity Module
SMS	Short Message Service
TCP	Transmission Control Protocol
TDMA	Time Division Multiple Access
UMTS	Universal Mobile Telecommunication System
USIM	User Service Identity
UTRAN	UMTS Terrestrial Radio Access Network
VHO	Vertical Handover
VHDM	Vertical Handover Decision Making
VHE	Virtual Home Environment

VoIP	Voice Over Internet Protocol
WCDMA	Wide Band Code Division Multiple Access
WiBro	Wireless Broadband
WLAN	Wireless Local Area Network

Contents

المستخلص

إن النمو السريع للطلبِ على الخدماتِ المُخْتَلِفةِ للإتصال اللاسلكي قاد قطاع الصناعةَ والباحثين للتحرّي حول جيلٍ جديدٍ مِنْ أنظمةِ الإتصال. أحد الأهدافِ الرئيسيةِ في تَصميم وتَطوير هذه الأنظمةِ أَنْ تُزوّدَ مستخدمي الهواتف النّقالةَ بالإتّصال المتين إلى الشبكاتِ المختلفةِ بحيث يتمكن المستعملون من أَنْ يَتحرّكوا بحرية بين الشبكاتِ المتباينةِ بينما يستعملون تطبيقاتَهم.

المناولة (HO) هي العملية التي تَبقي إتّصال مستخدم الهاتف النّقال نشطاً بينما يَتحرّكُ داخل الشبكة اللاسلكية ؛ والمناولة الرأسية (VHO) تُعنى بالمناولة بين الأنواع المختلفةِ مِنْ الشبكاتِ. إن تَحسين عمليةِ المناولة الرأسية يمثل قضيةً مهمةً للبحثِ، حيث تُؤدّي إلى تخفيض إشارات الاتخاطب داخل الشبكةِ وتقليل خسارةِ طاقة جهاز الجوال، ومن الناحية الأخرى تُحسّنُ نوعية الخدمةِ (QoS) للشبكةِ.

في هذا المشروع تم تصميم نموذج ذكي متعدد المعاييرَ للمناولة الرأسية بين شبكتين متباينتين؛ شبكة تستخدم تقنية الجيل الثالث (UMTS) والأخري تشمل شبكات محلية لاسلكية (WLAN). تَستعملُ هذه الخوارزميةِ طريقة ذكية بمعلوماتٍ متعددةِ المعالم تم الحصول عليها مِنْ بيئة الشبكاتِ المختلفة لتَحديد أفضل شبكةِ متوفرةِ لتتم عملية المناولة إليها.

تمت مُناقشةُ مفهومَ النموذج المُقتَرَح إضافةً إلى تصميم الطريقة الذكية ، وبعد ذلك أُستُخدِمت أدوات Matlab لإشْتِقاق المحاكاةِ.و أخيراً، تم تحليل أداء النموذج ؛و أظهرت النَتائِجُ تحسناً مطلوباً في عمليةِ المناولة الرأسية (VHO).

Abstract:

Rapid growth of demand for various wireless communication services has led the industry and the researchers to investigate on new generation of telecommunication systems. One of the main objectives in designing and developing these systems is to provide the mobile users with a robust connection to different networks so that the users can move freely between heterogeneous networks while running their applications.

Handoff (HO) is a process which maintains a mobile user's active connection as it moves within wireless networks; and vertical handoff (VHO) considers handoff between different types of networks. Optimizing vertical handoff process is an important issue of research, which leads to reduction of network signaling and mobile device power loss and on the other hand improves network quality of service (QoS).

In this project an intelligent handoff model uses multi-criteria has been proposed for two heterogeneous network; UMTS and WLAN. This model uses intelligent method with multi-parameters information acquired from the networks to decide the best available network to handoff to.

The concept of proposed model is discussed besides constructing the intelligent method, and then the MATLAB tools are used to derive the simulation. Finally, the model's performance is analyzed and results have shown desired enhancement in the VHO process.

1. INTRODUCTION

1.1. Background

Handoff process is one of the most important parts of any wireless network. A new handoff process considered for next generation wireless networks is vertical handoff. While handoff process defines the transfer of current mobile active connection between cells within a single layer, vertical handoff defines this process between different layers. Research on handoff is mainly categorized in two parts.

First, works done from upper network layers point of view, which focus on data transfer and performance of network protocols, like internet protocol (IP) and transmission control protocol (TCP).

Second, works done from physical layer point of view, which mostly deal with radio propagation characteristics and handoff decision process. Variation of signal power causes some unnecessary handoffs on cell boundaries, which is called ping pong effect. Ping pong increases network traffic load and this indirectly leads to increase of handoff blocking probability (HBP). It also causes power loss in transmitters and receivers. Crossover point is another metric to compare different vertical handoff algorithms and it is the distance between the point where the mobile makes handoff, and the cell boundary. An ideal algorithm makes handoff on cells and layers boundary. If crossover distance is long it causes increase of channel interference and link QoS degradation.

Heterogeneous networks environment suggests using different cellular networks with different sizes and probably different technologies, so that the user can still preserve its connection while going out of the coverage of a lower layer, because it is still in the coverage of an upper layer. Another benefit of such networks is the capability of adapting network parameters with user's specifications like speed and demanded service. Using different types of networks (heterogeneous networks), their coordination, integration and vertical handoff procedures is a significant challenge for third generation (3G) and fourth generation (4G) wireless networks.

1.2. Problem Statements:

When a mobile station is being under the coverage of one network or more, there are many factors that affect the strength and quality of the signals received and thus the QoS achieved by the mobile station.

This thesis is focusing in the problem of Vertical Handover Decision Making (VHDM) that determine which network the mobile should use, and when to move to that selected network based on the analysis of previous factors. The analysis and all work done will be from physical layer point of view.

1.3. Objective:

In this project an intelligent handoff model uses multi-criteria has been proposed for two heterogeneous network; UMTS and WLAN. This model uses intelligent method with multi-parameters information acquired from the networks to decide the best available network to handoff to.

So, the main idea behind this model is to build an intelligent Handover Decision Engine (HDE) to manage the multiple input parameters' data that collected from the heterogeneous environment; and decide which networks is preferred to be used.

1.4. Approach

To build the intelligent method, a combination of Fuzzy Logic System (FLS) and Adaptive Neural Network (ANN) is used. This method is known as Adaptive Neuro-Fuzzy Inference System (ANFIS), which is an adaptive network that permits the usage of neural network topology together with fuzzy logic. It not only includes the characteristics of both methods, but also eliminates some disadvantages of their lonely-used case.

First, the concept of proposed model is discussed besides constructing the intelligent method, and then the MATLAB tools are used to derive the simulation. Finally, the model performance is analyzed; results have shown desired enhancement in the VHO process compared with traditional algorithm.

1.5. Outline of Thesis

On the coming three chapters a literature review and background of the main topics related to the proposed model are figured out with highlighting its aspects and standards.

A Cellular Network and WLAN general preview, history and main concepts of heterogeneous networks are discussed in Chapter 2; with focusing on main features and problems that affect the strength and QoS.

In the 3rd chapter, the main concern will be the studying of handover between heterogeneous networks – Vertical Handover – and some algorithms applied in this field, and determining the main factors they relied on.

The used intelligent method, ANFIS, is detailed in Chapter 4. Starting with and introduction of the Fuzzy Logic System (FLS) and Adaptive Neural Network (ANN), and how the ANFIS combined them together besides its advantages over them.

Chapter 5 will be the core of the Thesis, here the structure and main concepts of the proposed model are figured, full analysis of its parts and simulation of the model is performed, and finally the results of simulation are compared with results of some conventional algorithms.

At last, all contributions and the work done through thesis are summarized besides highlighting the future works opportunities and recommendations on the last chapter, chapter 6.

2. CELLULAR NETWORK AND WLAN BACKGROUND

2.1. Mobile Network overview

The History and evolution of mobile services from the 1G (first generation) to fourth generation are discussed in this section. Table2.1 below presents a short summary of the history of mobile telephone technologies.

Table 1 Mobile Telephony Specifications

Technology	**1G**	**2G**	**2.5 G**	**3G**	**4G**
Design Began	1970	1980	1985	1990	2000
Implementation	1984	1991	1999	2002	2010
Service	Analog Voice, Synchronous data to 9.6 kbps	Digital Voice, short messages	Higher capacity, packetized data	Higher capacity, broadband data up to 2 Mbps	Higher capacity, completely IP-oriented, multimedia, data to hundreds of megabits
Standards	AMPS, TACS, NMT, etc	TDMA, CDMA, GSM, PDC	GPRS, EDGE, 1xRTT	WCDMA, CDMA2000	Single standard
Data Bandwidth	1.9 kbps	14.4 kbps	384 kbps	2 Mbps	200 Mbps
Multiplexing	FDMA	TDMA, CDMA	TDMA, CDMA	CDMA	CDMA?
Core Network	PSTN	PSTN	PSTN, packet network	Packet network	Internet

2.1.1. *First Generation, 1G*

The first wireless cellular systems (1G) started appearing in the 1980s. 1G network are based on the AMPS (Advanced Mobile Phone Service) standard. Unlike their predecessor wireless networks, 1G network are based on the idea of cells. 1G network provide analog voice service but no data service. 1G was analog, not digital. The spectral efficiency of 1G network was very low and the effective "energy/bit" was high. Handsets had short talk/standby times.

2.1.2. *Second Generation, 2G*

The 2G wireless networks are digital networks (for spectral efficiency and not for digital services). There are several 2G standards in use:

2.1.2.1. GSM:

Stands for Global System for Mobile Communications, widely used in Europe and countries other than USA, now appearing in the USA. *It* is an open, digital cellular technology used for transmitting mobile voice and data services. It supports voice calls and data transfer speeds of up to 9.6 kbit/s, together with the transmission of SMS (Short Message Service). GSM operates in the 900MHz and 1.8GHz bands in Europe and the 1.9GHz and 850MHz bands in the US. The 850MHz band is also used for GSM and 3G in Australia, Canada and many South American countries. By having harmonized spectrum across most of the globe, GSM's international roaming capability allows users to access the same services when traveling abroad as at home. This gives consumers seamless and same number connectivity in more than 218 countries. Terrestrial GSM networks now cover more than 80% of the world's population. GSM satellite roaming has also extended service access to areas where terrestrial coverage is not available.

GSM Technology Differentiator

One of the advantages of GSM is that it offers a subscriber identity module (SIM), also known as a smart card. The smart card contains a computer chip and

some non-volatile memory and is inserted into a slot in the base of the mobile handset.

The memory on the smart card holds information about the subscriber that enables a wireless network to provide subscriber services. The information includes:

- The subscriber's identity number
- The telephone number
- The original network to which the subscriber is subscribed

A smart card can be moved from one handset to another. A handset reads the information off the smart card and transmits it to the network.

GSM Network Elements are shown in Figure2-1.

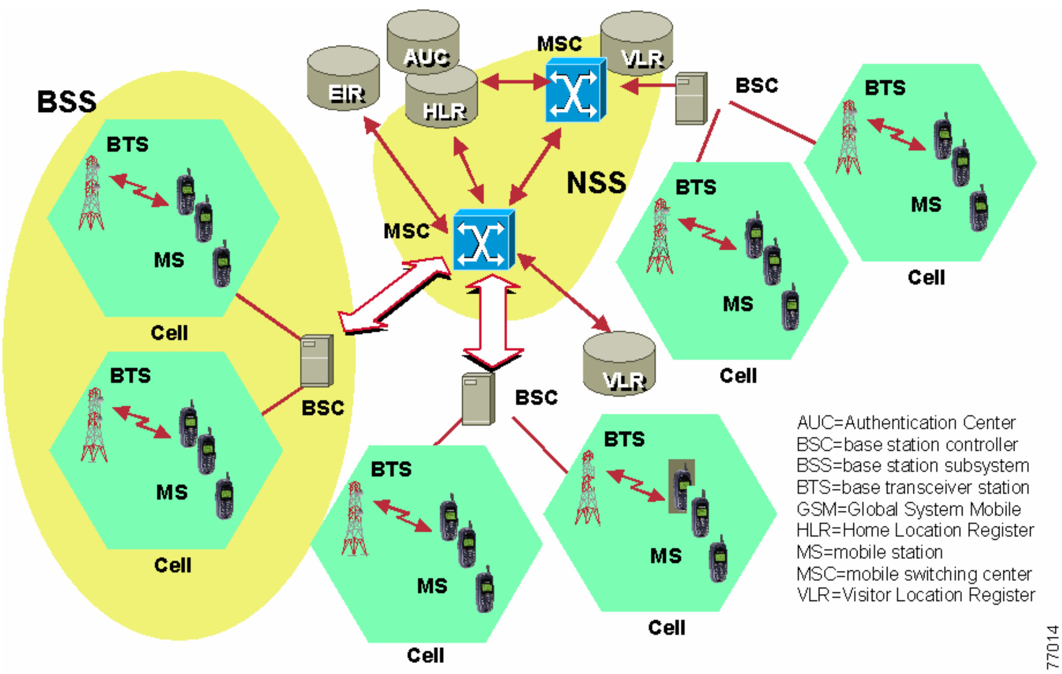

Figure 2-1 GSM Network Components

The GSM network components are shown in Figure2.1. It consists of the following network:

<u>*Mobile Station*</u>

The mobile station (MS) is the starting point of a mobile wireless network. The MS can contain the following components:

- Mobile terminal (MT)—GSM cellular handset
- Terminal equipment (TE)—PC or personal digital assistant (PDA)

The MS can be two interconnected physical devices (MT and TE) with a point-to-point interface or a single device with both functions integrated.

Base Transceiver Station When a subscriber uses the MS to make a call in the network, the MS transmits the call request to the base transceiver station (BTS). The BTS includes all the radio equipment (i.e., antennas, signal processing devices, and amplifiers) necessary for radio transmission within a geographical area called a cell. The BTS is responsible for establishing the link to the MS and for modulating and demodulating radio signals between the MS and the BTS.

Base Station Controller

The base station controller (BSC) is the controlling component of the radio network, and it manages the BTSs. The BSC reserves radio frequencies for communications and handles the handoff between BTSs when an MS roams from one cell to another. The BSC is responsible for paging the MS for incoming calls.

Base Station Subsystem

A GSM network is comprised of many base station subsystems (BSSs), each controlled by a BSC. The BSS performs the necessary functions for monitoring radio connections to the MS, coding and decoding voice, and rate adaptation to and from the wireless network. A BSS can contain several BTSs.

Mobile Switching Center

The mobile switching center (MSC) is a digital ISDN switch that sets up connections to other MSCs and to the BSCs. The MSCs form the wired (fixed) backbone of a GSM network and can switch calls to the public switched telecommunications network (PSTN). An MSC can connect to a large number of BSCs.

Equipment Identity Register

The equipment identity register (EIR) is a database that stores the international mobile equipment identities (IMEIs) of all the mobile stations in the network. The IMEI is an equipment identifier assigned by the manufacturer of the mobile station. The EIR provides security features such as blocking calls from handsets that have been stolen.

Home Location Register

The home location register (HLR) is the central database for all users to register to the GSM network. It stores static information about the subscribers such as the international mobile subscriber identity (IMSI), subscribed services, and a key for authenticating the subscriber. The HLR also stores dynamic subscriber information (i.e., the current location of the mobile subscriber).

Authentication Center

Associated with the HLR is the authentication center (AuC); this database contains the algorithms for authenticating subscribers and the necessary keys for encryption to safeguard the user input for authentication.

Visitor Location Register

The visitor location register (VLR) is a distributed database that temporarily stores information about the mobile stations that are active in the geographic area for which the VLR is responsible. A VLR is associated with each MSC in the network. When a new subscriber roams into a location area, the VLR is responsible for copying subscriber information from the HLR to its local database. This relationship between the VLR and HLR avoids frequent HLR database updates and long distance signaling of the user information, allowing faster access to subscriber information.

The HLR, VLR, and AuC comprise the management databases that support roaming (including international roaming) in the GSM network. These databases authenticate calls while GSM subscribers roam between the private network and the public land mobile network (PLMN). The types of information they store include subscriber identities, current location area, and subscription levels.

Network and Switching Subsystem

The network and switching subsystem (NSS) is the heart of the GSM system. It connects the wireless network to the standard wired network. It is responsible for the handoff of calls from one BSS to another and performs services such as charging, accounting, and roaming.

The other standards of 2G besides GSM are:

2.1.2.2. TDMA

Stands for Time Division Multiple Access, it is used primarily in the USA.

2.1.2.3. CDMA

Stands for Code Division Multiple Access. And it is used in USA and its use is spreading in the rest of the world.

2.1.2.4. PDC

Stands for Personal Digital Cellular. It is used only in Japan where iMode uses packet switched PDC.

Some characteristics of 2G networks are:

- Maximum data rates of 9.6 Kbits/second to 14.4 Kbits/second if you are in just the right place.
- Digital voice (results in a lower quality voice but uses less precious spectrum).
- Enhanced telephony features such as caller-id.
- Services such as text based messaging (big winner), downloads of still images and audio clips, etc.

2.1.3. *The 2.5G (Between 2G and 3G)*

2.5G networks, which are still not available everywhere, are essentially General Packet Radio Service (GPRS) packet overlays on 2G networks. Besides enhancing GSM and TDMA networks by making them packet-based networks, GPRS also increases their data rates. GPRS is primarily a software upgrade of GSM. Some characteristics of 2.5G networks are:

- Data rates of 64 - 144kb/second.
- Packet based.
- Always-on connectivity.
- Instant messaging with small attachments.

A new wireless standard, Enhanced Data GSM Environment (EDGE), has been developed to increase the bandwidth of GPRS. EDGE triples the bandwidth capacity of GPRS to 384 Kbits/second thus allowing GSM and TDMA operators to offer high-speed services. EDGE based networks fall in between 2.5G and 3G networks. The GPRS Network Elements are shown in Figure 2-2.

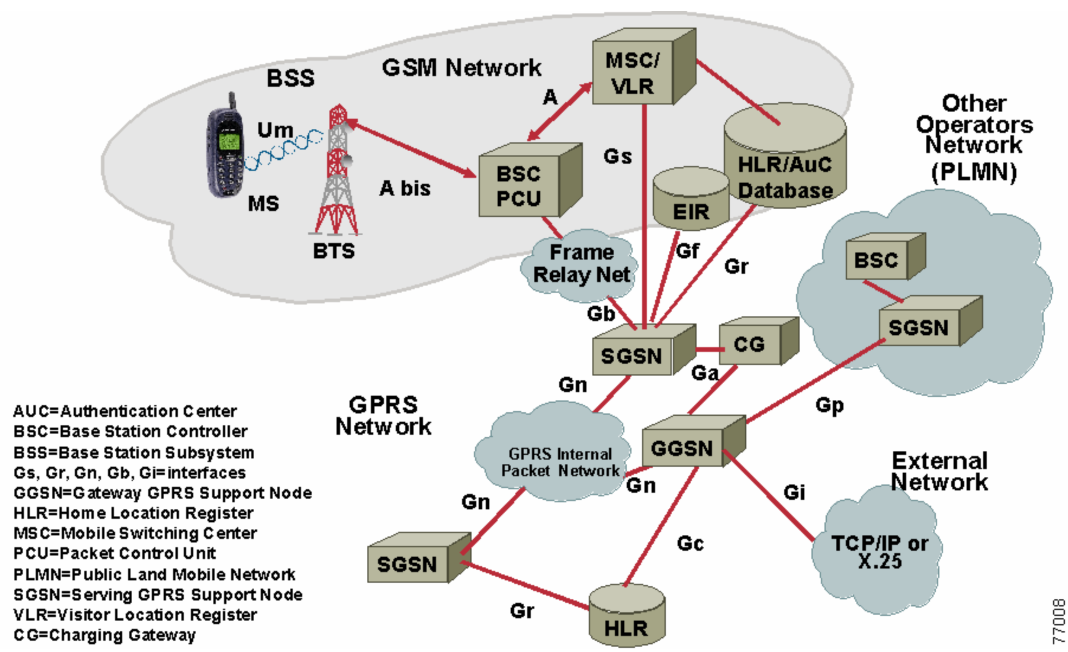

Figure 2-2 Network Elements of 2.5 G

There are some Modifications or Upgrade required for the existing GSM network in order to support GPRS Network. This upgrade includes the following elements:

- *Terminal Equipment (TE):* New terminal equipment is required to access GPRS services. These new terminals will be backward compatible with GSM for voice calls.
- *BTS:* A software upgrade is required in the existing base transceiver site.
- *BSC:* The base station controller (BSC) requires a software upgrade and the installation of new hardware called the packet control unit (PCU). The PCU directs the data traffic to the GPRS network and can be a separate hardware element associated with the BSC.
- *GPRS Support Nodes (GSNs):* The deployment of GPRS requires the installation of new core network elements called the serving GPRS support node (SGSN) and gateway GPRS support node (GGSN).

- *Databases (HLR, VLR, etc.):* All the databases involved in the network will require software upgrades to handle the new call models and functions introduced by GPRS.

Further enhancements to GSM networks are provided by Enhanced Data rates for GSM Evolution (EDGE) technology, which provides up to three times the data capacity of GPRS.

Using EDGE, operators can handle three times more subscribers than GPRS, triple their data rate per subscriber, or add extra capacity to their voice communications. It allows the delivery of advanced mobile services such as the downloading of video and music clips, multimedia messaging, high-speed Internet access and e-mail on the move.

EDGE uses the same structure, as today's GSM networks, which allows it to be overlaid directly onto an existing GSM network. For many existing GSM/GPRS networks, EDGE is a simple software-upgrade. Due to the very small incremental cost of including EDGE capability in GSM network deployment, virtually all new GSM infrastructure deployments are also EDGE capable and nearly all new mid- to high-level GSM devices also include EDGE radio technology.

2.1.4. 3G Systems

The 3G Systems are intended to provide a global mobility with wide range of services including telephony, paging, messaging, Internet and broadband data. International Telecommunication Union (ITU) started the process of defining the standard for third generation systems, referred to as International Mobile Telecommunications 2000 (IMT-2000). In Europe European Telecommunications Standards Institute (ETSI) was responsible of UMTS standardization process. In 1998 Third Generation Partnership Project (3GPP) was formed to continue the technical specification work. 3GPP has five main UMTS standardization areas: Radio Access Network, Core Network, Terminals, Services and System Aspects and GERAN.

Third Generation Partnership Project 2 (3GPP) was formed for technical development of cdma2000 technology which is a member of IMT-2000 family. In February 1992 World Radio Conference allocated frequencies for UMTS use. Frequencies 1885 - 2025 and 2110 - 2200 MHz were identified for IMT-2000 use. See the UMTS Frequency page for more details. All 3G standards are still under constant development. In 1999 ETSI Standardization finished for UMTS Phase 1 (Release '99, version 3) and next release is due December 2001. UMTS History page has a list of all major 3G and UMTS milestones. Most of the European countries and some countries round the world have already issued UMTS licenses either by beauty contest or auctions.

2.1.4.1. *UMTS Services*

UMTS offers teleservices (like speech or SMS) and bearer services, which provide the capability for information transfer between access points. It is possible to negotiate and renegotiate the characteristics of a bearer service at session or connection establishment and during ongoing session or connection. Both connection-oriented and connectionless services are offered for Point-to-Point and Point-to-Multipoint communication.

Bearer services have different QoS parameters for maximum transfer delay, delay variation and bit error rate. Offered data rate targets are:

- 144 Kbits/s satellite and rural outdoor
- 384 Kbits/s urban outdoor
- 2048 Kbits/s indoor and low range outdoor

UMTS network services have different QoS classes for four types of traffic:

- Conversational class (voice, video telephony, video gaming)
- Streaming class (multimedia, video on demand, webcast)
- Interactive class (web browsing, network gaming, database access)
- Background class (email, SMS, downloading)

UMTS will also have a Virtual Home Environment (VHE). It is a concept for personal service environment portability across network boundaries and between terminals. Personal service environment means that users are consistently presented with the same personalized features, User Interface customization and services in whatever network or terminal, wherever the user may be located. UMTS also has improved network security and location based services.

UMTS Architecture

A UMTS network consist of three interacting domains; Core Network (CN), UMTS Terrestrial Radio Access Network (UTRAN) and User Equipment (UE). The main function of the core network is to provide switching, routing and transit for user traffic. Core network also contains the databases and network management functions.

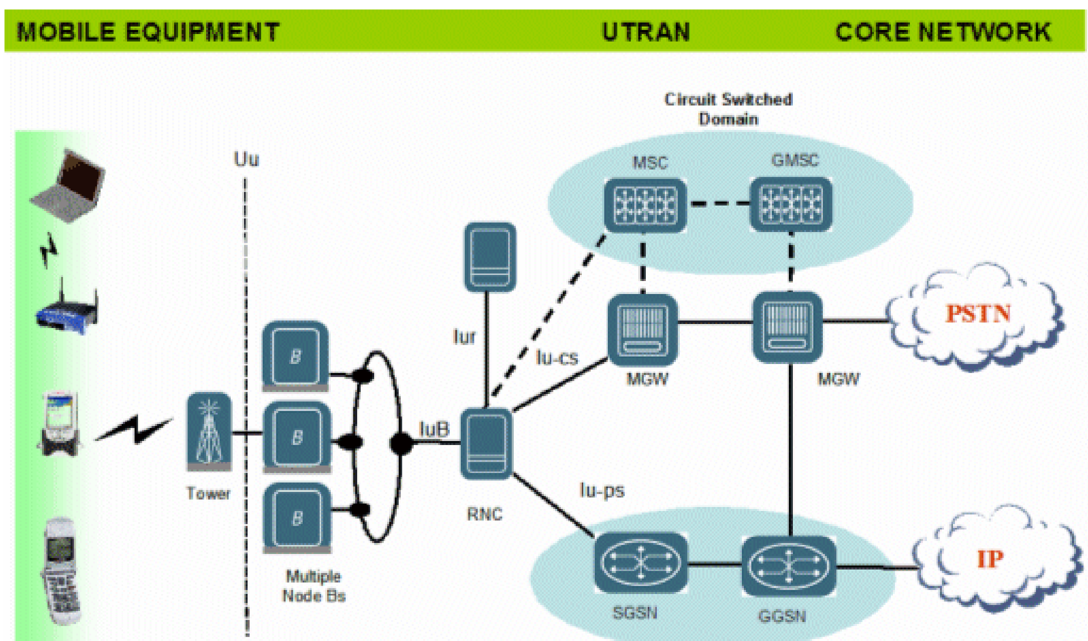

Figure 2-3 UMTS Architecture

The basic Core Network architecture for UMTS is based on GSM network with GPRS. All equipment has to be modified for UMTS operation and services. The UTRAN provides the air interface access method for User Equipment. Base Station is referred as Node-B and control equipment for Node-B's is called Radio Network

Controller (RNC). UMTS system page has an example, how UMTS network could be build.

It is necessary for a network to know the approximate location in order to be able to page user equipment. Here is the list of system areas from largest to smallest.

- UMTS systems (including satellite)
- Public Land Mobile Network (PLMN)
- MSC/VLR or SGSN
- Location Area
- Routing Area (PS domain)
- UTRAN Registration Area (PS domain)
- Cell
- Sub cell

2.1.4.2. Core Network

The Core Network is divided in circuit switched and packet switched domains. Some of the circuit switched elements are Mobile services Switching Centre (MSC), Visitor location register (VLR) and Gateway MSC. Packet switched elements are Serving GPRS Support Node (SGSN) and Gateway GPRS Support Node (GGSN). Some network elements, like EIR, HLR, VLR and AUC are shared by both domains.

The Asynchronous Transfer Mode (ATM) is defined for UMTS core transmission. ATM Adaptation Layer type 2 (AAL2) handles circuit switched connection and packet connection protocol AAL5 is designed for data delivery.

2.1.4.3. Radio Access

Wide band CDMA technology was selected to for UTRAN air interface. UMTS WCDMA is a Direct Sequence CDMA system where user data is multiplied with quasi-random bits derived from WCDMA Spreading codes. In UMTS, in addition to channelization, Codes are used for synchronization and scrambling. WCDMA has two basic modes of operation: Frequency Division Duplex (FDD) and Time Division Duplex (TDD). UTRAN interfaces are shown on UMTS Network page.

A summary of Node-B and RNC functions are illustrated in Table 2:

Table 2 Node-B and RNC functions

Node-B	***RNC***
Air interface Transmission / Reception	Radio Resource Control
Modulation / Demodulation	Admission Control
CDMA Physical Channel coding	Channel Allocation
Micro Diversity	Power Control Settings
Error Handing	Handover Control
Closed loop power control	Macro Diversity
	Ciphering
	Segmentation / Reassembly
	Broadcast Signaling
	Open Loop Power Control

2.1.4.4. User Equipment

The UMTS standard does not restrict the functionality of the User Equipment in any way. Terminals work as an air interface counterpart for Node-B and have many different types of identities. Most of these UMTS identity types are taken directly from GSM specifications.

- International Mobile Subscriber Identity (IMSI)
- Temporary Mobile Subscriber Identity (TMSI)
- Packet Temporary Mobile Subscriber Identity (P-TMSI)
- Temporary Logical Link Identity (TLLI)
- Mobile station ISDN (MSISDN)
- International Mobile Station Equipment Identity (IMEI)
- International Mobile Station Equipment Identity and Software Number (IMEISV)

UMTS mobile station can operate in one of three modes of operation:

- PS/CS mode of operation: The MS is attached to both the PS domain and CS domain, and the MS is capable of simultaneously operating PS services and CS services.
- PS mode of operation: The MS is attached to the PS domain only and may only operate services of the PS domain. However, this does not prevent CS-like services to be offered over the PS domain (like VoIP).
- CS mode of operation: The MS is attached to the CS domain only and may only operate services of the CS domain.

UMTS IC card has same physical characteristics as GSM SIM card. It has several functions:

- Support of one User Service Identity Module (USIM) application (optionally more than one)
- Support of one or more user profile on the USIM
- Update USIM specific information over the air
- Security functions
- User authentication
- Optional inclusion of payment methods
- Optional secure downloading of new applications

2.1.5. *Beyond 3G (4G and NGN)*

With operators worldwide starting to deploy High Speed Packet Access (HSPA) services, 3G has finally arrived and is now gaining momentum. As 3G enjoys its success, the evolution in the radio and core network space continues with Long Term Evolution (LTE) and Evolved Packet System (EPS) leading the way forward to the 4G revolution. The heralded "anytime, anywhere" access and entertainment medium shall become a reality with 4G.

2.1.5.1. *The Shift from 3G to 4G*

In today's market, the subscriber / end user is at the center of everything, and the services are driven by user demands. Initially more focused on short message

service (SMS), multimedia message service (MMS), and content downloading, the 3G market has moved rapidly toward video sharing, mobile video, and IPTV, all of which require very high data throughput and highly efficient radio.

The latest UMTS evolution brings improved spectral efficiency at lower latency and higher data speeds with almost 100 times improvement from 3G - the new promise known as 4G. To achieve this, the air interface is making use of new modulation techniques which effectively target the issues seen with 3G, such as symbol efficiency. The radio interface has adopted some best-in-class methods for the air interface, namely Orthogonal Frequency Division Multiplexing (OFDM) and Multiple-Input Multiple-Output (MIMO) systems as shown.

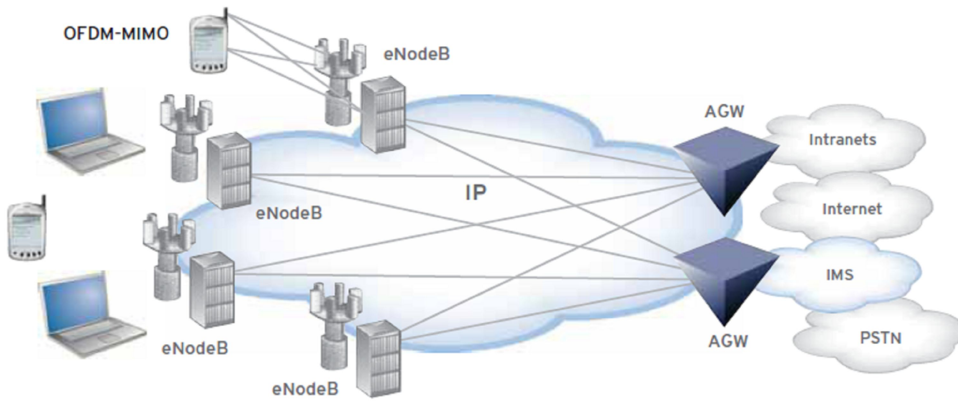

Figure 2-4 The 4G Network Architecture

To support the evolved UMTS Terrestrial Radio Access Network (UTRAN) with OFDM at radio access, the core has adopted a flatter IP-based architecture that gives better data performance. The 4G core shall support an open framework, allowing any kind of mobility protocol, quality of service (QoS), and Authentication, Authorization, and Accounting (AAA) services with support for multiple access technologies, and be able to provide value-added services which are personalized and context-aware.

The truth is, unlike 3G, which is clearly defined in IMT-2000 (International Mobile Telecommunications 2000), no standard body has yet to offer a clear definition of what comprises 4G. IMT-Advanced is the closest thing we could find

that offers a glimpse of the requirement of a 4G network. IMT-Advanced is defined as the following:

"Systems beyond IMT-2000, for which there may be a need for a new wireless access technology to be developed around the year 2010, capable of supporting high data rates with high mobility. High mobility here covers high speed on highways or fast trains (60 kilometers/hour to 250 km/h, or more). It is predicted that potential new radio interfaces will need to support data rates of up to approximately 100 Mbps (megabits per second) for high mobility such as mobile access and up to approximately 1 Gbps (gigabit per second) for low mobility such as nomadic/local wireless access, by around the year 2010."

Fourth-generation (4G) mobile systems dictate entirely new approaches and novel infrastructure solutions to seamlessly integrate the existing wireless technologies including wireless broadband (WiBro), 802.16e, CDMA, wireless LAN, Bluetooth, and etc.

2.1.5.2. NGN definition (ITU)

An NGN is a packet-based network able to provide telecommunication services and able to make use of multiple broadband, QoS-enabled transport technologies and in which service-related functions are independent from underlying transport-related technologies. It enables unfettered access for users to networks and to competing service providers and/or services of their choice. It supports generalized mobility which will allow consistent and ubiquitous provision of services to users.

2.1.5.3. Features of B3G systems

The first visions of B3G or 4G estimate this next generation to be realized around 2010 and based around five elements: fully converged services, ubiquitous mobile access, diverse user devices, autonomous networks and software dependency. In this vision, the seamless connection of heterogeneous networks include cellular data, 3G, WLAN/HIPERLAN, short range PAN/LAN/MAN and

broadcast DVB/DAB services integrated through an IP based core network. Implementing this type of integrated system invokes many challenges in mobile handset design, wireless system discovery, terminal mobility, topological fault tolerance and survivability. Realization of the envisioned B3G or 4G systems and technologies requires unified efforts in the areas of standardization and development of the enabling technologies, providing access network convergence in an evolutionary manner, designing system architectures with considerations on application, protocols, location, authentication and security. The fragmented nature of the heterogeneous radio spectrum calls for new software defined reconfigurable radio concepts utilizing economical smart and adaptive antennas in MIMO radio channels, multi-standard (3G/4G/802.11/GPS/Bluetooth) transceiver architectures, wideband OFDM/multi-carrier modulation, and other advanced solutions implemented in RFIC. They will provide new performance in achieving data rates up to 100 Mbps and support for streaming, multicasting, and downloads of 5-20Mbps even when traveling at 200 km/hr, over an IP-centric network. Terminal implementation technologies include system-on-chip (SoC) and system-in-a-package (SiP) technologies with reconfigurable building blocks.

From the service aspect one has to find a balance between public and private service domains by adapting multiple standards and service environments (home, office, outdoors, indoors) across multiple operators and service provider domains with ensured QoS, data privacy and information integrity, and taking into account user profile and terminal characteristics. This requires a service architecture that supports integrated mobility management for heterogeneous wireless network environment, session control and mobility, AAA (Authorization, Authentication and Accounting) functionality, and profile based personalizable service management. In order to achieve end-to-end QoS, the system architecture has to consider the handoff process in particular. During the handoff between heterogeneous systems the mobile user may experience significant variation is QoS due to handoff delay caused by message exchanges, multiple database accesses, and negotiation-renegotiation processes affecting the performance of both upper-layer protocols and applications. This type of problems requires considerations at many layers of the communications:

priority-based routing at the network layer, delay budget calculations and the optimization of individual parameters (e.g., the dwell-timer value in a handoff algorithm) in both locally (for a single user) and system-wide (for all users in a cell or a subnet).

Heterogeneous and fragmented cell architecture enables advanced scenarios for multi-hop and ad hoc type communications, although they present more challenges for QoS guarantees due to the nature of relayed communications. Location-awareness may bring a useful additional parameter for adaptive applications and routing methods.

2.1.5.4. The Path to 4G

Unlike 1G, 2G and 3G, 4G will not be about air interface technologies. On the contrary, 4G will be agnostic to the types of underlying air interfaces. Although, in order to support the targeted throughput as outlined by the ITU (International Telecommunication Union), it is likely to be based on OFDMA (Orthogonal Frequency Division Multiple Access) modulation with MIMO (multiple input, multiple output) and other smart antenna enhancements. 4G is a vision of an all-IP based, heterogeneous mobile broadband network that supports multiple air interfaces, converged fixed-mobile networks, and multiple device form factors and provides consumers with an always-best-connected, low latency, high QoS (quality of service) broadband experience.

Although WiMax supporters claim that 4G is WiMax, we believe this is a wrongful assertion. Future versions of WiMax may become potential 4G candidates, and OFDM (Orthogonal Frequency Division Multiplexing) modulation will be a key component of 4G, but 4G is definitely not WiMax. WiMax has served as a catalyst for 3GPP (Third Generation Partnership Project) and 3GPP2 to accelerate their next round of innovation, adopting OFDM modulation and implementing MIMO and other smart antenna technologies. Both camps have clearly defined their paths toward 4G.

The enhancements in data rate on cellular networks are shown in Figure 2-5.

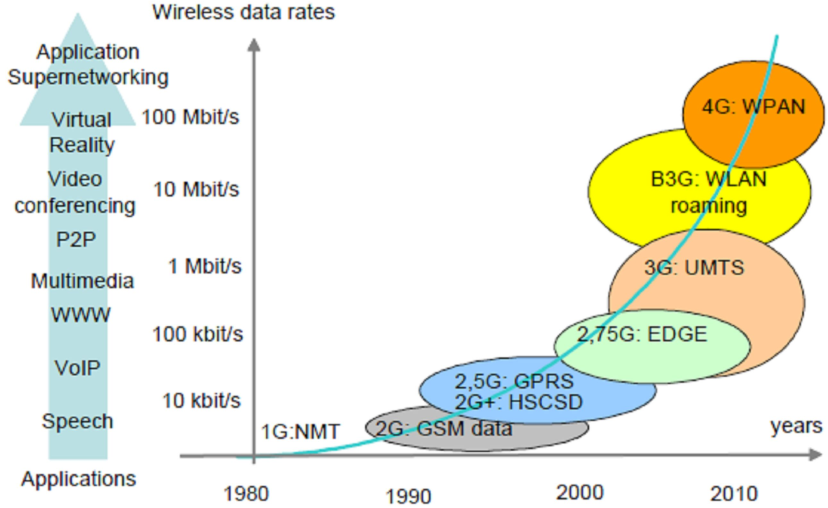

Figure 2-5 Data Rate Enhancement in Cellular Network

2.1.5.5. Components in the 4G Core

The new 4G architecture is evolving with fewer nodes and a flatter structure, thus giving lower latency. That brings in for a requirement of an all-IP-based core network to support the high data throughput and is general enough to be accessed by the different radio access networks through gateway interfaces.

As a result, 4G is not so much about all new standards, but is instead based on existing technologies (e.g., WLAN, 2G, 2.5G, 3G, and satellite) being used to better advantage. 4G is the evolution beyond 3G which addresses the limitations seen so far while working to enhance the quality of service and increase the bandwidth to make better use of resources.

2.2. WLAN (IEEE 802.11) Overview

In 1997 the IEEE adopted IEEE Std. 802.11-1997, the first wireless LAN (WLAN) standard, which belongs to the IEEE 802 family shown in Figure 2 9. This standard defines the media access control (MAC) and physical (PHY) layer-s for a LAN with wireless connectivity. It addresses local area networking where the connected devices communicate over the air to other devices that are within close proximity to each other. Here overview of the 802.11 architecture and the different topologies incorporated to accommodate the unique characteristics of the IEEE 802.11 wireless LAN standard are provided.

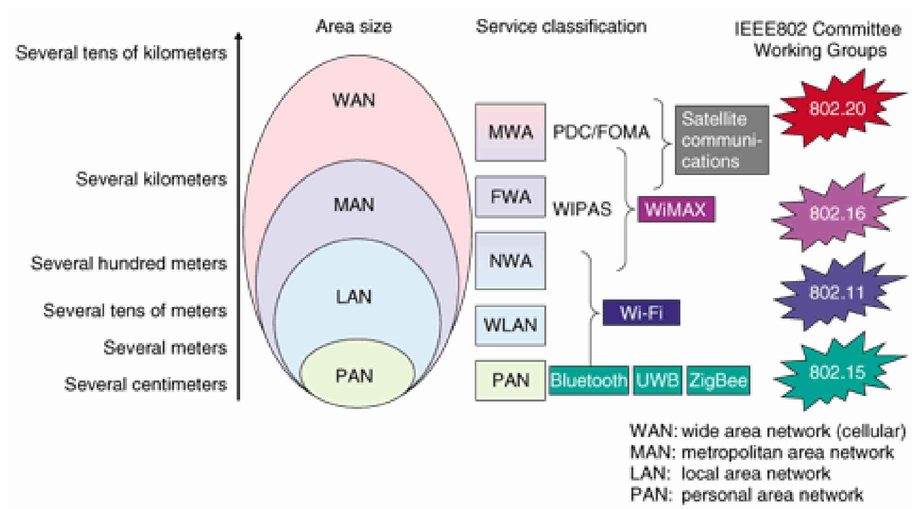

Figure 2-6 IEEE 802 Family

A wireless LAN or WLAN or wireless local area network is the linking of two or more computers or devices using spread-spectrum or OFDM modulation technology based to enable communication between devices in a limited area. This gives users the mobility to move around within a broad coverage area and still be connected to the network.

The 802.11 family includes over-the-air modulation techniques that use the same basic protocol. The most popular are those defined by the 802.11b and 802.11g protocols, and are amendments to the original standard. 802.11-1997 was the first wireless networking standard, but 802.11b was the first widely accepted one, followed by 802.11g and 802.11n. Security was originally purposefully weak due to

export requirements of some governments, and was later enhanced via the 802.11i amendment after governmental and legislative changes. 802.11n is a new multi-streaming modulation technique that is still under draft development, but products based on its proprietary pre-draft versions are being sold. Other standards in the family (c–f, h, j) are service amendments and extensions or corrections to previous specifications.

2.2.1. *Advantages of WLAN:*

The popularity of wireless LANs is a testament primarily to their convenience, cost efficiency, and ease of integration with other networks and network components. The majority of computers sold to consumers today come pre-equipped with all necessary wireless LAN technology. Benefits of wireless LANs include:

- **Convenience:** The wireless nature of such networks allows users to access network resources from nearly any convenient location within their primary networking environment (home or office). With the increasing saturation of laptop-style computers, this is particularly relevant.
- **Mobility:** With the emergence of public wireless networks, users can access the internet even outside their normal work environment. Most chain coffee shops, for example, offer their customers a wireless connection to the internet at little or no cost.
- **Productivity:** Users connected to a wireless network can maintain a nearly constant affiliation with their desired network as they move from place to place. For a business, this implies that an employee can potentially be more productive as his or her work can be accomplished from any convenient location. For example, a hospital or warehouse may implement Voice over WLAN applications that enable mobility and cost savings.
- **Deployment:** Initial setup of an infrastructure-based wireless network requires little more than a single access point. Wired networks, on the other hand, have the additional cost and complexity of actual physical cables being run to

numerous locations; which can even be impossible for hard-to-reach locations within a building.

- **Expandability:** Wireless networks can serve a suddenly-increased number of clients with the existing equipment. In a wired network, additional clients would require additional wiring.

- **Cost:** Wireless networking hardware is at worst a modest increase from wired counterparts. This potentially increased cost is almost always more than outweighed by the savings in cost and labor associated to running physical cables.

2.2.2. Disadvantages

Wireless LAN technology, while replete with the conveniences and advantages described in previous section, has its share of downfalls. For a given networking situation, wireless LANs may not be desirable for a number of reasons. Most of these have to do with the inherent limitations of the technology.

- **Security:** Wireless LAN transceivers are designed to serve computers throughout a structure with uninterrupted service using radio frequencies. Because of space and cost, the antennas typically present on wireless networking cards in the end computers are generally relatively poor. In order to properly receive signals using such limited antennas throughout even a modest area, the wireless LAN transceiver utilizes a fairly considerable amount of power. What this means is that not only can the wireless packets be intercepted by a nearby adversary's poorly-equipped computer, but more importantly, a user willing to spend a small amount of money on a good quality antenna can pick up packets at a remarkable distance; perhaps hundreds of times the radius as the typical user. In fact, there are even computer users dedicated to locating and sometimes even cracking into wireless networks, known as wardrivers. On a wired network, any adversary would first have to overcome the physical limitation of tapping into the actual wires, but this is not an issue with wireless packets. To combat this consideration, wireless networks users usually choose to utilize various encryption technologies available such as Wi-Fi Protected Access (WPA). Some of the older encryption methods, such as WEP are known to have weaknesses that a dedicated adversary can compromise.

o **Range:** The typical range of a common 802.11g network with standard equipment is on the order of tens of meters. While sufficient for a typical home, it will be insufficient in a larger structure. To obtain additional range, repeaters or additional access points will have to be purchased. Costs for these items can add up quickly. Other technologies are in the development phase, however, which feature increased range, hoping to render this disadvantage irrelevant.

o **Reliability:** Like any radio frequency transmission, wireless networking signals are subject to a wide variety of interference, as well as complex propagation effects (such as multipath, or especially in this case Rician fading) that are beyond the control of the network administrator. One of the most insidious problems that can affect the stability and reliability of a wireless LAN is the microwave oven. In the case of typical networks, modulation is achieved by complicated forms of phase-shift keying (PSK) or quadrature amplitude modulation (QAM), making interference and propagation effects all the more disturbing. As a result, important network resources such as servers are rarely connected wirelessly.

o **Speed:** The speed on most wireless networks (typically 1-108 Mbit/s) is reasonably slow compared to the slowest common wired networks (100 Mbit/s up to several Gbit/s). There are also performance issues caused by TCP and its built-in congestion avoidance. For most users, however, this observation is irrelevant since the speed bottleneck is not in the wireless routing but rather in the outside network connectivity itself. For example, the maximum ADSL throughput (usually 8 Mbit/s or less) offered by telecommunications companies to general-purpose customers is already far slower than the slowest wireless network to which it is typically connected. That is to say, in most environments, a wireless network running at its slowest speed is still faster than the internet connection serving it in the first place. However, in specialized environments, higher throughput through a wired network might be necessary. Newer standards such as 802.11n are addressing this limitation and will support peak throughput in the range of 100-200 Mbit/s.

2.2.3. *802.11 Physical Layer (PHY)*

The 802.11 physical layer (PHY) is the interface between the MAC and the wireless media where frames are transmitted and received. The PHY provides three functions. First, the PHY provides an interface to exchange frames with the upper MAC layer for transmission and reception of data. Secondly, the PHY uses signal carrier and spread spectrum modulation to transmit data frames over the media. Thirdly, the PHY provides a carrier sense indication back to the MAC to verify activity on the media.

802.11 provides three different PHY definitions: Both Frequency Hopping Spread Spectrum (FHSS) and Direct Sequence Spread Spectrum (DSSS) support 1 and 2 Mbps data rates. An extension to the 802.11 architecture (802.11a) defines different multiplexing techniques that can achieve data rates up to 54 Mbps. Another extension to the standard (802.11b) defines 11 Mbps and 5.5 Mbps data rates (in addition to the 1 and 2Mbps rates) utilizing an extension to DSSS called High Rate DSSS (HR/DSSS). 802.11b also defines a rate shifting technique where 11 Mbps networks may fall back to 5.5 Mbps, 2 Mbps, or 1 Mps under noisy conditions or to inter-operate with legacy 802.11 PHY layers.

2.2.3.1. *Spread Spectrum*

Spread spectrum is a technique trading bandwidth for reliability. The goal is to use more bandwidth than the system really needs for transmission to reduce the impact of localized interference on the media. Spread spectrum spreads the transmitted bandwidth of the resulting signal, reducing the peak power but keeping total power the same.

2.2.3.2. *Frequency Hopping Spread Spectrum (FHSS)*

Frequency Hopping utilizes a set of narrow channels and "hops" through all of them in a predetermined sequence. For example, the 2.4 GHz frequency band is

divided into 70 channels of 1 MHz each. Every 20 to 400 msec the system "hops" to a new channel following a predetermined cyclic pattern.

The 802.11 Frequency Hopping Spread Spectrum (FHSS) PHY uses the 2.4 GHz radio frequency band, operating with at 1 or 2 Mbps data rate.

2.2.3.3. Direct Sequence Spread Spectrum (DSSS)

The principle of Direct Sequence is to spread a signal on a larger frequency band by multiplexing it with a signature or code to minimize localized interference and background noise. To spread the signal, each bit is modulated by a code. In the receiver, the original signal is recovered by receiving the whole spread channel and demodulating with the same code used by the transmitter. The 802.11 Direct Sequence Spread Spectrum (DSSS) PHY also uses the 2.4 GHz radio frequency band.

3. VERTICAL HANDOVER

During the past decade both telecommunication and Internet technologies have been in a phase of rapid development. Till the beginning of the new millennium the development was mainly technology driven and the real user needs were many times forgotten. Still, the mobile Internet evolution has taken many important steps towards providing better quality wireless data services to a wide audience.

In cellular networks, evolution for the first three generations contributed to growing data rates and enhanced communication capabilities, achieving its current peak only recently in the third generation (3G) mobile networks and handsets. At the same time wireless local area networks have achieved enormous popularity in providing wireless broadband connection in public, enterprise and residential environments. Combining these two wireless technologies has attracted researchers now for about a decade, but there still remain issues to study. The next evolutionary steps after the third generation aim to provide extended mobility with optimized data rates and services. Nomadic users have more flexibility when using multi-service networks that provide services such as seamless connection to the Internet via heterogeneous networks, advanced spatial location and navigation services and true IP based real-time multimedia. One of the key challenges in future network management is end-to-end optimization that takes into account variables such as throughput optimization, routing optimization, delay profiles for heterogeneous wireless environments and also economical profitability. The door for next generation networks and services beyond 3G is opening and is soon ready for entering. These systems are called B3G (beyond 3G) or 4G. They will make heavy use of heterogeneous networking technologies.

3.1. Definition of Handover

Handoff process is one of the most important parts of any wireless network. While handoff process defines the transfer of current mobile active connection

between cells within a single layer, vertical handoff defines this process between different layers.

Handoff (or handover) is an event when a mobile station moves from one wireless cell to another. It can be classified into horizontal (intra-system) and vertical (inter-system) cases. Horizontal handoff means handoff within the same wireless access network technology, and vertical handoff means handoff among heterogeneous wireless access network technologies.

3.2. Need for Handover

In telecommunications there may be different reasons why a handoff (handover) might be conducted:

- When the phone is moving away from the area covered by one cell and entering the area covered by another cell the call is transferred to the second cell in order to avoid call termination when the phone gets outside the range of the first cell;
- When the capacity for connecting new calls of a given cell is used up and an existing or new call from a phone, which is located in an area overlapped by another cell, is transferred to that cell in order to free-up some capacity in the first cell for other users, who can only be connected to that cell;
- In non-CDMA networks when the channel used by the phone becomes interfered by another phone using the same channel in a different cell, the call is transferred to a different channel in the same cell or to a different channel in another cell in order to avoid the interference;
- Again in non-CDMA networks when the user behavior changes, e.g. when a fast-traveling user, connected to a large, umbrella-type of cell, stops then the call may be transferred to a smaller macro cell or even to a micro cell in order to free capacity on the umbrella cell for other fast-traveling users and to reduce the potential interference to other cells or users (this works in reverse too, when a user is detected to be moving faster than a certain threshold, the call can be transferred to a larger umbrella-type of cell in order to minimize the frequency of the handoffs due to this movement);

- In CDMA networks a soft handoff (see further down) may be induced in order to reduce the interference to a smaller neighboring cell due to the "near-far" effect even when the phone still has an excellent connection to its current cell.

The most basic form of handoff (handover) is when a phone call in progress is redirected from its current cell (called source) and its used channel in that cell to a new cell (called target) and a new channel. In terrestrial networks the source and the target cells may be served from two different cell sites or from one and the same cell site (in the latter case the two cells are usually referred to as two sectors on that cell site). Such a handoff, in which the source and the target are different cells (even if they are on the same cell site) is called inter-cell handoff. The purpose of inter-cell handoff is to maintain the call as the subscriber is moving out of the area covered by the source cell and entering the area of the target cell.

A special case is possible, in which the source and the target are one and the same cell and only the used channel is changed during the handoff. Such a handoff, in which the cell is not changed, is called intra-cell handoff. The purpose of intra-cell handoff is to change one channel which may be interfered or fading with a new clearer or less fading channel.

3.3. Types of Handover

In addition to the above classification of inter-cell and intra-cell classification of handoffs, they also can be divided into hard and soft handoffs:

A **hard handoff** is one in which the channel in the source cell is released and only then the channel in the target cell is engaged. Thus the connection to the source is broken before the connection to the target is made -- for this reason such handoffs are also known as break-before-make. Hard handoffs are intended to be instantaneous in order to minimize the disruption to the call. A hard handoff is perceived by network engineers as an event during the call.

A **soft handoff** is one in which the channel in the source cell is retained and used for a while in parallel with the channel in the target cell. In this case the connection to the target is established before the connection to the source is broken, hence this handoff is called make-before-break. The interval, during which the two connections are used in parallel, may be brief or substantial. For this reason the soft handoff is perceived by network engineers as a state of the call, rather than a brief event. A soft handoff may involve using connections to more than two cells, e.g. connections to three, four or more cells can be maintained by one phone at the same time. When a call is in a state of soft handoff the signal of the best of all used channels can be utilized for the call at a given moment or all the signals can be combined to produce a clearer copy of the signal. The latter is more advantageous, and when such combining is performed both in the downlink (forward link) and the uplink (reverse link) the handoff is termed as softer. Softer handoffs are possible when the cells involved in the handoff have a single cell site.

3.4. Definition of Vertical Handover

A new handoff process considered for next generation wireless networks is vertical handoff. The term *vertical* refers to overlapping wireless networks and their hierarchical and asymmetric relationship. The problem of vertical mobility can be illustrated with a simple tale. Let us think of a relay team of a rabbit and a mouse. They have a task to carry a carrot as fast as they can. The mouse cannot carry very heavy loads, but it can go far. The rabbit, in this illustration, is restricted inside a fence. Inside a fence it can carry a heavy load of carrots, but at the fence it has to give the carrots to a mouse. In this thesis the rabbit represents a wireless local area network (WLAN), and the mouse a GPRS network. This work focuses on the event where the rabbit and mouse exchange the carrot load. This event is called *handoff* and in more precise terms *vertical handoff*. The corresponding process is referred as a *transition* from GPRS to WLAN and vice versa. In *mobility management*, the mouse and rabbit communicate with each other about how and when they plan to exchange the carrot, and they may inform others in the community about it too. The carrot represents the *payload* that the mouse (GPRS) and the rabbit (WLAN) are carrying.

The payload means data such as bits or bytes, which are parts of bigger files (such as multimedia files) in a certain format. The problem is to make the best out of a rabbit and a mouse carrying a carrot together. There are physical and practical limits in achieving that, but it is needed to build a systematic approach in the analysis to find out what is possible, what is desirable – and what is not.

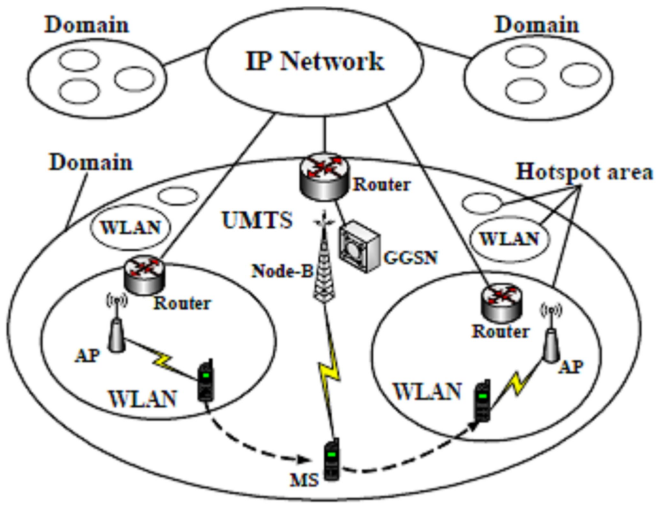

Figure 3-1 Vertical handoff in radio heterogeneous network

The convergence of telecommunication and data networks has been driven by the integration of Internet technologies with wireless and cellular networks. This integration involves both radio carrier switching, network level routing information updating, and transport and application level adaptation. The primary goal has been to enable these processes to appear seamless to the users. In addition, services provided by the operators have to adapt to the vertical mobility scenarios, and in some cases provide additional value through content-aware solutions. Service quality while switching between dissimilar systems has to be managed in a systematic way. These issues are discussed in this chapter with references to the key technical papers in the literature.

3.5. Need for Vertical Handover

The handoff procedure in general includes events such as registration, association, re-association, and dissociation. The vertical handoff procedure involves these same basic events. Figure 3-2 illustrates the vertical handoff procedure.

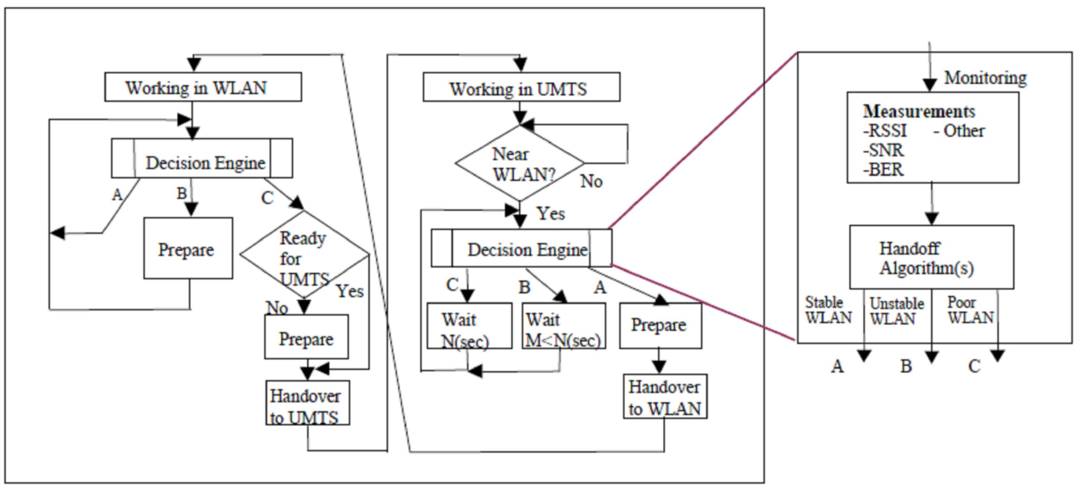

Figure 3-2 Vertical Handoff procedure between WLAN and UMTS

When the MH is associated with UMTS, it monitors at repeated intervals to see whether or not a better high data rate WLAN service is available. Like suggested in Paper I, location information can aid in optimizing this interval so that it is not monitored when the MH is not close to a WLAN hotspot. In the MH there are algorithms that can be implemented as a Handoff Decision Engine (HDE) to provide rules for decision making.

As soon as such a WLAN becomes available, the HDE should initiate an association procedure to the newly discovered AP. Depending on the HDE configuration, the MH may or may not dissociate from UMTS while roaming to a WLAN. In the tight integration model, no dissociation is taking place, only the serving point of network attachment changes. In the loose integration model using Mobile IP, corresponding relations called bindings in the binding caches are updated.

As depicted in Fig. 3.2., the HDE can give one of three possible action outputs: A, B and C. When the MH is associated to a WLAN, the output A (Relax)

indicates that the link quality of the WLAN is satisfactory and there is no imminent need for handoff. Output B (Alert) indicates that that the WLAN signal is weakening and approaching the transition region. HDE allows the application to prepare for the coming handoff by, e.g., momentarily increasing buffer size and potentially inform upper layer protocols such as to tailor the TCP congestion window size or other mechanisms to improve the performance during handoff. Finally, output C (Handoff) indicates that a handoff is needed and the handoff procedure to the overlay network (UMTS) is invoked.

The user should also be able to configure the HDE for certain options such as "use WLAN when available". In an ideal case the HDE would be aware of the application requirements for delay and throughput and can make decisions without user intervention.

3.6. Vertical Handover Algorithm

There are several challenging issues on vertical handoff support. The vertical handoff decision may depend on the bandwidth available for each wireless access network, the ISP (Internet Service Provider) charge for the network connection, the power usage requirements, and the current battery status of the mobile device. The vertical handoff operation should provide authentication of the mobile users, incur a low control overhead, and maintain the connections such that packet loss and transfer delay are minimized.

In general, the vertical handoff process can be divided into three main steps, namely system discovery, handoff decision, and handoff execution. During the system discovery phase, mobile terminals equipped with multiple interfaces have to determine which networks can be used and the services available in each network. The networks may also advertise the supported data rates for different services. During the handoff decision phase, the mobile device determines which network it should connect to. The decision may depend on various parameters including the available bandwidth, delay, jitter, access cost, transmit power, current battery status of the mobile device, and the user's preferences as shown in Figure 3-3. During the

handoff execution phase, connections need to be re-routed from the existing network to the new network in a seamless manner. This phase also includes the authentication and authorization, and the transfer of user's context information.

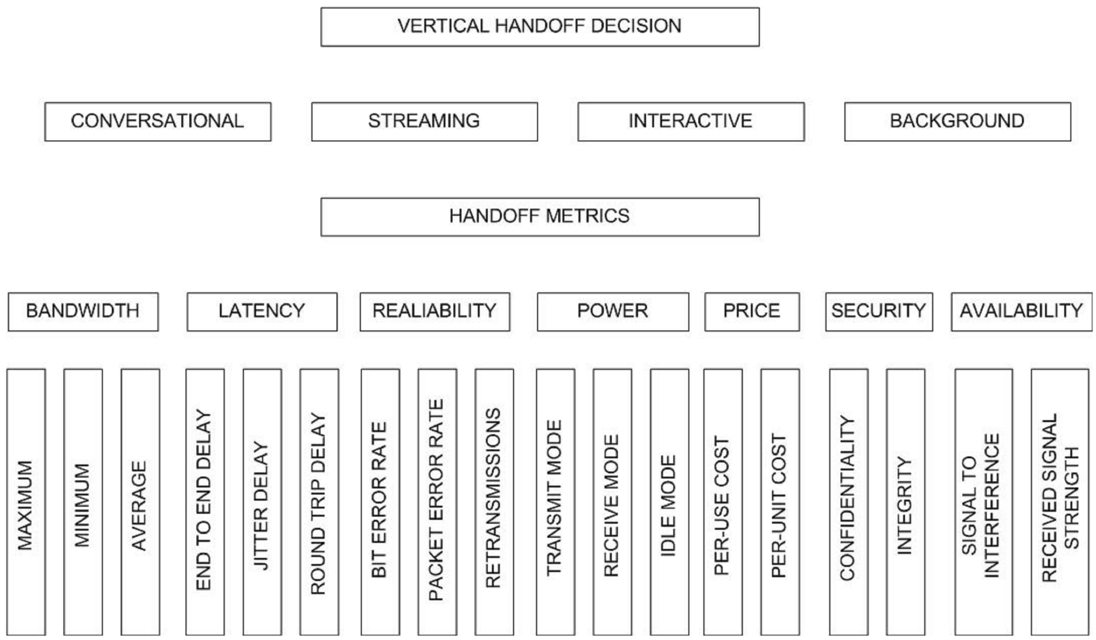

Figure 3-3 Handoff metrics and traffic classes

We choose these inputs because:

1) The feasible bandwidth of UMTS and WLAN are very different. UMTS supports traditional voice service and low bandwidth data-oriented traffics, whereas WLAN applies itself to high-bandwidth and high data-rate traffics. So the traffic type must be considered during the handoff.

2) Velocity is also a very important criterion since the coverage of WLAN is so small that a high-speed MS would quickly pass through it. The frequently passing in and out WLAN cells will bring excessive VHO, which will do great harm to the QoS. So low-speed MS and high-speed MS are suited to be serviced by WLAN and UMTS respectively.

3) The number-of-users has a strong impact on the interference power of UMTS and WLAN. UMTS adopts WCDMA as its radio interface technique, which is a direct sequence CDMA system. The system treats each user separately as a signal, with other users considered as noise or MAI (Multiple Access Interference). So the

capacity of WCDMA is interference limited and the number-of-users exerts important influence on the QoS of UMTS. In WLAN, it is only one channel for users to compete using the ALOHA protocol. Hence, the QoS of WLAN is also number-of-users sensitive.

Though there is other criteria can be chosen, too many inputs will greatly increase the computational complexity. So we only choose the most important criteria. Besides, since the number-of-users in an after-handoff network is time varying it needs to be predicted. In the AMVHO algorithm, the MENN is chosen to do the prediction of the number-of-users and its output will act as an input of the FIS. The FIS makes the analysis of relevant criteria and does the final decision according to the inputs and the if-then rules.

4. ADAPTIVE NEURO-FUZZY INFERENCE SYSTEM (ANFIS)

ANFIS is an adaptive network which permits the usage of neural network topology together with fuzzy logic. It not only includes the characteristics of both methods, but also eliminates some disadvantages of their lonely-used case.

So, in order to understand the concept of ANFIS we must first have a background of Fuzzy Logic System and Adaptive Neural Network Systems as per the first two sections in this chapter, and then, on the third section, have a look at the structure and basic features of ANFIS besides its advantages over FLS and ANN.

4.1. Fuzzy Logic System

The concept of Fuzzy Logic (FL) was conceived by Lotfi Zadeh, a professor at the University of California at Berkley, and presented not as a control methodology, but as a way of processing data by allowing partial set membership rather than crisp set membership or non-membership. This approach to set theory was not applied to control systems until the 70's due to insufficient small-computer capability prior to that time. Professor Zadeh reasoned that people do not require precise, numerical information input, and yet they are capable of highly adaptive control. If feedback controllers could be programmed to accept noisy, imprecise input, they would be much more effective and perhaps easier to implement. Unfortunately, U.S. manufacturers have not been so quick to embrace this technology while the Europeans and Japanese have been aggressively building real products around it.

4.1.1. *Definition of Fuzzy Logic*

In this context, FL is a problem-solving control system methodology that lends itself to implementation in systems ranging from simple, small, embedded micro-controllers to large, networked, multi-channel PC or workstation-based data acquisition and control systems. It can be implemented in hardware, software, or a combination of both. FL provides a simple way to arrive at a definite conclusion based upon vague, ambiguous, imprecise, noisy, or missing input information. FL's

approach to control problems mimics how a person would make decisions, only much faster.

4.1.2. *Fuzzy System Components*

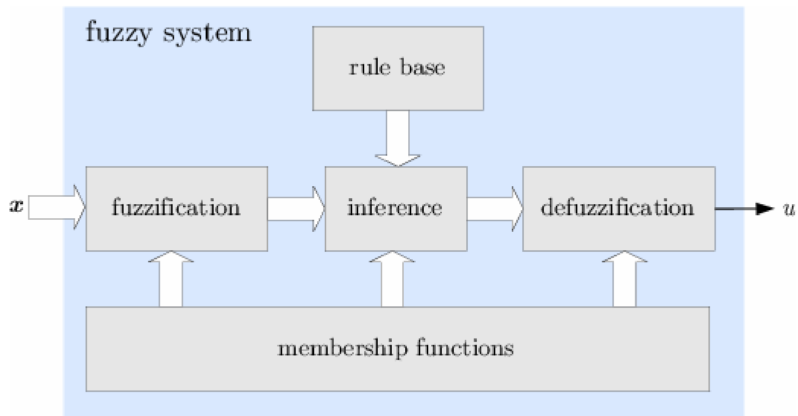

Figure 4-1 Components of a fuzzy system

4.1.2.1. Fuzzification:

The fuzzification comprises the process of transforming crisp values into grades of membership for linguistic terms of fuzzy sets. The membership function is used to associate a grade to each linguistic term.

For the fuzzification of the car speed value $(\chi_0) = 0.75$ km/h the two membership functions μA and μB, which characterize a low and a medium speed fuzzy set, respectively. The given speed value of $(\chi_0) = 0.75$ km/h belongs with a grade of $\mu A(\chi_0) = 0.75$ to the fuzzy set ``low'' and with a grade of $\mu B(\chi_0) = 0.25$ to the fuzzy set ``medium''.

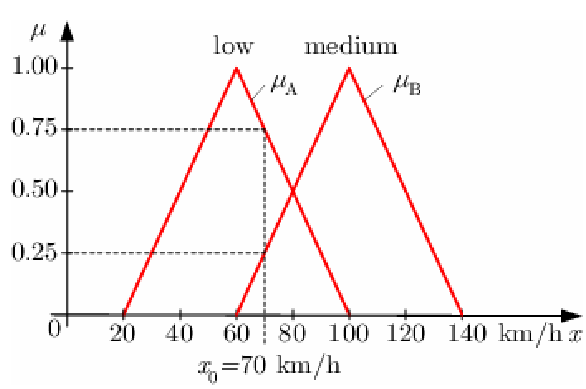

Figure 4-2 Fuzzy System for Car Speed

4.1.2.2. Fuzzy inference machine:

The core section of a fuzzy system is that part, which combines the facts obtained from the fuzzification with the rule base and conducts the fuzzy reasoning process.

A simple fuzzy system is given, which models the brake behavior of a car driver depending on the car speed. The inference machine should determine the brake force for a given car speed. The speed is specified by the two linguistic terms ``low'' and ``medium'', and the brake force by ``moderate'' and ``strong''. The rule base includes the two rules:

(1) IF the car speed is low THEN the brake force is moderate

(2) IF the car speed is medium THEN the brake force is strong

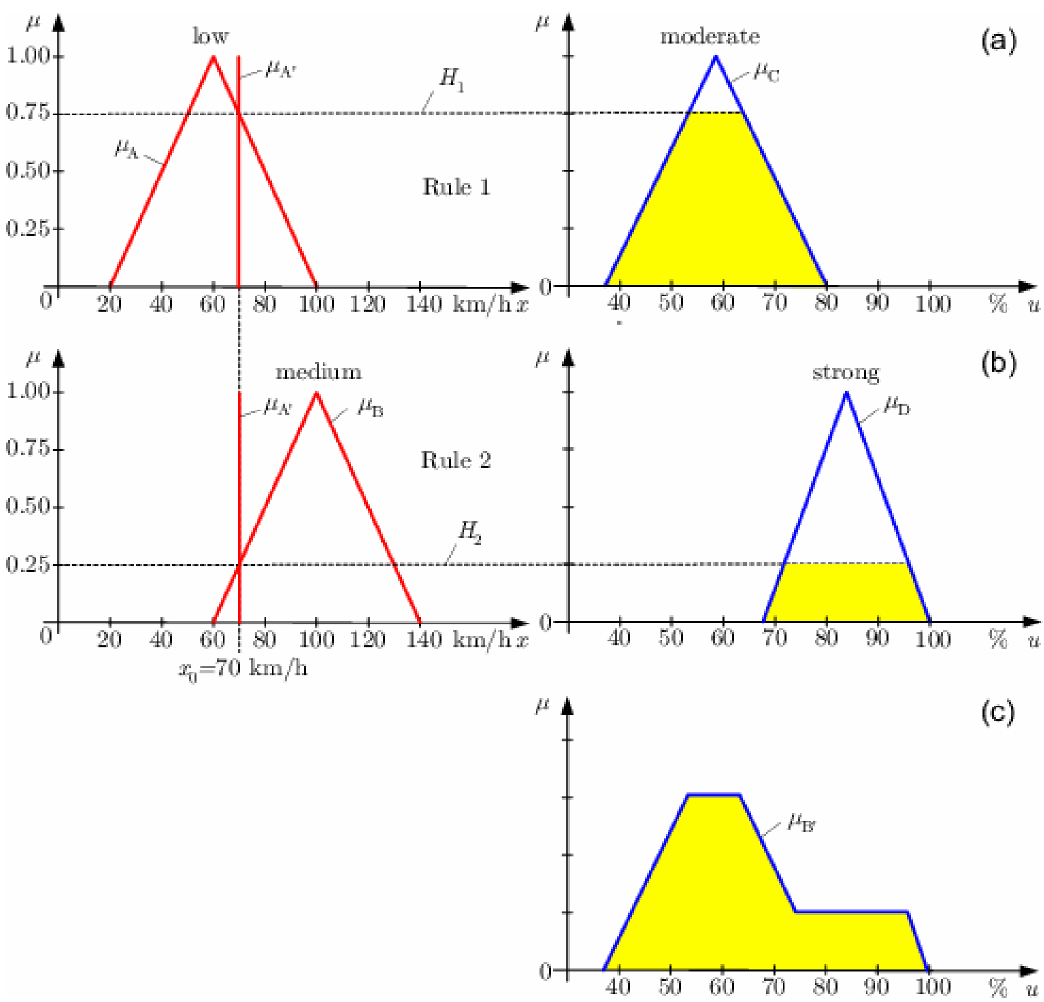

Figure 4-3 Fuzzy Inference example: (a) with rule 1, (b) with rule 2, and (c) final fuzzy set

4.1.2.3. Defuzzification:

As a result of applying the previous steps, one obtains a fuzzy set $\mu_{B'}(u)$ from the reasoning process that describes, for each possible value u, how reasonable it is to use this particular value. In other words, for every possible value u, ne gets a grade of membership that describes to what extent this value u is reasonable to use. Using a fuzzy system as a controller, one wants to transform this fuzzy information into a single value u` that will actually be applied. This transformation from a fuzzy set to a crisp number is called a defuzzification. It is not a unique operation as different approaches are possible.

4.1.3. FL Versus Conventional Control Methods

FL incorporates a simple, rule-based IF X AND Y THEN Z approach to a solving control problem rather than attempting to model a system mathematically. The FL model is empirically-based, relying on an operator's experience rather than their technical understanding of the system. For example, rather than dealing with temperature control in terms such as "SP =500F", "T <1000F", or "210C <TEMP <220C", terms like "IF (process is too cool) AND (process is getting colder) THEN (add heat to the process)" or "IF (process is too hot) AND (process is heating rapidly) THEN (cool the process quickly)" are used. These terms are imprecise and yet very descriptive of what must actually happen. Consider what you do in the shower if the temperature is too cold: you will make the water comfortable very quickly with little trouble. FL is capable of mimicking this type of behavior but at very high rate.

4.1.4. Work Method of FL

FL requires some numerical parameters in order to operate such as what is considered significant error and significant rate-of-change-of-error, but exact values of these numbers are usually not critical unless very responsive performance is required in which case empirical tuning would determine them. For example, a simple temperature control system could use a single temperature feedback sensor whose data is subtracted from the command signal to compute "error" and then time-

differentiated to yield the error slope or rate-of-change-of-error, hereafter called "error-dot". Error might have units of degs F and a small error considered to be 2F while a large error is 5F. The "error-dot" might then have units of degs/min with a small error-dot being 5F/min and a large one being 15F/min. These values don't have to be symmetrical and can be "tweaked" once the system is operating in order to optimize performance. Generally, FL is so forgiving that the system will probably work the first time without any tweaking.

4.1.5. *Need for FL*

FL offers several unique features that make it a particularly good choice for many control problems.

1) It is inherently robust since it does not require precise, noise-free inputs and can be programmed to fail safely if a feedback sensor quits or is destroyed. The output control is a smooth control function despite a wide range of input variations.

2) Since the FL controller processes user-defined rules governing the target control system, it can be modified and tweaked easily to improve or drastically alter system performance. New sensors can easily be incorporated into the system simply by generating appropriate governing rules.

3) FL is not limited to a few feedback inputs and one or two control outputs, nor is it necessary to measure or compute rate-of-change parameters in order for it to be implemented. Any sensor data that provides some indication of a system's actions and reactions is sufficient. This allows the sensors to be inexpensive and imprecise thus keeping the overall system cost and complexity low.

4) Because of the rule-based operation, any reasonable number of inputs can be processed (1-8 or more) and numerous outputs (1-4 or more) generated, although defining the rule-base quickly becomes complex if too many inputs and outputs are chosen for a single implementation since rules defining their interrelations must also be defined. It would be better to break the control system into smaller chunks and use several smaller FL controllers distributed on the system, each with more limited responsibilities.

5) FL can control nonlinear systems that would be difficult or impossible to model mathematically. This opens doors for control systems that would normally be deemed unfeasible for automation.

4.1.6. Uses of FL

There are many wide spread uses of FL could be summarized in the following points:

1) Define the control objectives and criteria: What am I trying to control? What do I have to do to control the system? What kind of response do I need? What are the possible (probable) system failure modes?

2) Determine the input and output relationships and choose a minimum number of variables for input to the FL engine (typically error and rate-of-change-of-error).

3) Using the rule-based structure of FL, break the control problem down into a series of IF X AND Y THEN Z rules that define the desired system output response for given system input conditions. The number and complexity of rules depends on the number of input parameters that are to be processed and the number fuzzy variables associated with each parameter. If possible, use at least one variable and its time derivative. Although it is possible to use a single, instantaneous error parameter without knowing its rate of change, this cripples the system's ability to minimize overshoot for a step inputs.

4) Create FL membership functions that define the meaning (values) of Input/Output terms used in the rules.

5) Create the necessary pre- and post-processing FL routines if implementing in S/W, otherwise program the rules into the FL H/W engine.

6) Test the system, evaluate the results, tune the rules and membership functions, and retest until satisfactory results are obtained.

4.2. Artificial Neural Network (ANN)

An Artificial Neural Network (ANN) is a massively parallel distributed processor made up of simple processing units, which has a natural propensity of storing experiential knowledge and making it available for use. It is composed of a large number of highly interconnected processing elements (neurons) working in unison to solve specific problems. ANNs, like people, learn by example. An ANN is configured for a specific application, such as pattern recognition or data classification, and recently in ATM network, through a learning process. Learning in biological systems involves adjustments to the synaptic connections that exist between the neurons. This is true of ANNs as well.

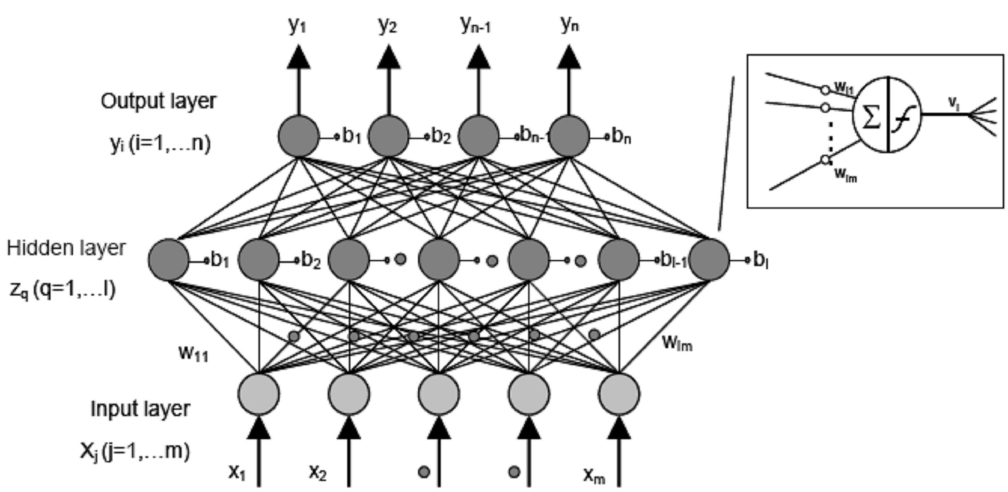

Figure 4-4 Typical Neural Network Architecture

It resembles the brain in two respects:

- Knowledge is acquired by the network from its environment through a learning process.
- Interneuron connection strengths, known as synaptic weights, are used to store the acquired knowledge.

There are many types of artificial neural networks, designed to mimic different capabilities of natural neural networks. Some artificial neural networks are designed

purely to solve particular types of problems and bear little or no resemblance to natural neurons.

4.2.1. *Need for Neural Networks*

Neural networks, with their remarkable ability to derive meaning from complicated or imprecise data, can be used to extract patterns and detect trends that are too complex to be noticed by either humans or other computer techniques. A trained neural network can be thought of as an "expert" in the category of information it has been given to analyses. This expert can then be used to provide projections given new situations of interest and answer "what if" questions.

Other advantages include:

1. Adaptive learning: An ability to learn how to do tasks based on the data given for training or initial experience.
2. Self-Organization: An ANN can create its own organization or representation of the information it receives during learning time.
3. Real Time Operation: ANN computations may be carried out in parallel, and special hardware devices are being designed and manufactured which take advantage of this capability.
4. Fault Tolerance via Redundant Information Coding: Partial destruction of a network leads to the corresponding degradation of performance. However, some network capabilities may be retained even with major network damage.

4.2.2. *Network layers*

The commonest type of artificial neural network consists of three groups, or layers, of units: a layer of "input" units is connected to a layer of "hidden" units, which is connected to a layer of "output" units. (See Figure 4-5)

- The activity of the input units represents the raw information that is fed into the network.

- The activity of each hidden unit is determined by the activities of the input units and the weights on the connections between the input and the hidden units.
- The behavior of the output units depends on the activity of the hidden units and the weights between the hidden and output units.

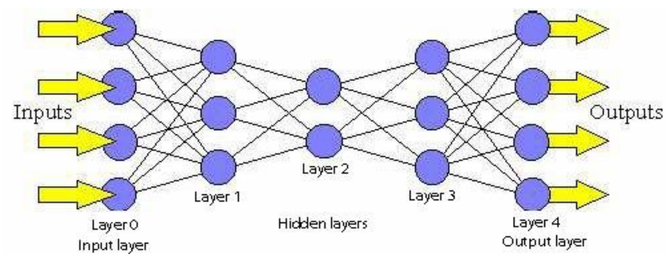

Figure 4-5 An example of a simple feed forward network

4.2.3. *Feed-forward networks*

Feed-forward ANNs (Figure 4-5) allow signals to travel one way only; from input to output. There is no feedback (loops) i.e. the output of any layer does not affect that same layer. Feed-forward ANNs tend to be straight forward networks that associate inputs with outputs. They are extensively used in pattern recognition. This type of organization is also referred to as bottom-up or top-down.

4.2.4. *Feedback networks*

Feedback networks can have signals traveling in both directions by introducing loops in the network. Feedback networks are very powerful and can get extremely complicated. Feedback networks are dynamic; their 'state' is changing continuously until they reach an equilibrium point. They remain at the equilibrium point until the input changes and a new equilibrium needs to be found. Feedback architectures are also referred to as interactive or recurrent, although the latter term is often used to denote feedback connections in single-layer organizations.

4.2.5. Training the network (Back-Propagation Algorithm)

Considering a network with Q feed-forward layers and let ${}^{q}net_i$ and ${}^{q}y_i$ denote the net input and output of the ith unit in the qth layer, respectively. The network has m input nodes and n output nodes. Let ${}^{q}w_{ij}$ denote the connection weight from ${}^{q-1}y_j$ to ${}^{q}y_j$,

A set of training pairs $x^{(k)}, d^{(k)} \mid k = 1, 2, ..., p$ is used as basic input, where the input vectors are augmented with the last elements as -1, i.e. $x^k_{m+1} = -1$. the learning is achieved in following phases:

Phase 1 (initialization): choose η>0 and $E_{\max}$. Initialize the weights to small random values.
Phase 2 (training): Apply the kth input pattern to the input layer

$${}^{q}y_i = {}^{1}y_i = x_i^k \text{ for all } i.$$

Phase 3 (forward propagation): Propagate the signal forward through the network by

using $${}^{q}y_i = a({}^{q}net_i) = a\left(\sum_j {}^{q}w_{ij}^{q-1} y_j\right)$$

Phase 4 (Error measure): compute the error value and error signals ${}^{Q}\delta_i$ for the output layer:

$$E = \frac{1}{2}\sum_{i=1}^{n} \left(d_i^k - {}^{Q}y_i\right)^2 + E$$

$${}^{Q}\delta_i = (d_i^k - {}^{Q}y_i)a'({}^{Q}net_i)$$

Phase 5 (Error back-propagation): Propagation the errors backward to update the weights and compute the error signals ${}^{q-1}\delta_i$ for the preceding layers:

$$\Delta {}^{q}w_{ij} = \eta {}^{q}\delta_i^{q-1} y_j \text{ and } {}^{q}w_{ij}^{new} = \Delta {}^{q}w_{ij}^{old} + \Delta {}^{q}w_{ij}$$

$${}^{q-1}\delta_i = a'({}^{q-1}net_i)\sum_j {}^{q}w_{ij}\, {}^{q}\delta_j \quad \text{for q = Q, Q-1, ...,2.}$$

Phase 6 (One epoch looping): Check whether the whole set of training data has been cycled once. If k<p, then $k = k+1$ and go to phase 1; otherwise, go to phase 6.

Phase 7 (Total error checking): Check whether the current total error is acceptable: if $E < E_{\max}$, then terminate the training process and output the final weights; otherwise, E=0, k=1, and initiate the new training epoch by going to phase 1.

4.3. ANFIS

A fuzzy inference system (FIS) can use human expertise by storing its essential components in the rule base and the database and can perform fuzzy reasoning to infer the overall output value. The derivation of if – then rules and corresponding membership functions (MFs) depends heavily on the a priori knowledge about the system under consideration. However, there is no systematic way to transform experiences of knowledge of human experts into the knowledge base of an FIS. There is also a need for adaptability or some learning algorithms to produce outputs within the required error rate. On the other hand, ANN learning mechanism does not rely on human expertise. Due to the homogenous structure of ANN, it is hard to extract structured knowledge from either the weights or the configuration of the artificial neural network (ANN). The weights of the ANN represent the coefficients of the hyperplane that partition the input space into two regions with different output values. If we can visualize this hyperplane structure from the training data, then the subsequent learning procedures in an ANN can be reduced. However, in reality, the a priori knowledge is usually obtained from human experts; it is most appropriate to express the knowledge as a set of fuzzy if – then rules, and it is not possible to encode into an ANN 0. Table 3 summarizes the comparison of FIS and ANN.

Table 3 ANN Vresus FIS

ANN	FIS
Black box	Interpretable
Learning from scratch	Making use of linguistic knowledge

As we have already seen, fuzzy systems present particular problems to a developer:

- Rules. The if-then rules have to be determined somehow. This is usually done by 'knowledge acquisition' from an expert. It is a time consuming process that is fraught with problems.
- Membership functions. A fuzzy set is fully determined by its membership function. This has to be determined.

The ANFIS approach learns the rules and membership functions from data.

ANFIS is an *adaptive network*. An adaptive network is network of nodes and directional links. Associated with the network is a learning rule - for example back propagation. It's called adaptive because some, or all, of the nodes have parameters which affect the output of the node. These networks are learning a relationship between inputs and outputs.

Adaptive networks cover a number of different approaches but for our purposes we will investigate in some detail the method proposed by Jang known as ANFIS.

4.3.1. ANFIS architecture

The ANFIS architecture is shown below. The circular nodes represent nodes that are fixed whereas the square nodes are nodes that have parameters to be learnt.

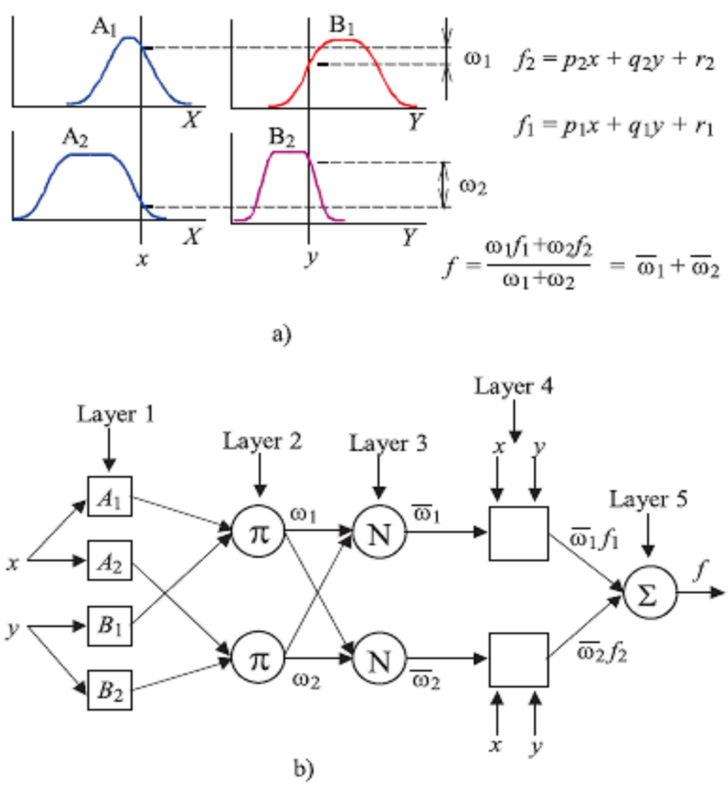

Figure 4-6 An ANFIS architecture for a two rule Sugeno system

(a) – Two-input, one-output, two-rule Sugeno model, (b) – equivalent ANFIS structure

A Two Rule Sugeno ANFIS has rules of the form:

$$If \ \ x \ is \ A_1 \ and \ \ y \ is \ B_1 \quad THEN \ \ f_1 = p_1 x + q_1 y + r_1$$

$$If \ \ x \ is \ A_2 \ and \ \ y \ is \ B_2 \quad THEN \ \ f_2 = p_2 x + q_2 y + r_2$$

For the training of the network, there is a forward pass and a backward pass. We now look at each layer in turn for the forward pass. The forward pass propagates the input vector through the network layer by layer. In the backward pass, the error is sent back through the network in a similar manner to back-propagation.

$O_{l.i}$ is the output of the i[th] node of the layer *l, and* μ_x is the membership function of the input x.

Layer 1

The output of each node is:

$$O_{1,i} = \mu_{A_i}(x) \text{ for } i = 1, 2, \text{ or}$$
$$O_{1,i} = \mu_{B_{i-2}}(x) \text{ for } i = 3, 4$$

So, the $O_{1,i}(x)$ is essentially the membership grade for x and y.

The membership functions (μ_x) could be anything; but for illustration purposes we will use the bell shaped function given by:

$$\mu_A(x) = \frac{1}{1 + \left| \frac{x - c_i}{a_i} \right|^{2b_i}}$$

Where a_i, b_i, c_i are parameters to be learnt. These are the premise parameters.

Layer 2

Every node in this layer is fixed. This is where the t-norm is used to 'AND' the membership grades - for example the product:

$$O_{2,i} = w_i = \mu_{A_i}(x)\mu_{B_i}(y), \quad i = 1,2$$

Layer 3

Layer 3 contains fixed nodes which calculate the ratio of the firing strengths of the rules:

$$O_{3,i} = \overline{w}_i = \frac{w_i}{w_1+w_2}, \quad i = 1, 2$$

Layer 4

The nodes in this layer are adaptive and perform the consequent of the rules:

$$O_{4,i} = \overline{w_i} f_i = \overline{w_i}(p_i x + q_i y + r_i)$$

The parameters in this layer (p_i, q_i, r_i) are to be determined and are referred to as the consequent parameters.

Layer 5

There is a single node here that computes the overall output:

$$overall\ output = O_{5,1} = \sum_i \overline{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i}$$

4.3.2. Model Learning and Inference through ANFIS

The basic idea behind these neuro-adaptive learning techniques is very simple. These techniques provide a method for the fuzzy modelling procedure to learn information about a data set, in order to compute the membership function parameters that best allow the associated fuzzy inference system to track the given input/output data. This learning method works similarly to that of neural networks.

The major training routine for Sugeno-type fuzzy inference systems uses a hybrid learning algorithm to identify parameters of Sugeno-type fuzzy inference systems. It applies a combination of the least-squares method and the back-propagation gradient descent method for training FIS membership function parameters to emulate a given training data set.

5. THE MODEL DESIGN

In this chapter the main idea of the proposed model will be illustrated with its basic elements, and then describe the structure and principles of its operation. After that, a simulation of this model will take place and result analysis carried out to measure the efficiency and performance improvement gained by using the enhancement model.

5.1. Model Specification

As shown in the previous chapter, an essential element in each vertical handover model is the Handoff Decision Engine (HDE). Here the HDE will be implementing by using the **ANFIS** in order to achieve more accurate and adaptable mechanism uses intelligent calculations to take the HO decision.

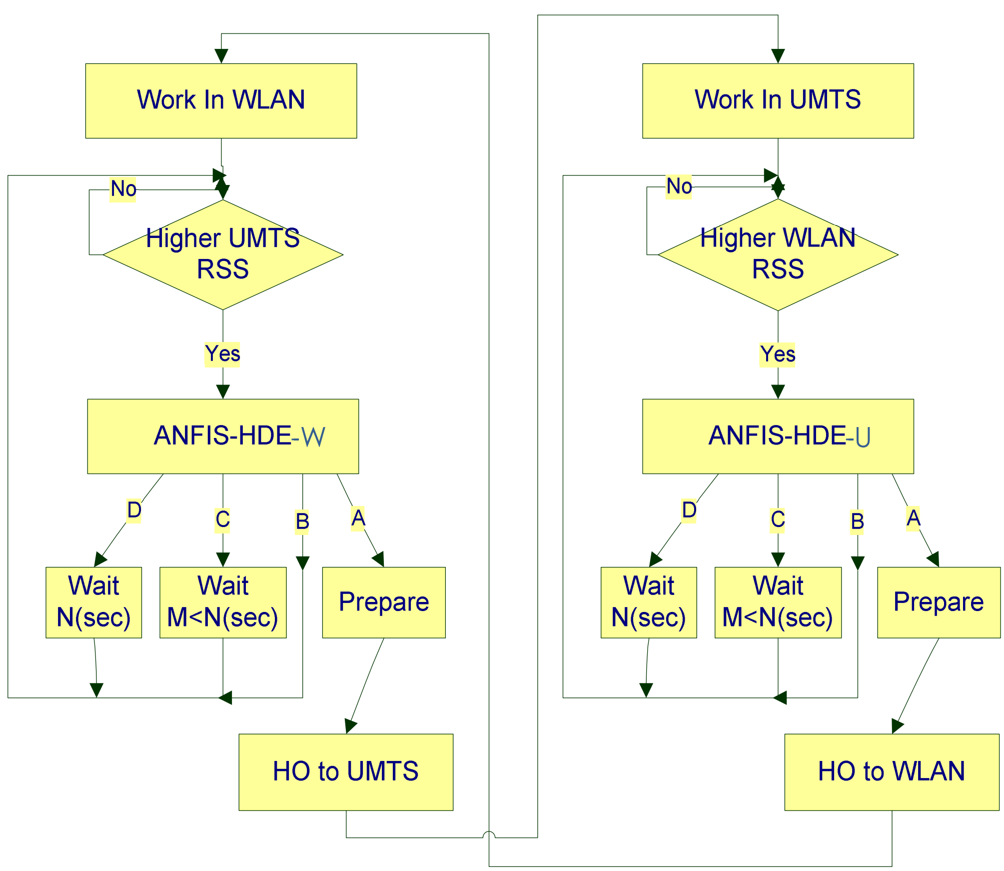

Figure 5-1 Intelligent ANFIS algorithm

The proposed Intelligent ANFIS algorithm, as shown in figure 5-1, contains two HDEs which are **ANFIS-HDE-W** works with WLAN and takes the decision to handoff to UMTS, and the other one **ANFIS-HDE-U** works with UMTS and takes the decision to handoff to WLAN.

The **ANFIS-HDE** is constructed from three parts:

i. **Fuzzification:**

   Convert real values data into fuzzified representations with the help of membership functions.

ii. **Fuzzy Interface:**

   Contains the set of rules and conditions of the inference system, besides, the neural network is trained with the fuzzified information.

iii. **De-Fuzzification:**

   Reconstruct the results to produce real values of the desired output.

5.2. Model Design

Each one of the two ANFIS –HDEs has one **Output** determines the degree by which the handover process to the other network is suitable to take place, i.e. Fit of Not-Fit, as follow:

i. **Output of ANFIS –HDE-W**

W-to-U-HO : represents the degree by which the handover process is fit to occur from UMTS to WLAN network.

ii. **Output of ANFIS –HDE-U**

U-to-W-HO : represents the degree by which the handover process is fit to occur from UMTS to WLAN network

While the main **Inputs** of ANFIS -HDE to be used in my model are:

i. **Received Signal Strength** (**R**ss):

The strength of the signal received, as the RSS of the neighboring network rises above the threshold the Vertical Handoff is feasible i.e. the handoff takes place if and only if RSS of the BS or AP is above the threshold.

ii. Number of Users (UN):

The QoS of WLAN is UN sensitive. As the number of users increase, the collisions increase and results in poor QoS.

iii. Velocity of Mobile Terminal (VMT):

The velocity with which the mobile terminal (MT) is moving. For high speed MT, UMTS is preferred because of greater coverage area.

iv. Traffic Type (TT):

It could be either real time or non-real time. For real time applications i.e. time bound services cellular networks are preferred and for non-real time applications WLAN is preferred.

v. Bandwidth Available (BAV):

It is the amount of unused bandwidth of the candidate Base Station (BS) or Access point (AP). WLAN has greater bandwidth than cellular Network (UMTS).

5.3. Simulation

The parameters shown in Table-4 are used as average estimated values in the simulation.

Table 4 UMTS and WLAN parameters

	UMTS	**WLAN**
Cell Bandwidth	384 Kbps to 2 MBit/s	11 Mbps (IEEE 802.11b)
Cell Diameter	1.5 Km (average)	up to 150 m
Mobile RSS HO threshold	-85 dBm	-80 dBm
Transmitted Power	33 dBm	20 dBm
Maximum acceptable load	Up to 98.0%	Up to 80%

For the simulation process, a Matlab Tool called **"ANFIS Editor"** is used to construct the Model and to implement the FIS parameters. Figure 5-2 shows a snapshot of the model.

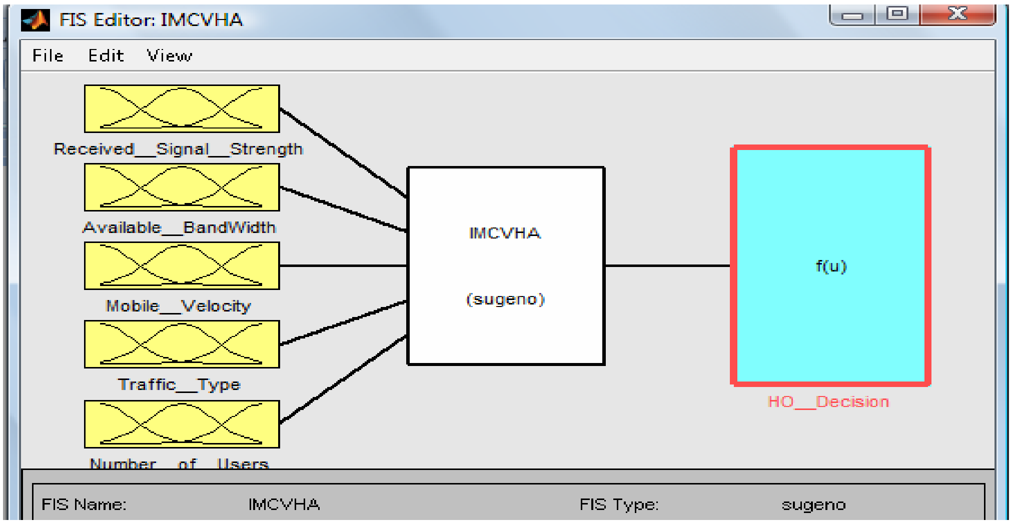

Figure 5-2 FIS Structure

5.3.1. *Fuzzification & Membership-Functions(MF)*

Fuzzification process exists on the fuzzy interface before an interface converts real valued data into a fuzzified representation with the help of membership functions.

At the beginning the training data set that supposed to be observed and acquired from the heterogeneous environment to the ANFIS is loaded, and the number of Inputs and the number of MFs per each input is determined.

The ANFIS then create MFs with default values within the input range for each input. These default values of the MFs are shown on the following figures:

- **R_{SS} = [Below-Thresh, Above- Thresh] = [BT, AT]** , measured in dBm as in figure 5-3.

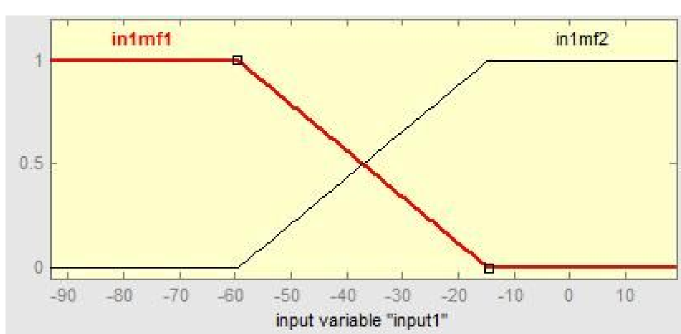

Figure 5-3 Received Signal Strength MF before Training

- U_N = **[Few, Less, Medium, More] = [FE, LE, ME, MO]**, percentage ratio as in figure 5-4.

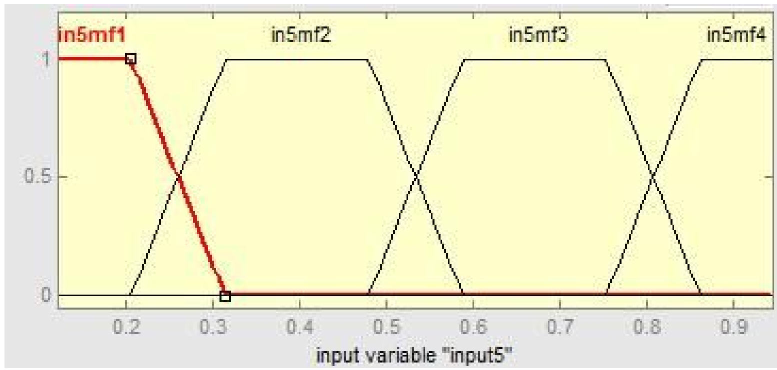

Figure 5-4 Number of User MF before Training

- V_{MT} = **[Very Slow, Slow, Medium, High Speed] = [VS, SL,MS, HS]**, measured in Km/Hr as in figure 5-5.

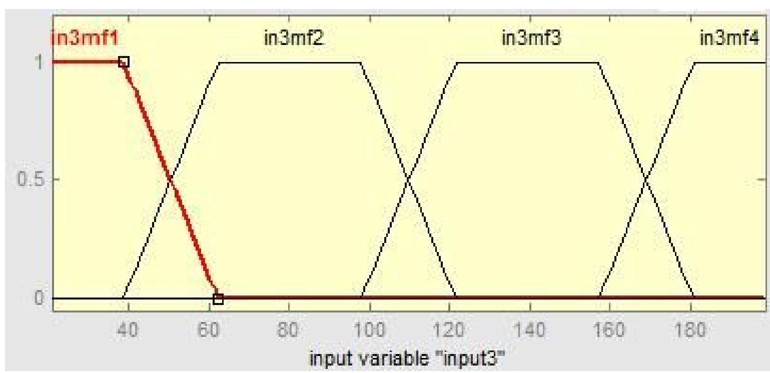

Figure 5-5Mobile Velocity MF before Training

- T_T = **[Non-real Type, Real Type] = [NT, RT]**, percentage ratio as in figure 5-6.

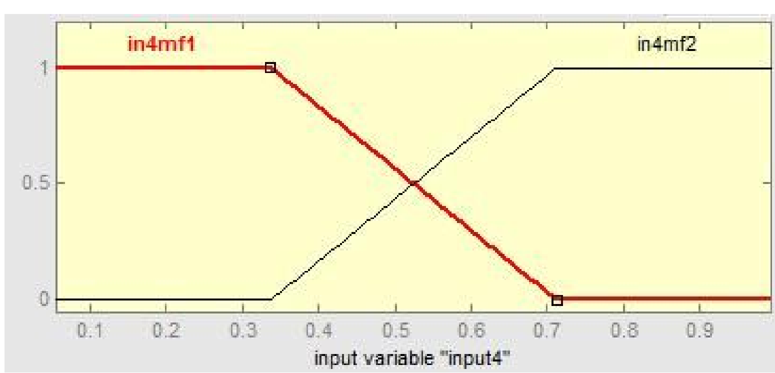

Figure 5-6 Traffic Type MF before Training

- **B_{AV} = [Low, Medium, High, Very High] = [LO, ME, HI, VH]**, measured in MBit/s as in figure 5-7.

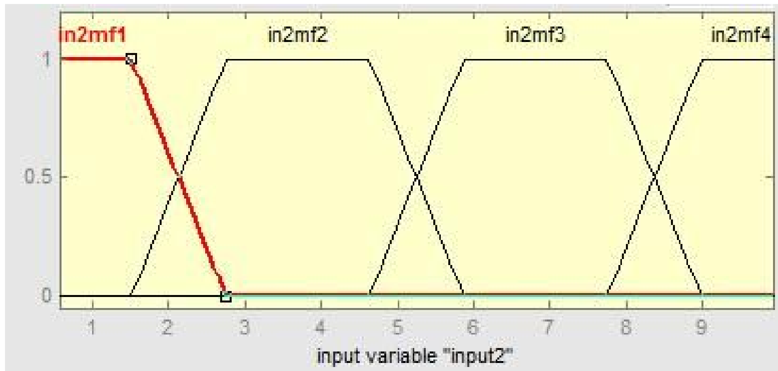

Figure 5-7 Available Bandwidth MF before Training

5.3.2. *Fuzzy Inference System (FIS)*

This level contains the determination of the rules of the FIS that govern the flow of inputs toward output through the ANFIS.

The main rules of the VHO here are:

- RSS must be over threshold to make HO decision
- WLAN is highly affected by the number of users
- Mobile Network is preferred at high speed
- For Non-Real time application WLAN is advantage, and vice versa
- WLAN have greater bandwidth

In the light of these conditions; a set of FIS rules are constructed to manage the HO Decision Output of the HDE depending on the input data. Table.5 shows the main FIS Rules used in our Model.

Table 5. FIS Rules

	B_{AV}	V_{MT}	U_N	T_T	UMTS to WLAN	WLAN to UMTS	Notes
1	LO/ME/HI	VS/SL	FE/LE	NT	FI	MF	This is ideal condition with BAV requirement not too high & only difference is TT=NT is supported by WLAN & TT = RT by UMTS.
2	LO/ME/HI	VS/SL	FE/LE	RT	MF	FI	
3	LO/ME/HI	VS/SL	ME/LR	NT	LF	MF	As UN increases the QoS

4	LO/ME/HI	VS/SL	ME/LR	RT	NF	FI	decreases as no. of users in WLAN share same channel ALOHA.
5	LO/ME/HI	ME/HI	FE/LE	NT	LF	FI	For Medium or High speed HO to WLAN is not supported
6	LO/ME/HI	ME/HI	FE/LE	RT	NF	FI	
7	LO/ME/HI	ME/HI	ME/LR	NT	NF	MF	High speed with large number of user make HO to WLAN totally not supported
8	LO/ME/HI	ME/HI	ME/LR	RT	NF	FI	
9	VH	VS/SL	FE/LE	NT	FI	LF	When High BW required the WLAN is preferred
10	VH	VS/SL	FE/LE	RT	FI	MF	
11	VH	VS/SL	ME/LR	NT	MF	LF	High BW required with high number of users
12	VH	VS/SL	ME/LR	RT	LF	MF	
13	VH	ME/HI	FE/LE	NT	LF	MF	Low fit because of HIGH SPEED and High UN
14	VH	ME/HI	FE/LE	RT	NF	LF	
15	VH	ME/HI	ME/LR	NT	NF	LF	Not fit for WLAN because of HIGH SPEED and High UN
16	VH	ME/HI	ME/LR	RT	NF	MF	

The final structure of the ANFIS model after applying the Membership functions and FIS Rules will be as shown in figure 5-8.

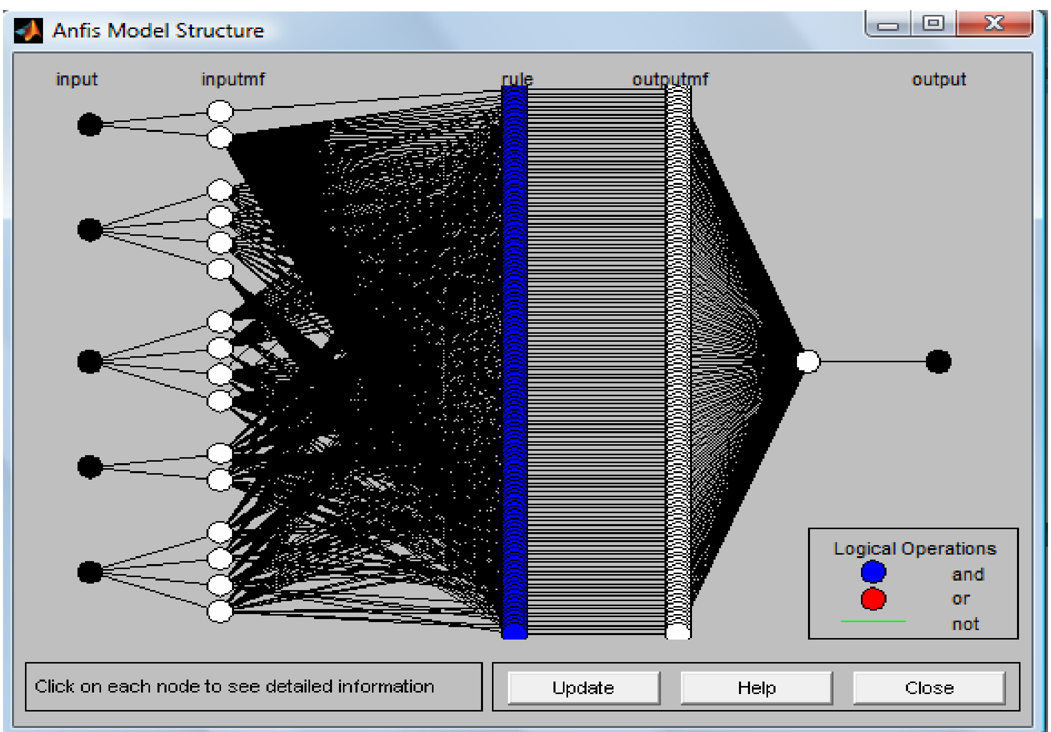

Figure 5-8 ANFIS Model Structure

5.4. Results and Analysis

The training data set is used to train the ANFIS till reach the acceptable value of error already determined and then the training will stop.

Since the data set is the same but with different output value depending on the type of the used and target networks, thus the training process will lead to the same results for both ANFIS-W and ANFIS-U. So the following training process and results represent the two models.

The error value is set to 10^{-10}, and that it is reached after 10 epochs of training process as shown in figure 5-9.

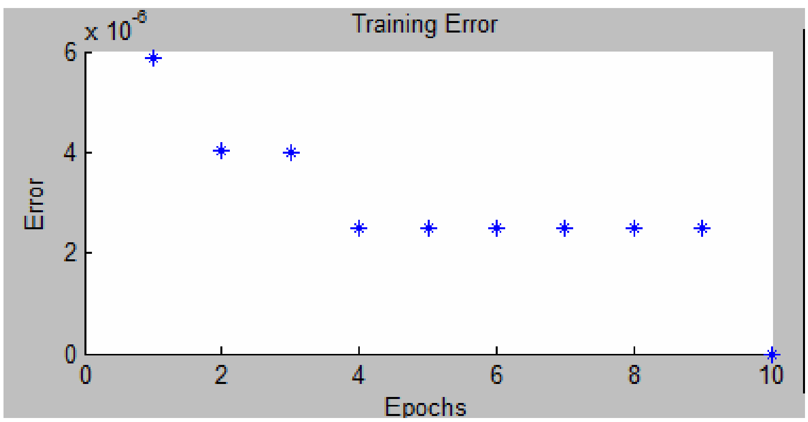

Figure 5-9 Training Result

While the ANFIS info was as follow:

Number of nodes: 552

Number of linear parameters: 1536

Number of nonlinear parameters: 64

Total number of parameters: 1600

Number of training data pairs: 260

Number of checking data pairs: 100

Number of fuzzy rules: 256

After training the ANFIS with the training data set, it comes out with the optimized values of the ANFIS MFs that fit the used heterogeneous environment. The optimal values of MFs of our inputs used in this study are shown on the following figures:

- **R_{SS} = [Below-thresh, Above- thresh] = [BT, AT]** , measured in dBm as in figure 5-3.

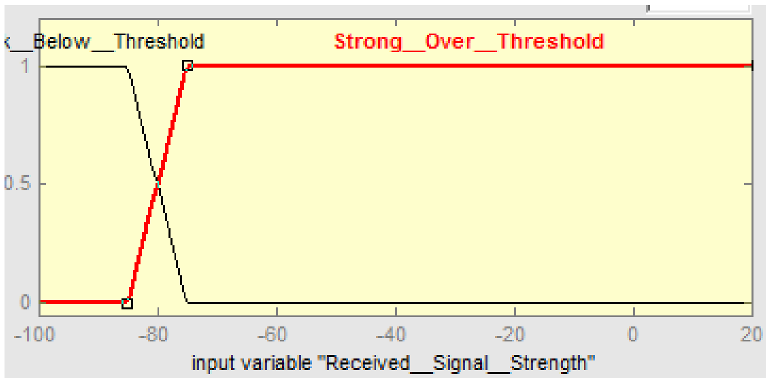

Figure 5-10 Received Signal Strength MF after Training

- **U_N = [Few, less, medium, more] = [FE, LE, ME, MO]**, percentage ratio as in figure 5-4.

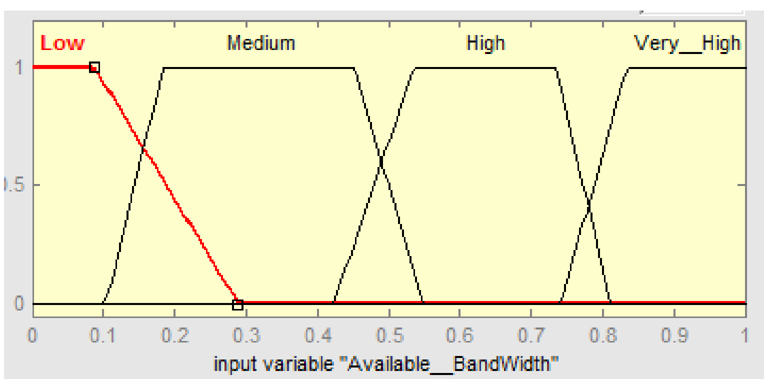

Figure 5-11 Number of User MF after Training

- **V_{MT} = [Very slow, slow, Medium, High Speed] = [VS, SL,MS, HS]**, measured in Km/Hr as in figure 5-5.

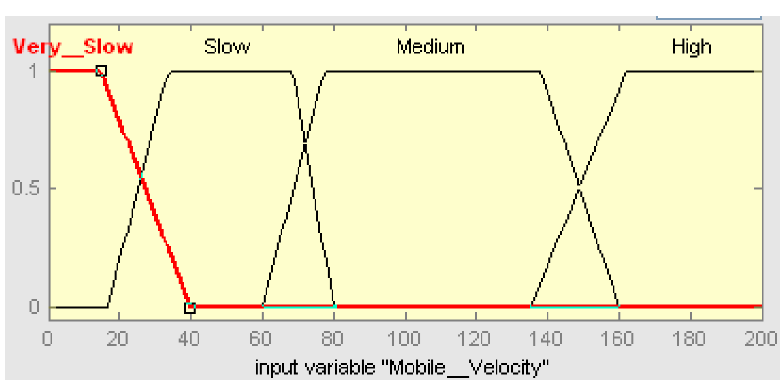

Figure 5-12Mobile Velocity MF after Training

- **T_T = [Non-real type, real type] = [NT, RT]**, percentage ratio as in figure 5-6.

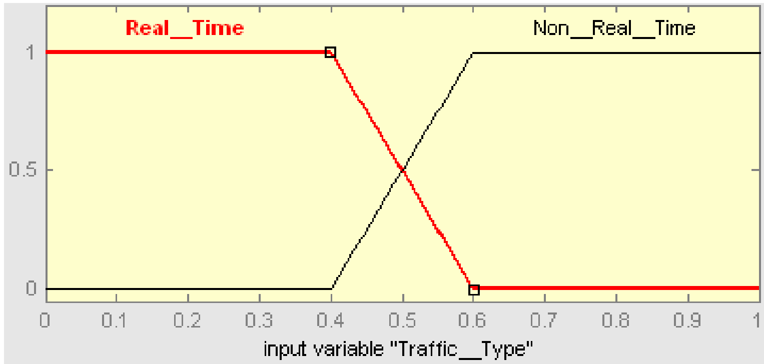

Figure 5-13 Traffic Type MF after Training

- **B_{AV} = [Low, Medium, High, Very High] = [LO, ME, HI, VH]**, percentage ratio as in figure 5-7.

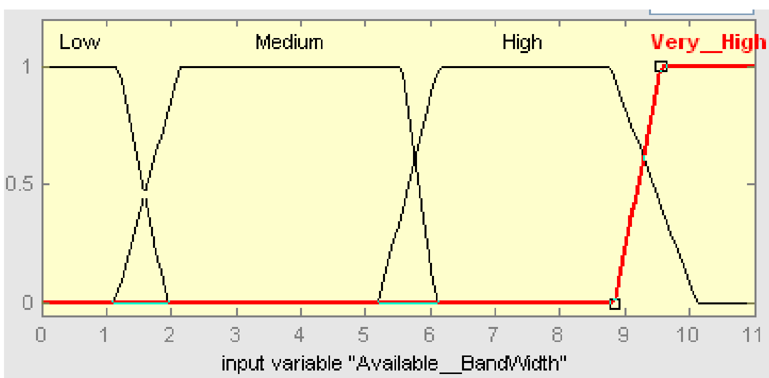

Figure 5-14 Available Bandwidth MF after Training

In order to evaluate our model a date set similar to that used in training process is used. The Matlab ANFIS tool provides the Testing ability by using the testing data set and applies them through the ANFIS layers, the output error should be less than the tolerance error that determined before and used on the training process.

After the optimized MFs come out after training process, they are used by the testing data set with the Test function of the ANFIS. The results show an excellent match with the desired output, and all errors are within the acceptable tolerance as shown in figure 5-15.

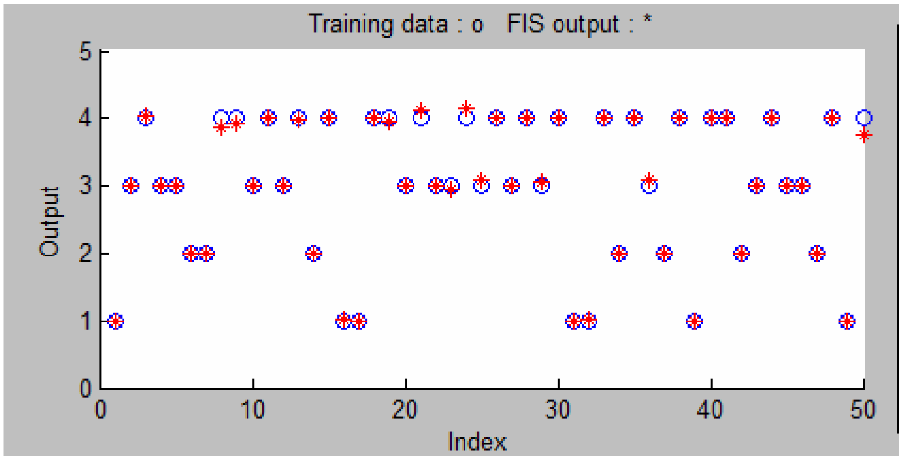

Figure 5-15 Testing Result

Finally, it could be said that the end point in structuring the intelligent model is reached, and the model was trained and tested by the data of a heterogeneous environment, and it proved that it is highly trusted and very efficient especially on the following points:

- Decreasing VHO blocking probability and then Call-Drop ratio
- Decreasing VHO Ping-Pong scenarios
- Achieving highest QoS
- Adaptability to various Networks
- Expandability to include additional parameters

6. CONCLUSION & FUTURE WORK

6.1. Conclusion

Vertical mobility has been a topic of research for over a decade now, and it is deployed in commercial products and field tests. Yet, the popularity of seamless services has not taken its place in the everyday life of consumers in the same way that talking to a mobile phone or using Internet from a home PC. While there is some doubt if vertical mobility will ever have significant enough revenue creating ability for operators, the future challenge is to "put into action" services and applications that utilize vertical roaming with both technical and economical excellence.

This thesis developed an intelligent model using multiple criteria by collecting the required information from the both networks used UMTS and WLAN in the heterogeneous environment beside the mobile status, and analysis them based on the intelligent ANFIS system and then decide the best available network to handoff to in this heterogeneous network environment.

The developed model has shown excellent performance when simulated by Matlab in a heterogeneous network environment. At the beginning, a set of input sequence was generated and used to train the ANFIS system that uses ANN and back-propagation algorithm. After getting the accepted error tolerance from the training it's used to adjust the values of MF's of the FIS system. Finally, the new values and weights are used to construct the intelligent handover decision engine ANFIS-HDE.

Comparing with the conventional HO method which depends only on Received Signal Strength, the developed model indicates better performance and more reliable QoS. Overall advantages could be summarized on the following points:

- Multiple parameters are used to analysis the networks and make the handover decision, that ensures more accuracy and reliability based on variety of the acquired input information.

- By using the Fuzzy Logic and membership functions, the model provide the ability to classify the collected data which enable it to make a set of rules and conditions to deal with each group of classes and create the correspondence handover decision.
- The learning and adaptability features gained by using the ANN has very advantageous effects on setting the suitable values and rules of the model based on specification of the used networks in that heterogeneous network environment.
- As an intelligent model, the ANFIS-HDE has the ability to respond to changes that might happen on the networks and modify the parameters accordingly to maintain the required connectivity and QoS.

Although of all these advantages, our developed model – like any other technology – has some weakness that might slightly affect its performance, and they have been taken in consideration. Those disadvantages are:

- Applying Fuzzy Logic and ANN leads to increase the complexity of the model and then lead to increase the vertical handover delay time.
- If the training process is not well performed or imprecise data is used this will result in inaccurate values of parameters and case performance degradation.
- Due to Matlab Tool "ANFIS Editor" limitations, it requires to design two different ANFIS systems for the two used networks which led to more complexity and delay.

6.2. Suggestions and Future Work

There are many areas of research concerning the vertical handover and NGN. In the same line of our developed model, we can figure out some points for future work and development:

- More parameters could be added to the multi-criteria model and involved on the ANFIS-HDE. And these parameters may include other types of information like: Cost, Mobile Battery, User Preference, etc.
- The training process could achieve more accurate values if a larger number of real data training sequences is used; taking in consideration that values of ANFIS-HDE might be different for different network used.

In another side, the produced model on this thesis – which concerns with the physical layers only – can be used in implementing a complete system of vertical handover in a heterogeneous network environment or NGN alongside with other studies and researches that deal with the other layers of the telecommunication network.

Hybrid networking requires a top-down approach that takes into account novel communication concepts such as application super-networking with new lead application scenarios in the higher layers of the OSI model, and hand-in-hand provisioning tailored wireless and wired network services on the lower OSI layers to meet the application and service requirements. For these aspects the technical solutions presented in this thesis open the door to the future only partially.

REFERENCES

[1]. 3GPP. 3rd Generation Partnership Project, Technical Specification Group Services and System Aspects. http://www.3gpp.org/TB/home.htm.

[2]. Universal Mobile Telecommunication System (UMTS) website, http://www.umtsworld.com/technology/overview.htm .

[3]. J. McNair and F. Zhu, "Vertical Handoffs in Fourth-generation Multi-network Environments," *IEEE Wireless Comm*.

[4]. W. Chen, J. Liu, and H. Huang, "An Adaptive Scheme for Vertical Handoff in Wireless Overlay Networks," in *Proc. of ICPADS'04*.

[5]. MATLAB V.7.0.0 (R14) Help Documents

[6]. Abolfazl Mehbodniya and Jalil Chitizadeh, "An Intelligent Vertical Handoff Algorithm for Next Generation Wireless Networks," *IEEE Wireless Comm*.

[7]. Defining 4G, 3G Americas June 2007, www.3gamericas.org

[8]. Luis B Almeida**, Multilayer perceptrons**;

'http://www.lx.it.pt/~lbalmeida/papers/AlmeidaHNC.pdf'; Mar 2009.

[9]. http://www.intelligraphics.com/articles/80211_article.html

[10]. http://www.technewsworld.com/story/lfWhY0AZUuXQwP/Who-Will-Win-the-4G-Race.xhtml

[11]. http://www.ccpu.com/papers/3Gto4Gmigration/

[12]. http://www.cisco.com/en/US/netsol/ns341/ns396/ns177/ns278/networking_solutions_design_guide_chapter09186a00801219ac.html

[13]. http://en.wikipedia.org