AALBORG UNIVERSITY

Implementation of a self-healing framework

Vikramajeet Khatri

Supervisors:
Venkata Ramkumar
Neeli R. Prasad
Rasmus Hjorth Nielsen

A thesis submitted for the degree of Master of Science at Department of Electronic Systems, Aalborg University.

31 May 2011

Table of Contents

List of Figures

List of Tables

Chapter 1 . Introduction

1.1 Introduction

Radio Communication is a rapidly growing technology, which provides ease of access to different services backed by different technologies and seamless connectivity anywhere anytime. Due to the services offered, the public interest has been high and it has become an integral part of the daily life. In the business concept, running a wireless communication system or network with lower cost and reduced complexity is always preferred over all the other options. Keeping an eye on the current market trend and user demands, self-organizing networks (SONs) are an emerging approach towards a successful business [1].

In this thesis, a self-healing framework has been analyzed, designed and implemented. A self-organizing network aims at self-planning, self-configuration, self-optimization, and self-healing capabilities and improves the overall network performance. Self-healing in simple terms is an automated fault management system, which automatically detects any fault occurring in the network, diagnoses those fault avoiding any service breakdown as well as maintaining Service Level Agreements (SLAs), and reduces costs of the network i.e. capital expenditures (CAPEX) as well as operational expenditures (OPEX) [1].

The thesis is organized as follows:

- Chapter 1 gives the introduction, motivation and problem statement
- Chapter 2 presents Long Term Evolution (LTE) networks: introduction, an overview of network architecture and power control.
- Chapter 3 introduces self-healing systems: introduction, terminology, related work, architecture of framework, thresholding techniques, methods to determining probabilities, and fault modelling.
- Chapter 4 presents the simulation results.
- Chapter 5 gives the conclusions and future work.

1.2 Motivation

Nowadays, users prefer a wireless network that can offer them different kinds of applications in addition to the traditional telephony and messaging services e.g. browsing, VOIP, video streaming, gaming, RSS feeds and updates onto their device.

Considering the need of users to have different applications in the wireless networks, each of the applications has a different threshold for QoS, which needs to be carefully monitored in line with SLAs. In any wireless network, a variety of network elements are responsible for maintaining QoS at the user end. Keeping an eye on the market growth, future 4G wireless networks i.e. Long Term Evolution (LTE) are considered for defining and implementing self-healing mechanisms.

The fault and performance management are the two key processes for a wireless network, which help keep a network running smoothly, satisfy SLAs, generate revenues and gain popularity in the market. In case faults occur either in the software sector or the hardware sector of the network, network services fail, which in turn violate SLAs and the network operator experiences losses in the revenue. Self-healing mechanisms fall into the maintenance category of Optimization and Maintenance Section (O&M) of the network management, as they are automated fault detection and diagnosis algorithms or processes. Therefore, the self-healing plays an important role, and is an emerging area of interest [2].

1.3 Problem Statement

The objective is to find, deal and diagnose any faults that occur in 4G wireless network LTE and implement a self-healing framework. In order to deliver applications satisfying SLAs on the other end in LTE network, the self-healing (automated fault management) reaches a higher level of priority and complexity,

The self-healing task/framework can further be split into three sections as :

- Find the symptoms of a problem from Key Performance Indicator (KPIs)
- Model the fault in accordance to KPIs
- Detect the thresholds of symptoms for each user in the network
- Perform the fault diagnosis i.e. apply corrections to the system

The primary task was to study related work and methods for self-healing mechanism, and then different scenarios for self-healing were considered. The possible scenarios considered for LTE is: call drop scenario. The call drops in a network are subject to a cause of interference and coverage problems.

The different symptoms for the scenarios/faults along with the root causes of the fault were figured out, and the symptom chosen is: SNIR at both the uplink and downlink pathways. The symptom is further mapped to the corresponding root cause, and is updated in the database / codebook. Due to the time limitations, only the symptoms have been detected and the solutions for the root causes/faults i.e. fault diagnosis will be set up in the future.

The fault localization is not considered in this thesis, where as different methods for KPI thresholding and fault diagnosis have been studied, but the implementation has been successful only for one method. It was planned to implement all the methods, and compare the performance analysis for each of the method and then decide which method is better at the performance level of the system. Finally, the algorithm for a specific scenario has been implemented in Matlab, and the performance of the defined algorithm and method has been tested by showing the detection accuracy.

References for Chapter 1

[1]. E3. (2008, December 22). *White paper : Self-x in Radio Access Networks.* (I. G. Eckard Bogenfeld, Ed.) Retrieved March 4, 2010, from End to End Efficiency (E3) FP7 EU Project: https://ict-e3.eu/project/white_papers/Self-x_WhitePaper_Final_v1.0.pdf

[2]. Ramkumar, V. (2011). *Self-healing architecture and methods.* Literature Survey, Networking and Security, Aalborg University, Denmark.

[3]. Moazzam Islam Tiwana, B. S. (2010). Statistical Learning-based Automated Healing: Application to Mobility in 3G LTE Networks. *21st Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications* (ss. 1746-1751). IEEE.

[4]. Moreno, R. B. (2007). Bayesian modelling of fault diagnosis in mobile communication networks. Higher Technical School of Telecommunication Engineering. University of Malaga, Spain.

Chapter 2 . Long Term Evolution (LTE)

2.1 Introduction

Long Term Evolution (LTE) is a new technology, which migrates traditional circuit-based application and services to an all-IP network environment. The Third Generation Partnership Programme (3GPP) initiated LTE project in 2004, aimed at optimizing radio access architecture and enhancement of Universal Terrestrial Radio Access (UTRA) technology, which turned out to a new radio access technology called LTE and is currently deployed. The 3GPP aims at the evolution of 3rd generation (3G) mobile systems, by lowering complexity of systems, decreasing costs, improving data rates and quality of service. LTE offers a variety of services, LTE network design includes the coordination between application servers, devices as well as support and compatibility with the existing technologies. LTE has offered an increase in spectral efficiency up to four times and cells can support 10 times more users than what was supported in preceding technologies such as Global System for Mobile communication (GSM).

Initially, the LTE radio interface was designed to use Orthogonal Frequency Division Multiple Access (OFDMA) scheme in the downlink, and Single Carrier – Frequency Division Multiple Access (SC-FDMA) in the uplink, according to 3GPP specifications. Later on, 3GPP in LTE Release 8 (according to technical documentation 3GPP TS 36.101 V8.13.1) has specified to move the uplink radio access scheme from SC-FDMA to Digital Fourier Transform Spread OFDMA (DFTS-OFDMA) scheme. The major parameters of LTE networks according LTE Release 8 (3GPP TS 36.101 V8.13.1) can be seen in table 2.1 below [1].

Parameter		**Value / Schemes**
Access Scheme	UL	DFTS-OFDM
	DL	OFDMA
Bandwidth		1.4, 3, 5, 10, 15, 20 MHz
Minimum TTI		1 msec
Sub-carrier spacing		15 kHz
Cyclic prefix length	Short	4.7
	Long	16.7
Modulation		QPSK, 16QAM, 64QAM
Spatial multiplexing		Single layer for UL per UE Up to 4 layers for DL per UE MU-MIMO supported for UL and DL

Table 2.1: Different parameters of LTE

The network operators planning to deploy LTE technology also need to consider service mobility issues that allow users to enjoy service mobility and continuity between existing mobile technologies such as GSM, GPRS, UMTS, etc.

2.2 Overview of LTE architecture

While comparing the system architecture of second (2G) and third (3G) generations, the System Architecture Evolution (SAE) standardized by 3GPP, is aimed at helping at the minimization of the total number of nodes as well as an increase in the data efficiency in the network. In order to reduce the inter-node data traffic delays, the evolved architecture has removed and replaced some of the elements in the network architecture, such as the Radio Network Controller (RNC), the Serving GPRS Support Node (SGSN) and the Gateway GPRS Support Node (GGSN) which are all removed and replaced by the SAE Gateway (GW) [2].

In LTE, the evolved Node B (eNB) acts as the Radio Network Controller which is similar to the Base Station in 2G networks. In LTE, the RNC functions have been distributed among eNB and the Mobility Management Entity (MME). The interface X2 is used for the interconnection of eNBs in the network, where as the interface S1 is used for connection between eNBs and MME. The interface S1 can further be classified into two types: S1-MME for connection between eNB and MME which carries the control plane, and S1-U for connection between eNB and Serving Gateway (S-GW) which carries the user plane [3]. The architecture can be seen in figure 2.1 below.

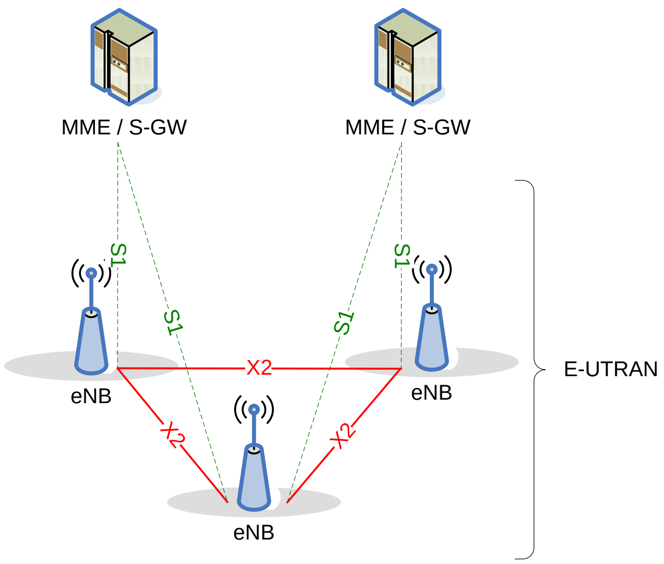

Figure 2.1: Architecture of the LTE Radio Access Network [3]

2.3 Power Control in LTE

Considering the specific consequences of radio access technology in the LTE, the need of power control (PC) can be justified:

- From the system point of view: the need of a reduced impact of inter-cell interference level.
- From the user point of view: achieving a required Signal to Interference Noise Ratio (SINR) level.

The above two requirements are interrelated to each other, as interference is caused by increased level of transmitted power. The PC works as part of the Link Adaption (LA) unit with the purpose of controlling the transmitting power spectral density (PSD) of the users, for hazards reason meeting required signal levels with respect to compensation of channel variations typical of mobile communications systems [4].

Considering the variation of the channel to be compensated, the PC schemes have further been divided into two types: slow and fast power control schemes.

- Slow PC: aimed at compensating for slow channel variations (distance-dependent pathloss, antenna losses, and shadow fading).
- Fast PC: aimed at compensating for fast channel variations (fast fading)

Considering the information transfer mode between user and serving node, PC schemes have further been classified into two types: open loop and closed loop PC schemes.

- Open Loop PC: In this scheme, using parameters and measures obtained from signals sent by the eNB towards the user equipment (UE), the power is adjusted at the UE. Therefore, in this scheme no feedback is sent to the eNB regarding the power used for transmission.
- Closed Loop PC: This scheme is similar to the open loop PC schemes, but in this scheme a feedback is sent to the eNB by UE, which information is then utilized to rectify the user transmit power.

References for Chapter 2

[1]. *3GPP - Long Term Evoluation Technology.* (n.d.). Retrieved April 22, 2011, from 3GPP: http://www.3gpp.org/LTE

[2]. Yilmaz, O. N. (2010). *Self-Optimization of Coverage and Capacity in LTE using Adaptive Antenna Systems.* Department of Communications and Networking. Espoo, Finland: Aalto University.

[3]. (2007). *Technical Specification Group Radio Access Network (3GPP TS 36.300); EUTRA and EUTRAN, Release 8 overall description.* 3GPP.

[4]. Quintero, N. J. (2008). *Advanced Power Control for UTRAN LTE Uplink.* Electronic Systems. Aalborg University, Denmark.

Chapter 3 . Self-healing systems

3.1 Introduction

A self-organizing network aims at self-planning, self-configuration, self-optimization, and self-healing capabilities and improves the overall network performance. In order to meet the user demands, next generation wireless networks will support heterogeneity of different elements in a single network architecture, which has added complexity to the network management. In such heterogeneous networks, if any fault occurs in the network, it will cause losses to the network and will need experts from different areas in order to resolve the issue. Therefore, in order to tackle these problems, self-organizing networks are preferred, and are an emerging area of interest.

Self-healing in simple terms is an automated fault management system, which automatically detects any fault occurring in the network, and/or diagnoses these avoiding any service breakdown. Self-healing systems help a network maintain SLAs, and reduce the costs of the network i.e. CAPEX and OPEX [1]. Self-healing systems quickly diagnose any fault occurring in the network, thus reducing the system downtime as well as the experts from different areas needed to solve the issue i.e. human intervention. The self-healing systems and mechanisms fall into the maintenance category of the network management of wireless networks.

The different challenges that a self-healing system experiences in a network are mentioned below [2]:

- Fault evidence/indication may be
  - ambiguous, same alarm may be generated as an indication of many faults
  - inconsistent, disagreement between NEs, one may think its fault other may not
- Generation of multiple alarms
- Generation of same alarm for multiple faults
- Isolation of multiple simultaneous faults
- Less time to follow each of the triggered alarm, as alarms occur in bursts.
- The characteristics of alarms and alarm sequences may change as the network grows i.e. scalability.
- When self-healing system is implemented for wireless networks, it has more challenges in addition to the above mentioned challenges, which are:
  - Dynamically changing topology
  - Higher fault frequencies, due to high data rates
  - Higher symptom loss rates, due to fading and frequently changing radio conditions
  - Increased number of transient faults
  - Less computing resources available

3.2 Basic definitions

3.2.1 Fault

Fault refers to the defected behaviour or malfunction of any physical or logical network element (hardware or software), which initiates some failures and finally becomes a bottleneck in the overall network performance. In the self-healing mechanism, fault is often referred to as the cause, and is identified from the respective symptoms. Considering the cause of triggering of faults, the faults can be classified into two types: primary and secondary. The primary faults are the ones which are not caused or triggered by any other events or faults in the network; whereas the secondary faults are the ones which are caused or triggered by consequences of any other events or faults occurred in the network. Considering the duration of validity of faults, these are further classified into three types: permanent, intermittent and transient faults. The intermittent faults occur usually at regular intervals for shorter duration of time and are vanished in the notification database; whereas the permanent faults are the ones which occur for longer duration of time, are critical ones, and will not vanish from the notification database unless they are resolved. The transient faults are similar to the concept of volatile memory, these faults occur for shorter duration of time, and will reside in the notification database unless the power is turned off [2].

3.2.2 Alarm

Alarm refers to an event, which has taken place due to occurrence of a fault in the network. An alarm is raised, when any symptom (parameter value) goes below the expected value i.e. threshold, which eventually degrades the overall network performance [3]. It can be raised by every network element, which has either directly or indirectly affected by the fault. An alarm may reside in the notification database unless the fault is rectified. Considering the effect that an alarm has on the overall performance of the network, the alarms can be prioritized into three types: critical, high and low alarms [2].

3.2.3 Symptom

A symptom refers to the performance indictor which value shows up as evidence/proof of a fault, which can further be optimized or improved [4], such as SINR, throughput, etc.

3.2.4 Condition

A condition is a factor which value makes the probability of certain causes occurring either high or low [4]. In simple terms, the condition is the root cause e.g. creating the interference by turning off the PC. In this example, turning PC on and off makes the probability of a user being faced by interference either high or low.

3.3 Fault to Symptom mapping for LTE

In order to apply a self-healing framework on LTE network, a call drop scenario is considered. The reason for calls being dropped may be insufficient radio resources, users may be roaming, users may have experienced interference or coverage issues, there may be some problem in call connection procedure i.e. paging and call admission control (CAC) issues. In this thesis, roaming of users is not considered; instead when users reach the boundary of their respective cell, they bounce back into their respective cell boundaries. The interference and coverage issues are considered as the causes for the call drop scenario in LTE. The SNIR at downlink and uplink has been chosen as a common symptom to analyze these causes, and the condition for these causes is PC.

If PC is on, when any or all of the users face any interference or loss in SNIR, the PC will try to overcome those losses to some extent and users will not have much effect on the SNIR. On the contrary, if the PC is turned off, when any or all of the users face any interference or loss in SNIR, the PC algorithm will not increase the transmitted power and overcome the situation, therefore the SNIR will be affected.

The cause for interference losses is the interference experienced from the users of neighboring cells, when any user reaches closer to the boundary of the respective cell. In the coverage cause, the additional obstacles have been created in the cell boundary, which add a shadowing loss to the path loss that eventually affects the SNIR for a user. The visual diagrams about these causes can be seen in chapter 4.

The relationship between symptom and causes for the call drop scenario can be seen in table 2.1 below.

Scenario. Call drops in a network

Causes	Symptoms
▪ Interference ▪ Coverage losses	▪ Average Downlink User Equipment SNIR ▪ Average Uplink User Equipment SNIR

Table 3.1

Table 3.1: Causes to symptom mapping for LTE

3.4 Architecture of Self-healing framework

Currently, a lot of research has been going on in the area of SON, and many architectures have been proposed for self-healing mechanisms. In [5], the interfaces between various blocks inside the SON and the interfaces between terminal and network, for a wireless network are specified. In this architecture, the module is connected to and communicates with all the other modules present in the network such as common radio resource management (CRRM). The architecture is aimed at detection and diagnosis of the faults occurring in the network, by keeping an eye onto the symptoms specified in the different modules or elements of a SON.

According to [4], the global fault management service architecture utilizes model-based diagnosis module, and is aimed at heterogeneous wireless networks. In [6], using the belief networks and case based reasoning (CBR), a hybrid diagnosis model is proposed. A hybrid diagnosis model for SON is also specified in [7], which uses neural network and CBR techniques. Using CBR, another fault diagnosis model has been proposed in [8]. The CBR principle works on the past history of the fault, it looks into the database for any past history of the fault and then diagnoses that fault. From some architectures it has been concluded that designing any fault diagnosis model which only utilizes CBR has not been successful so far, the only reason is the inadaptability of CBR to the newer faults. Therefore while designing fault diagnosis model, CBR is always used with a combination of any other techniques, as it has been in [6] and [8].

In statistical learning automated healing (SLAH) [9], the self-healing for the handover margin in LTE network has been proposed and implemented. Using the Logistic Regression (LR) mechanisms for statistical learning the parameters: Block call rate (BCR) and average bit rate (ABR) considered as KPIs are optimized in an iterative procedure. This paper reflects more towards optimizing the parameters, and falls into the self-optimization category rather than the maintenance or self-healing.

In [2], self healing architecture has been proposed, which is aimed at locating the fault occurring at any point in the network, finding and raising the alarm by looking at the symptom values in any point of the network. The proposed architecture for self healing mechanism has been a step by step process of four different modules namely fault localization, KPI (Key Performance Indicator) thresholding and monitoring, fault modelling and fault diagnosis, which can be seen in figure 3.1.

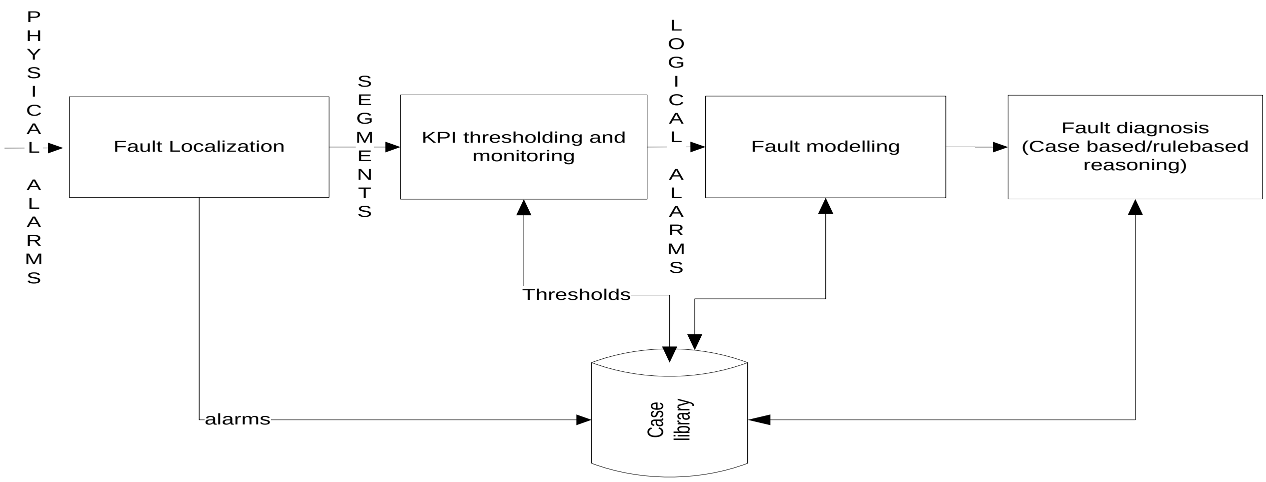

Figure 3.1: The proposed architecture for self-healing [2]

In this architecture, whenever any fault occurs in the network, single or multiple alarms will be activated in the relative network element or entity. In order to resolve the conflicting alarm activation issues i.e. multiple alarms for the same fault or similar alarm for multiple faults, alarm correlation techniques are applied in the fault localization module. Once the root cause of fault is found, the next step is to find the threshold for the symptom; the process takes place in the KPI thresholding and monitoring module. Afterwards, the fault is modelled by fault modelling techniques such as Bayesian modelling in the fault modelling module, and the final process takes place in the fault diagnosis module by applying the solution strategy such as CBR. The data from each module i.e. alarms, thresholds are stored in the database i.e. case library, and the CBR analyzes prior history of faults from database and makes diagnosis decisions, therefore the update of data from each module makes the self-healing diagnosis efficient. A brief description of all the four modules is mentioned.

- **Fault localization module:** This module is aimed at identifying the location of the fault, e.g. the segment of a network in which the fault has occurred, which may be a network element or a software module. The module receives alarms from management system and the alarm correlation techniques are applied to resolve the conflicting alarm issues. A variety of models have been proposed for alarm correlation, some of them are: model based reasoning alarm correlation, hierarchical reasoning method for alarm correlation, and AC view model. There are different approaches for fault localization process, such as fuzzy logic fault localization, distributed fault localization, and neural network fault localization approach. In this thesis, the fault localization module has been taken into consideration for self-healing framework.

- **KPI thresholding and monitoring module:** The different parameters from the network are chosen, the data is statistically analyzed, and the symptoms of the fault are calculated in the form of KPIs. The thresholds for each symptom are calculated; when any symptom goes below the threshold, then an alarm is raised. In case of occurrence of a fault, if the threshold level is exceeded by any KPI then KPI is declared as the symptom and a logical alarm is raised for the respective KPI, which is transferred to the next module. The correct selection of KPI, which are relevant to a fault, is very essential that helps to determine the most suitable root cause of occurrence of a fault. There are different methods to calculate the threshold for a symptom from a training set of data, which are further classified: entropy based methods and beta pdf method. All these methods have been detailed in the next section. In this thesis, only beta pdf method has been implemented in the self-healing framework.

- **Fault modeling:** This module maps the symptoms which are forwarded by the preceding module, with the possible causes for a fault, and the root cause of a fault is known. The probability of the root cause given the symptoms is calculated in this module. After KPI thresholding and monitoring module, this module is the essential one that helps to find the root cause and diagnose the fault with accuracy. Some of the models used in this module are Bayesian networks, incremental hypothesis updating method and code book approach.

- **Fault diagnosis:** In this module, a new case is created for the root cause along with related symptoms obtained. The newly created case is matched against the cases stored in the database using any event correlation techniques such as CBR, rule based reasoning etc. If the case is not matched then it is stored in the database, and the solution strategy is applied to diagnose the fault. If the case is found in the database, then the solution strategy is applied accordingly.

3.5 Bernoulli and Beta distribution:

Let's suppose a binary experiment, whose results will be either success (1) or failure (0). It is said that a random variable X, related to the previous experiment follows a Bernoulli distribution with parameter β (0 ≤ β ≤ 1), if X can take only the values 0 and 1 with probabilities 1- β and β respectively [10] [11] [12].

A beta random variable with parameter (a,b) has the following probability density function:

$$f(x) = \frac{\lceil(a+b)}{\lceil(a)\lceil(b)} x^{a-1}(1-x)^{b-1} \qquad (3.1)$$

where 'a' and 'b' are the positive real number, and $\lceil(x)$ is the gamma function, defined by:

$$\lceil(x) = \int_0^{\infty} x^{t-1} e^{-x} dx \qquad (3.2)$$

The mean and variance of the beta function are:

$$\mu = \frac{a}{a+b} \qquad \sigma^2 = \frac{ab}{(a+b)^2(a+b+1)} \qquad (3.3)$$

Let $x_1, \dots \dots, x_n$ be a random sample from a Bernoulli distribution for which the value of the parameter β is unknown (0 < β < 1). Suppose that the prior distribution of β is a beta distribution with parameters 'a' and 'b'. Then the posterior distribution of β, given that the random sample $x_1, \dots \dots, x_n$ is observed, is a beta distribution with parameters $a + y$ and $b + n - y$, where $y = \sum_{i=1}^{n} x^i$[13]. Therefore, if the prior distribution of β is a beta distribution, then the posterior distribution at each stage of the sampling / training set data will also be a beta distribution.

3.6 Methods to calculate threshold

3.6.1 Entropy based methods

From the information theory, entropy is defined as the measure of the uncertainty associated with a random variable. It is also defined as the average information of a random variable, given by [14]:

$$H(X) = E[-log_2(p_X(x))] = -\sum_{x\epsilon \ X} p_X(x)log_2(p_X(x)) \quad (3.4)$$

Entropy has the following interpretations:

- Average information obtained from an observation.
- Average uncertainty about X before the observation.

Considering the simplest example of two classes, we have data:

Binary random variable X , $\Omega_X = \{0,1\}$, $p_X(1) = q$

Therefore, the entropy is given as: $H(X) = -qlog_2(q) - (1-q)log_2(1-q)$

The plot of entropy for above binary random variable can be seen in figure 3.2 below:

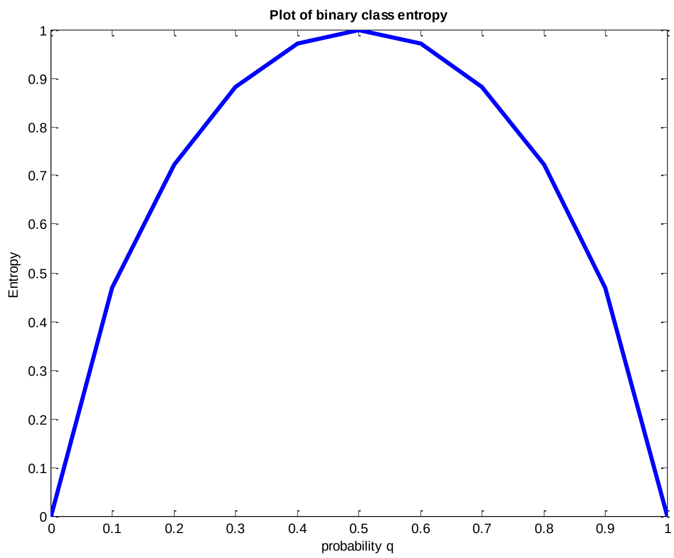

Figure 3.2: Binary class entropy

From the above example, it is observed that the entropy is minimum when q=0 or q=1, i.e. when any one of the two classes is certain; whereas the entropy is maximum when classes are equally possible i.e. q=0.5 .

3.6.1.1 Entropy Minimization Discretization (EMD)

This entropy based method for discretization has been proposed in [3] [15].

For each continuous-valued symptom S_i^c, the algorithm for EMD has been defined as:

- The cases of the training set D are first sorted out by increasing value of the symptom or attribute S_i^c. The midpoint m_j between the symptom values in each successive pair of cases in the sorted sequence is considered as a potential threshold m_j.
- Each candidate threshold m_j partitions the data set of cases D into two subsets D_1 and D_2. The class entropy of the partition is evaluated as:

$$Ent\left(D, m_j, S_i^c\right) = \sum_{k=1}^{2} \frac{|D_k|}{|D|} . Ent(D_k) \tag{3.5}$$

  where k refers to the number of causes, and $Ent(D_k)$ is the entropy of subset D_k, which is calculated using the formula 3.6.
- The best threshold t_i for S_i^c is the candidate threshold which minimizes the class entropy of the partition.

Let $|P|$denote the number of cases in a subset P and let$|P(c_i)|$ be used for the number of cases in P with c_i. The class entropy of the subset P is defined as:

$$Ent(P) = -\sum_{i=1}^{K} \frac{|P(c_i)|}{|P|} \cdot log_2 \frac{|P(c_i)|}{|P|} \tag{3.6}$$

The entropy minimization discretization (EMD) method is aimed at searching for partitions where all the cases in any of the subsets belong to the same cause c_i , so that if a symptom or an attribute value is within that interval, we can certainly assess that the cause is c_i . Therefore, the goal of the heuristic should be aimed at minimizing the class entropy of each subset.

Thus, a binary discretization for D is determined by selecting the threshold t_i for which $Ent\left(D, m_j, S_i^c\right)$ is minimal among the entire candidate cut points m_j. By applying the above method, the discretization is extended to multiple intervals, in such cases Minimum description length (MDL) method, which decides when to stop the discretization process and declare the threshold t_i .

3.6.1.2 Selective Entropy Minimization Discretization (SEMD)

The EMD is aimed at selecting the best threshold for a continuous symptom S_i^c, so that the resulting discrete symptom S_i helps to discriminate among all the causes present in a network or system. The EMD calculates the class entropy of a subset by adding over all the K causes present in a network, and this is not a realistic situation. The symptom S_i^c is related only to some causes in a network, e.g. SINR in a network is only related to causes of interference and coverage, it may not be related to the causes in the hardware or transmission equipments. Therefore, in all the causes in a network, only some cause's(s) leads to anomalous symptoms, which yield an accuracy of threshold.

In order to overcome with such the problem, the SEMD have been proposed [3], which differentiates all the causes in a network as:

- Causes related to symptom(C_r)
- Causes not related to symptom (C_n).

The causes related and not related to symptom must be sorted out first, and requires the knowledge of domain experts. Let $C_r^i = \{c_{r1}^i, \ldots\ldots, c_{rR_1}^i\}$ be the causes, which are related to symptom S_i^c, and let $C_n^i = {}^{C}/_{C_r^i}$ be the causes, which are not related to symptom S_i^c. The entropy SEMD of a subset P can be calculated according to formula (3.7) as:

$$Ent(P) = -\frac{|P(C_r^i)|}{|P|} \cdot log_2 \frac{|P(C_r^i)|}{|P|} - \frac{|P(C_n^i)|}{|P|} \cdot log_2 \frac{|P(C_n^i)|}{|P|} \tag{3.7}$$

where $|P(C_r^i)|$ denotes the number of cases in P whose cause belongs to C_r^i, and $|P(C_n^i)|$ denotes the number of cases in P whose cause does not belong to C_r^i .

The best threshold for each symptom is calculated according to the similar algorithm as of EMD. The complexity of SEMD is lower than EMD, which are evident from the equations 3.6 and 3.7. The number of operations of SEMD is 2/K times the number of operations of EMD, where K refers to the number of causes occurred in the network.

3.6.2 Beta probability density function method

In this thesis, thresholds have been calculated using beta probability density function method, known as Beta Maximum a Posteriori (BMAP) that follows directly from the theory of hypothesis testing. In the statistical hypothesis testing, the probability distributions are distinguished on the basis of random variables generated from those distributions [16].

When the BMAP is applied to univariate discretization, then it is aimed at the selection of the cause that maximizes its posterior probability, given a continuous attribute or symptom S_c^i that can be mathematically expressed as:

$$\max_j \left[P\left(\left(c_j|s_i^C\right)\right)\right] = \max_j \left[f_{S_i^C}\left(s_i^C|c_j\right).P\left(c_j\right)\right] \qquad (3.8)$$

where $f_{S_i^C}\left(s_i^C|c_j\right)$ is the conditional pdf of symptom S_c^i in s_c^igiven cause c_j .

Similar to the SEMD method, a distinction among the causes is made i.e. which causes are related to the symptom i.e. C_r^i , and which causes are not related to the symptom i.e. C_n^i. Therefore, the equation 3.8 simplifies to a form:

$$\max_{n,r}\left[P\left(C_r^i|s_i^C\right), P\left(C_n^i|s_i^C\right)\right] \qquad (3.9)$$

The threshold t_i is the cross-over point of the curves defined by (3.9), corresponding to causes related to symptom S_c^i , and causes not related to symptom S_c^i , the following equation has been accomplished:

$$\sum_{c_j \epsilon C_r^i} f_{S_i^C}\left(t_i|c_j\right).P\left(c_j\right) = \sum_{c_j \epsilon C_n^i} f_{S_i^C}\left(t_i|c_j\right).P\left(c_j\right) \qquad (3.10)$$

where it has been assumed that the causes are exclusive and the Bayes' rule has been applied.

An example of BMAP method for finding the threshold can be seen in figure 3.3 below:

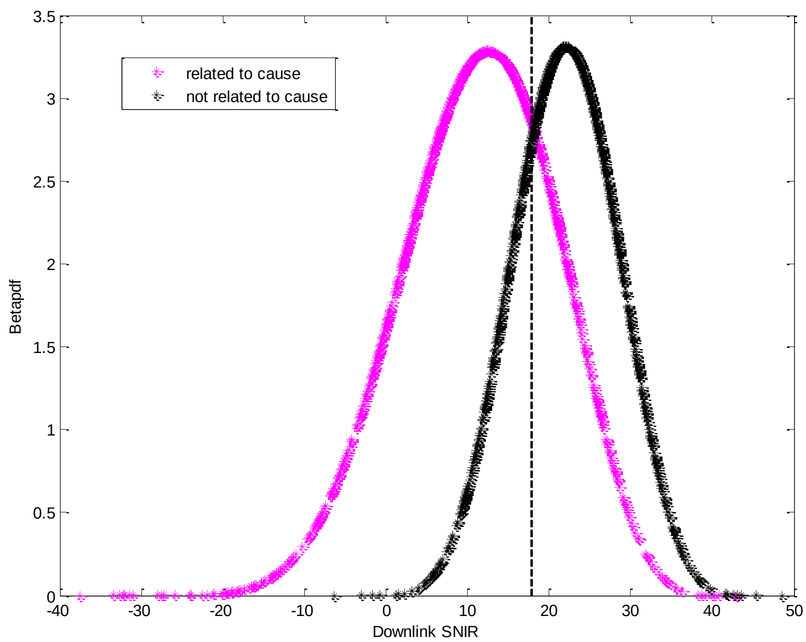

Figure 3.3: Threshold for a user for symptom downlink SINR using BMAP method

In the equation 3.10, the prior probabilities of causes can be easily obtained on the experience of occurrence of cause, but the main problem appears at the determination of conditional pdf of symptoms given the causes i.e. $f_{S_i^c}(t_i|c_j)$. Since the symptoms are modelled as continuous variables, defining the conditional pdf for such symptoms is not easy, unless they are discretized. To every problem, there is an alternate solution as well. The symptoms follow any known pdf (in this case beta pdf), then the parameters for the pdf are to be estimated to plot a known pdf that fits to the symptoms. An attention should be drawn first towards the tow known distributions: Bernoulli and beta distributions, and know how the pdf of any symptom is always estimated to a beta pdf, which is explained in section 3.4.

In order to find the parameters, it should be considered that how the symptoms which are related to the KPI are defined. This information can be known from analysis of a percentage of symptom value from training set of data, complying with a condition. In other words, the symptoms are the relative frequency of one of the two possible outcomes of an experiment (condition to create the scenario) is achieved or not, such cases can be defined and described by a Bernoulli random variable X. Let Y be another random variable, having real values in the interval [0,1], representing the expert's belief concerning the relative frequency with which X = 1 is randomly chosen. The pdf of Y can be seen as the prior belief in the β parameter of the Bernoulli distribution followed by X. Under these conditions the most adequate pdf is the beta pdf, as explained in section 3.4.

Therefore, it is concluded that the pdfs of most of the symptoms given the causes in the diagnosis model can be approximated by beta pdfs [17]. In order to find the parameters 'a' and 'b', the method of estimation is explained further in the last chapter. When the beta pdf parameters 'a' and 'b' are acquired for both the case i.e. $C_r^i \ and \ C_n^i$, then the area of cross over of these beta pdf of symptom is declared as the threshold for that symptom.

The BMAP discretization method can be summarized as a whole:

- Prior probabilities of causes are determined based on data or on knowledge.
- For each continuous symptom S_i^c and cause c_j , the parameters 'a' and 'b' of the beta pdf are estimated, which fit the training data.
- The threshold t_i is computed as the cross point between the probability of the related causes given the continuous symptom and the probability of the non-related causes given the continuous symptom.

3.7 Methods for estimating probability

After calculating the thresholds for symptoms, the probabilities are calculated. There are different methods for calculation of probability from a training set of data or symptoms, which are explained in this section. The probability can also be computed based on knowledge only, this method fits to the experts who have been working in the diagnosis section for a long time and have an estimate of probability of occurrence of a fault [3].

3.7.1 Maximum Likelihood Estimation

In MLE, the probabilities are calculated based on the frequency of occurrence of a fault or cause in the training data set. Assuming $|D|$ is the number of cases in the training set data D, $|D(c_i)|$ is the number of cases where the fault has occurred and has been observed, and$\left|D\left(c_i\,,s_j,1\right)\right|$ is the number of cases where both the cause c_i and the state $s_{j,1}$ for the symptom S_j are observed, then the conditional probability using MLE is calculated as [2][4]:

$$P\left(s_{j,1}\middle|c_j\right) = \frac{\left|D\left(c_i\,,s_j,1\right)\right|}{|D(c_i)|} \qquad (3.11)$$

$$P(c_i) = \frac{|D(c_i)|}{|D|} \qquad (3.12)$$

This method of calculating probability is not accurate enough, the inaccuracy will arise at the situation when any or both of the elements $\left|D\left(c_i\,,s_j,1\right)\right|\ or\ |D(c_i)|$ are low or even zero, and the estimated probabilities would be inaccurate. This situation is normally experienced with the MLE method, when the number of cases is limited or scarce and some probabilities are lower.

3.7.2 M-estimate (MEST)

In order to overcome the problems faced in the MLE, and improving the accuracy of probability estimation, MEST has been introduced, which works in two stages. In the first stage in MEST, the beta prior probabilities for a cause or fault are estimated and computed. The prior probability of a cause c_i is estimated by the probability of the cause happening in the next trial when there were $|D(c_i)|$ cases, where the cause c_i previous cases was $|D|$. Assuming that the initial distribution of causes is uniform, the posterior probability of the cause or fault c_i is estimated as [2][4]:

$$P(c_i) = \frac{|D(c_i)| + 1}{|D| + K} \qquad (3.13)$$

where K is the number of causes.

In the second stage, instead of uniform beta pdf, the beta pdf have been preferred as the initial probability distributions. Nevertheless, prior probabilities of causes are still calculated according to the equation 3.13. The conditional probability of symptoms given the causes $P(s_{j,1}|c_j)$ is estimated as:

$$P(s_{j,1}|c_j) = \frac{|D(c_i\,, s_j, 1)| + m.P(c_i)}{|D(c_i)| + m} \qquad (3.14)$$

where $P(c_i)$ is estimated by Laplace's law of succession according to equation 3.13, and m is a constant parameter related to the parameters 'a' and 'b' of the beta pdf.

This method has been widely used; the only disadvantage of this approach is its dependency on the history to acquire prior and conditional probabilities.

3.8 Self-healing models

The relationship between different abnormalities and causes is found by the self-healing models. These models also calculate the probability of a root cause given the system, and update it to the database. Bayesian network is very popular approach, which was originated in the early decades of the 20th century [19]. Other models include incremental hypothesis updating and codebook approach. The models for fault diagnoses help to solve two important issues in self-healing systems: how the knowledge about symptoms will be represented, and what is the cause of problems, based on the known values of some symptoms and conditions. The two main components of any diagnosis system are: model and inference method. The model is a representation of how the "world works" in the area under study; whereas the interference method is an algorithm which identifies the cause of problems based on the evidences i.e. symptoms [4].

3.8.1 Bayesian modelling

Bayesian modelling is one of the popular approaches for fault modelling till date, which finds the root cause of the problem given the symptoms. Bayesian models are capable of working in the situations, even when the training data is uncertain or imprecise [2]. In Bayesian modelling, the probability of occurrence of a fault given a symptom depends on the frequency of occurrence of the symptom, and the probability is subject to a change when a symptom appears. The two popular models of Bayesian modelling are: Bayesian Classifier and Bayesian Networks, the models based on Bayesian Networks will be considered in our self-healing framework. The basic difference between these two approaches is the design principle; Bayesian classifier assigns an unlabelled example e.g. symptoms to a cause, whereas Bayesian networks also called probabilistic belief networks and is modelled by representing the relationship between variables i.e. causes, symptoms and conditions [3][18][19].

The symptoms in a network are modelled as KPIs. The KPIs can be modelled using Bayesian modelling algorithms in two different ways: continuous model and discrete model. For defining and designing the Bayesian continuous model that delivers the performance at an acceptable level, the model demands a large amount of training data. The discrete model is preferred over the continuous model in the situation, where a lack of cases exist and obtaining a large amount of training data is not possible [2][4].

The fault modelling using Bayesian networks approach is a process consisted of 2 steps:

- Knowledge acquisition, which represents the knowledge required to identify root cause that includes the selection of the right symptoms and causes.
- Applying Inference method algorithm, which identifies the cause of the problem based on symptoms

The symptoms received from the preceding module i.e. KPI thresholding and monitoring module, could be in the form of either alarms or KPIs. In Bayesian Networks, causes and alarms are modelled as discrete random variables, having binary states: yes/no; whereas the KPIs are modelled as discrete random variables, having a discrete number of states each state corresponds to a range of KPIs i.e. {normal, high, medium, and low} [4][19].

3.8.1.1 Knowledge acquisition:

The process of knowledge acquisition (KA) in the Bayesian Networks have further been split into two steps namely the qualitative and quantitative. The qualitative part, different variables (symptoms, causes and conditions) and their dependency will be identified; whereas in the quantitative part, the probabilities that link the different variables are specified [2] [4] [18].

In quantitative model, probabilities of the KPIs in the network can be obtained from a variety of sources, which can be used for automatic build up of the Bayesian Network [4]. Since the model is designed as a discrete model here, therefore the information about the discretization of continuous variables i.e. symptoms should be specified here as well. The probabilities which are needed in the knowledge acquisition process are: prior probabilities of causes i.e. $P(c_i)$, and conditional probabilities of the symptoms given causes i.e. $P\left(s_j \middle| c_j\right)$.

The knowledge acquisition process is a step by step process, which can be seen in figure 3.4 [3] [19]. A Knowledge Acquisition Tool (KAT) may be built, which is based upon all the six steps mentioned in the figure 3.4 above. This tool makes its use easier, and can be used by the network management troubleshooting experts, the non-familiarity with the Bayesian Networks remains its major advantage [3]

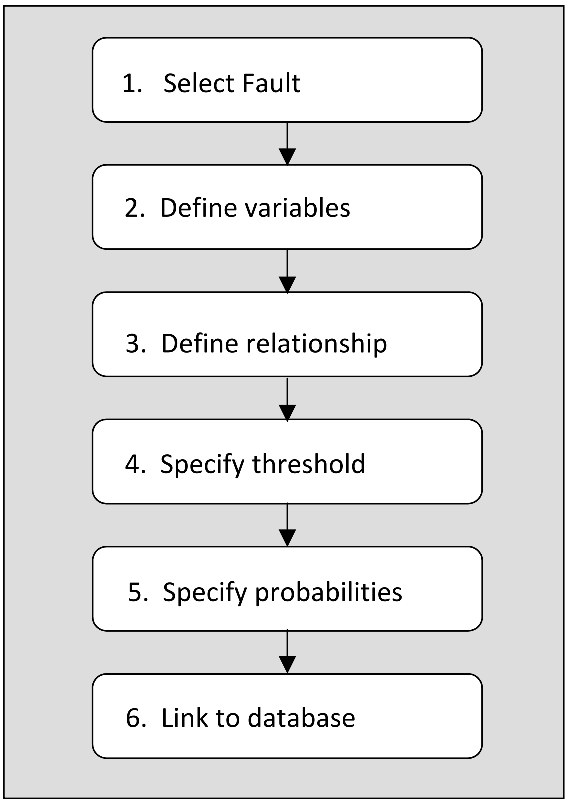

Figure 3.4: Different phases in knowledge acquisition process for Bayesian Networks

3.8.1.2 Applying Inference Method

After the knowledge acquisition process has been completed, the inference method is applied by calculating the probability of each possible cause, given the values of symptoms $\{S_1, S_2, \ldots\ldots, S_M\}$, the probability of cause c_i can be computed as [3] [19]:

$$P(c_1|S_1, \ldots\ldots, S_M) = \frac{P\,(c_i).\prod_{j=1}^{M} P\,(S_j|c_i)}{\sum_{n=1}^{K} P\,(c_i).\prod_{j=1}^{M} P\,(S_j|c_i)} \qquad (3.15)$$

where $P\,(c_i)$ is the prior probabilities of the cause c_i , and $P\,(S_j|c_i)$ are the probabilities of the symptoms given the causes.

The prior probabilities of the causes $P\,(c_i)$ are usually obtained by experts or calculated from training data as the frequency of occurrence of each type of fault.

After the Bayesian Network model has been designed, the root cause of a fault is found by taking the values of variables from the preceding module i.e. KPIs or symptoms and alarms. The final outcome of the system is a list of causes that are ranked by their respective probabilities [3].

3.8.2 Code book approach

The term codebook refers to a subset of symptoms, which are chosen to offer the desired level of distinction between the causes. In order to represent the dependency of causes and events, the root cause(s) of symptom(s) are considered as nodes, and is shown by directed edges. Such a graph of dependence is then transformed in to correlation matrix, representing root causes in column and events in row of matrix [4].

This approach uses the binary concept, and a problem is represented by a binary vector. Consider that there are two causes C_1, C_2 and two symptoms S_1, S_2 , the correlation graph for these causes and symptoms can be seen in figure 3.5 below:

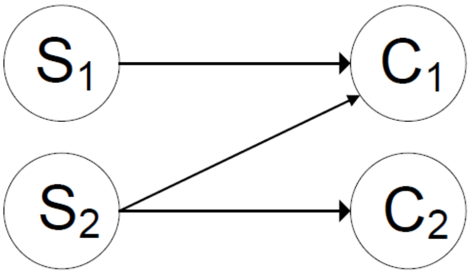

Figure 3.5: Correlation graph between causes $\boldsymbol{C_1, C_2}$ and symptoms $\boldsymbol{S_1, S_2}$

The code for C_1 is (1,1) and C_2 is (0,1), where the first bit for each code represent symptom S_1 and second bit represent symptom S_2. The codebook for the above example can be seen in table 3.2 below:

	C_1	C_2
S_1	1	1
S_2	1	0

Table 3.2: Codebook for causes $\boldsymbol{C_1, C_2}$ and symptoms $\boldsymbol{S_1, S_2}$

In order to find the root cause of a problem, the codes for each cause are matched against an observed symptom vector, and the one whose code optimally matches the symptom vector, is declared as the root cause of the fault. The hamming distance measures the distinction among the causes, and half of the minimal distance between codes is known as the radius. In the example shown in table 3.2, the hamming distance is 1 and the radius is 0.5.

This method can be used to detect and rectify the lost or a spurious symptom. In general, the code book approach can rectify observation errors in $k-1$ symptoms, and detect errors in symptoms as long as $k \le radius\ of\ codebook$ [2].

References for Chapter 3

[1]. E3. (2008, December 22). White paper : Self-x in Radio Access Networks. (I. G. Eckard Bogenfeld, Ed.) Retrieved March 4, 2010, from End to End Efficiency (E3) FP7 EU Project: https://ict-e3.eu/project/white_papers/Self-x_WhitePaper_Final_v1.0.pdf

[2]. Ramkumar, V. (2011). Self-healing architecture and methods. Literature Survey, Aalborg University, Denmark, Networking and Security.

[3]. Clark, V. (2000). *To Maintain an Alarm Correlator.* Wales, Australia: The University of New South Wales.

[4]. Moreno, R. B. (2007). Bayesian modelling of fault diagnosis in mobile communication networks. Higher Technical School of Telecommunication Engineering. University of Malaga, Spain.

[5]. Ulrich Barth, E. K. (2010). Self-organization in 4G Mobile Networks : Motivation and Vision. Wireless Communication Systems (ISCWS). IEEE.

[6]. Usama M. Fayyad, a. K. (1993). Multi-Interval Discretization of Continuous-Valued Attributes for Classification Learning. *The International Journal of Computational Intelligence and Applications (IJCIA)*, *2*, ss. 1022-1027.

[7]. Gardner, R.;& Harle, D. (1997). Alarm Correlation and Network Fault Resolution using the Kohonen Self-Organizing Map. *Global Telecommunication Conference (GLOBECOM). 3*, ss. 1398 - 1402. IEEE.

[8]. Chao, C.;Yang, D.;& Liu, A. (1999). An Automated Fault Diagnosis System Using Hierarchical Reasoning and Alarm Correlation. *IEEE Workshop on Internet Applications* (ss. 120 - 127). IEEE.

[9]. Moazzam Islam Tiwana, B. S. (2010). Statistical Learning-based Automated Healing: Application to Mobility in 3G LTE Networks. 21st Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (ss. 1746-1751). IEEE.

[10]. Srihari, S. N. (n.d.). *Machine Learning and Probabilistic Graphical Models Course.* Retrieved March 10, 2011, from University at Buffalo: http://www.cedar.buffalo.edu/~srihari/CSE574/

[11]. Beta Distribution. (2011, April 3). Retrieved from Mathworks: http://www.mathworks.com/help/toolbox/stats/brn2ivz-4.html

[12]. Binomial Distribution. Retrieved April 4, 2011, from Mathworks: http://www.mathworks.com/help/toolbox/stats/brn2ivz-9.htm

[13]. Parmigiani, G. (2002). Modeling in Medical Decision Making. John Wiley & Sons.

[14]. Irani, U. F. (1993). Multi-interval discretization of continuous valued attributes. International Joint Conference on Artificial Intelligence, (ss. 1022-1027). Chambery, France.

[15]. Renfors, M. (2008). TLT-5406 Digital Transmission Lecture Notes. Tampere, Finland: Tampere University of Technology.

[16]. Rice, J. (2006). *Mathematical statistics and data analysis* (3rd Edition p.). USA: Cengage Learning.

[17]. R. Barco, V. W. (2005). System for automated diagnosis in cellular networks. European Trans. Te*lecommunications , 16* (5), 399–409.

[18]. Paola Sebastiani, M. M. (2010). *Data Mining and Knowledge Discovery Handbook* (2nd Edition p.). SpringerLink

[19]. Rana M. Khanafer, B. S. (July 2008). Automated Diagnosis for UMTS Networks Using Bayesian Network Approach. IEEE Transcations on Vehicular Technology , 57 (4), ss. 2451-2461.

Chapter 4 . Simulation Results

4.1 Description of Matlab Simulation

In order to analyze the effect of different parameters on the network performance, an LTE network is simulated in Matlab. The simulated environment consists of 4 base stations i.e. eNB, and 50 physical resource blocks (PRBs) which are randomly allocated to each user depending upon the type of user, where type of user relates to the data rate of the user. If the PRBs are in decreasing order i.e. [5 2 1] then the user data rate would be 2 Mbps, 800 Kbps and 400 Kbps respectively. The PRBs are allocated equally among all the users i.e. [1 1 1], therefore all the users will have same data rate and the interference caused can be seen easily.

A typical view of the network setup in the simulation can be visualized in figure 4.1.

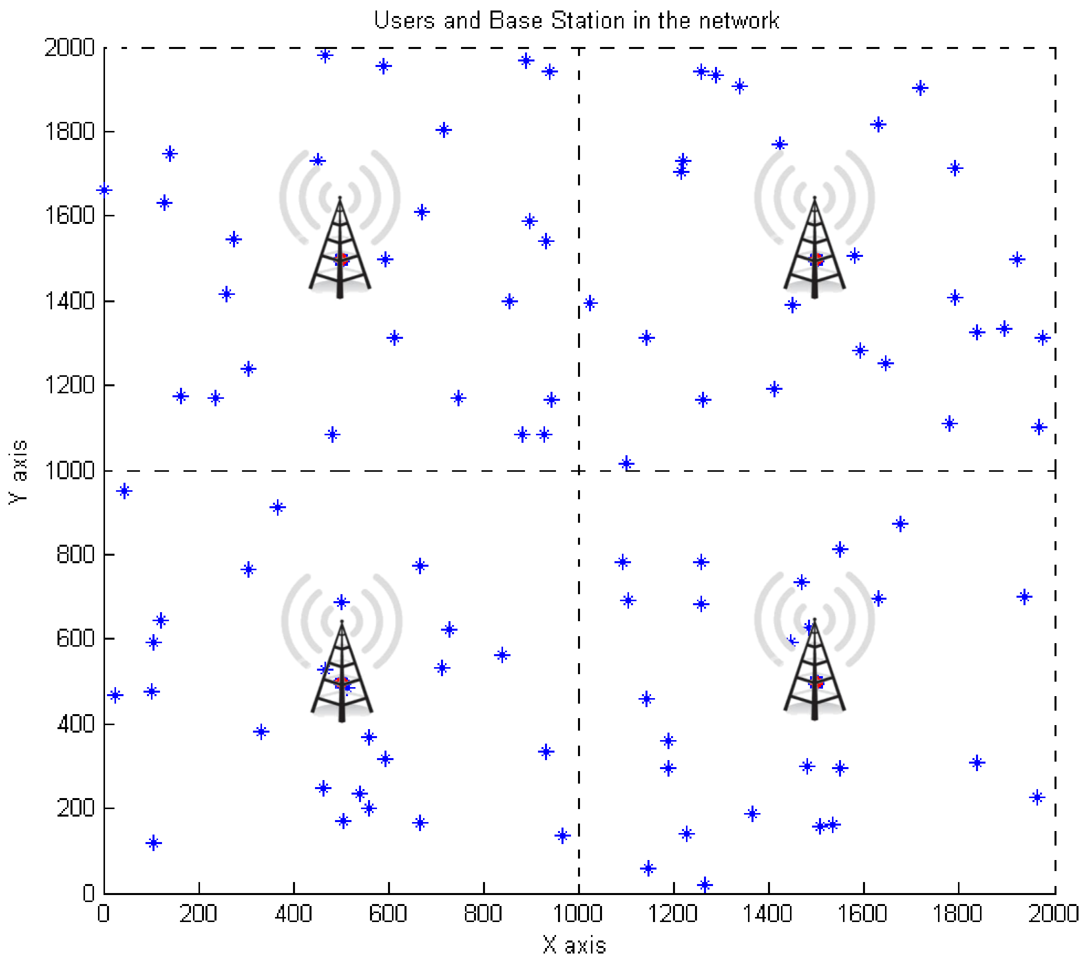

Figure 4.1: A view of network setup in Matlab simulation

The system is set to have full load, the resources are fully utilized, and the target SINR to be achieved is set to 20 dBm. The PC at uplink as well as downlink can be set to five different states: no PC, full or partial PC, disabled, closed loop PC with target SINR having same eNB power across spectrum, and closed loop PC with target SINR having same eNB power for each UE. The default simulation is set for 100 frames, where each frame has duration of 10 msec in the network. All the users in the network are mobile, and have a default velocity of 1 km/h. The radius of the each cell by default is 500m, the coordinates of eNBs and users are allocated randomly according to the customized functions.

The total time (t_{total}) taken by the network and the total no. of frames ($frames_{total}$) needed to reach the boundary of each cell, can be calculated as:

$Velocity = 1$ km/h $= \frac{1000}{3600}$ m/sec = 0.2778 m/sec

$t_{total} = \frac{Radius}{Velocity} = \frac{500}{0.2778} = 1800$ sec

$frame_{total} = \frac{t_{total}}{Frame\ duration} = \frac{1800}{10\times10^{-3}} = 180000$ frames

The velocity with which the users will move is decided, and then the total no. of frames ($frames_{total}$) are calculated, which help to determine that for how much frames the simulation should run in order to create the cause and analyze the respective symptoms. After calculating these parameters, the simulation is run, and all the users reach the boundary of their respective cell, which helps in creating causes similar to realistic situations in the simulation. Although in the real situations, the users roam in the nearby cells, but the roaming situations are not considered in the simulations, instead the users are limited within the cell boundary by bouncing them back into the respective cell coordinates.

4.2 Interference at Uplink and Downlink

One of the major cause for call drops in a network, may be interference occurred at either uplink or downlink or at both pathways. The interference in the network is created by setting down the power control at both pathways i.e. uplink and downlink power control is turned off completely. By turning off the power control, in the uplink all the users transmit with the maximum power of 0.3162 W, and in the downlink the available power is equally distributed among all the PRBs i.e. BS transmits with equal power to all users in the downlink.

In this simulation, some parameters are modified such as the radius of the cell is reduced to 150 m, the velocity is increased to 100 km/h, and similarly the total no. of frames is set as 1100 frames. It has been ensured that while the user is moving from one cell with a velocity, it does not cross the boundaries of that cell and does not mix with the users of other cell. The users in each cell move with a specific velocity with a specified angle, when the user reaches the boundary of the respective cell, 180 degree is added to the angle of that user, and therefore it bounces back and moves further with a defined velocity. The symptoms which represent interference at uplink and downlink are average user equipment SINR uplink (AV_UE_SINR_UL) and average user equipment SINR downlink (AV_UE_SINR_DL) respectively.

A view of the network focusing the position of a randomly chosen user at the start and the end of simulation can be seen in figure 4.2, and the movement of users in each cell can be seen in figure 4.3.

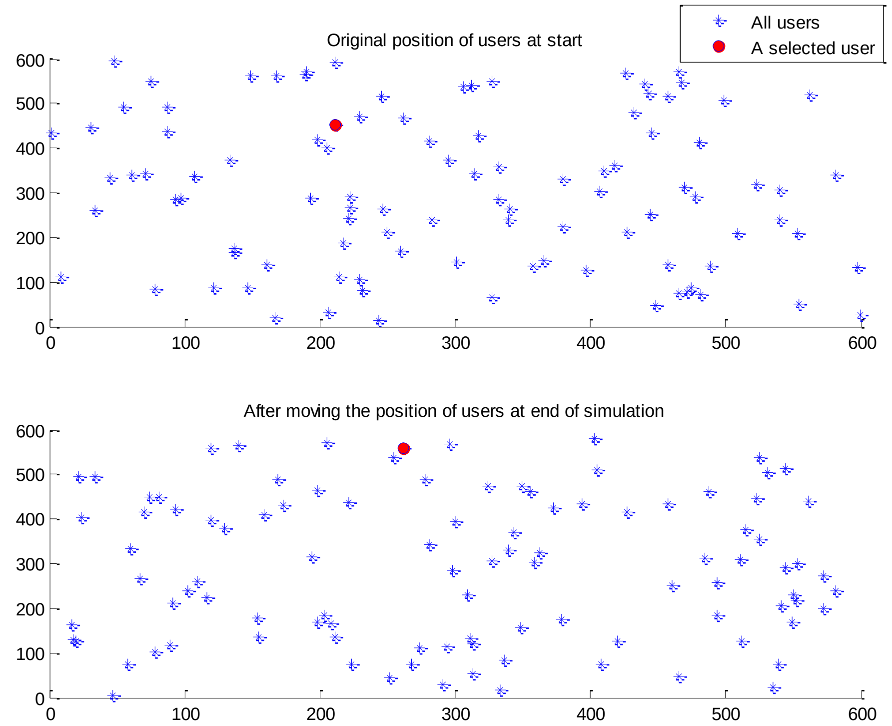

Figure 4.2: The position of a randomly chosen user at the start and the end of simulation

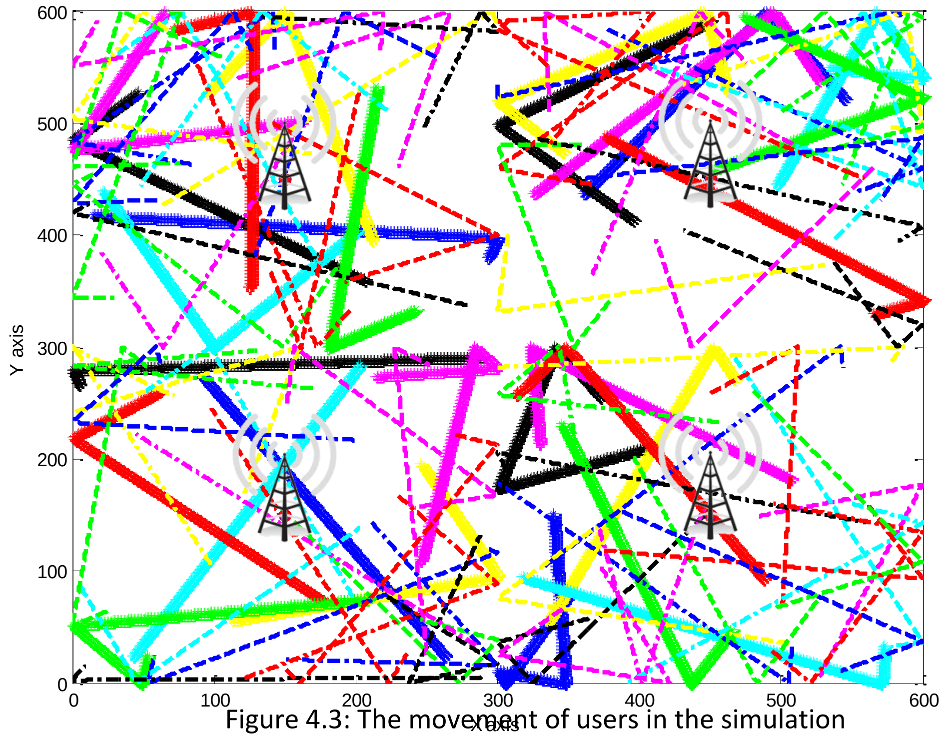

Figure 4.3: The movement of users in the simulation

4.2.1 Interference at Uplink

By turning off power control at uplink, all the users transmit with the maximum power of 0.3162 W. After the simulation runs, the symptom average user equipment SINR uplink (AV_UE_SINR_UL) is analyzed for all those users, and if there has been any interference caused, the plots of symptom will show a decrease in the AV_UE_SINR_UL. The interference is subject to the users from the adjacent cells, therefore at initial, the users present in each cell are determined, and then user coordinates are checked first to see that during which frames user is closer to boundary, it touches the boundary and then bounces back. Once the information about the frames during has been known, then the adjacent cells are checked during the same frames, to see whether there are any users closer to the boundary of adjacent cells. If the users have been found during the same frames, then it is confirmed that the caused interference is due to the users from neighbouring cells.

In the first simulation, after checking out the coordinates and analyzing symptom plots of each user, these users were detected to have caused uplink interference: 3, 4, 5, 9, 10, 20, 25, 26, 27, 29, 33, 35, 39, 40, 46, 55, 71, 81, 85, 89.

The users 4 and 20 experiences interference caused at uplink during the frames 700-900 and 700-800, which can be seen in figure 4.4 and figure 4.5 respectively. The neighbouring users also experience interference while they reach the boundary and bounce back. The user 4 (cell 4) reaches the boundary and then bounces back, during the same frames user 26 from adjacent cell (cell 3) reaches the boundary during frames 200-400. Therefore both experience the interference, which can be seen in figure 4.6.

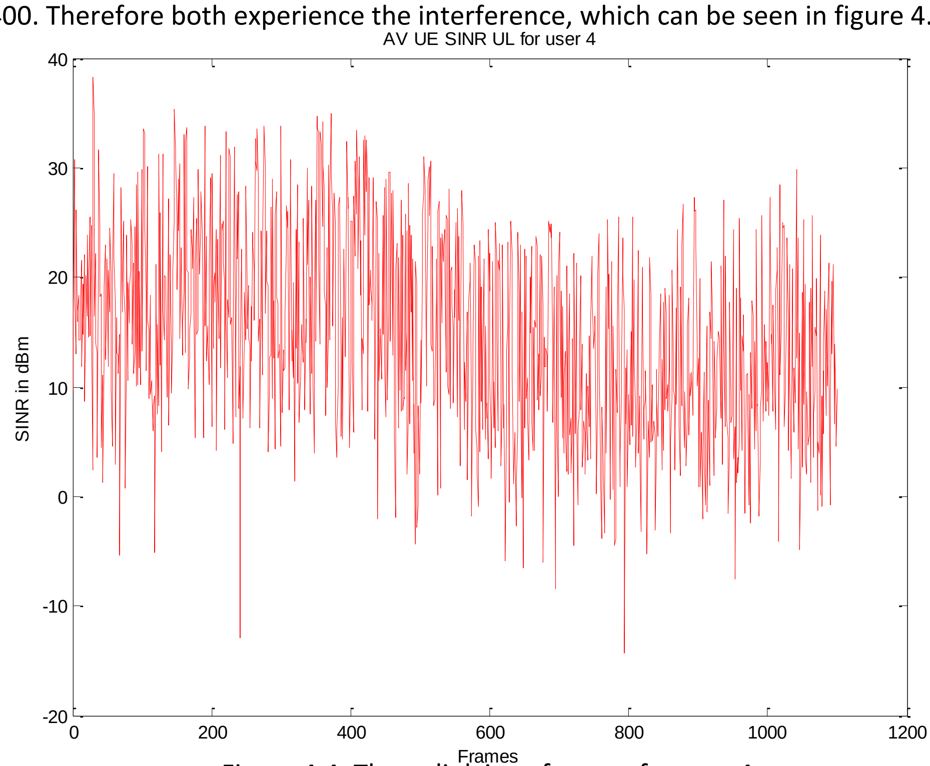

Figure 4.4: The uplink interference for user 4

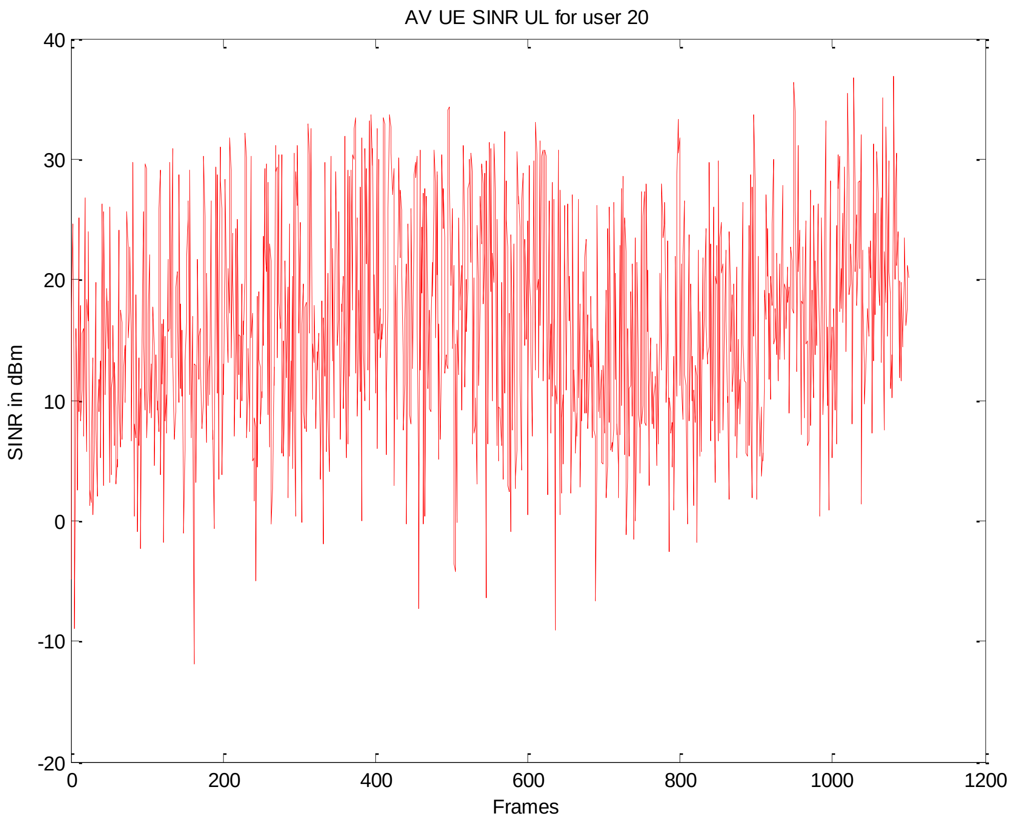

Figure 4.5: The uplink interference for user 20

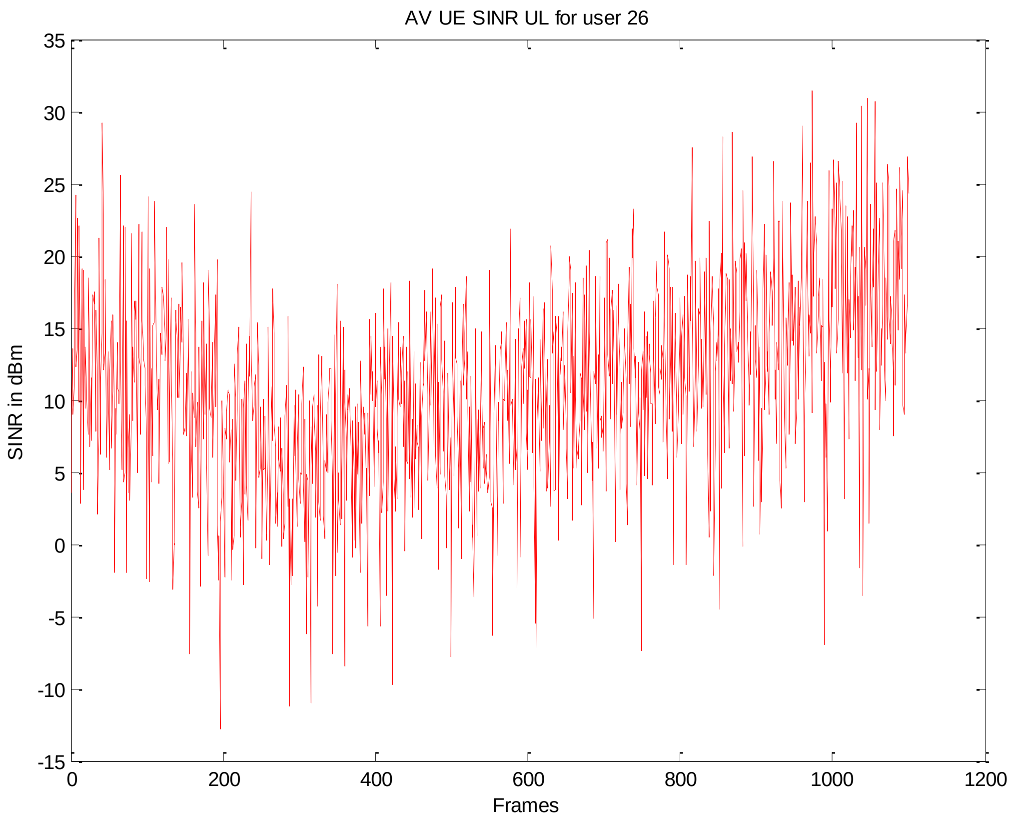

Figure 4.6: The uplink interference for user 26

4.2.2 Interference at Downlink

By turning off power control at downlink, the base stations transmit with equal power, equally distributed among all the available PRBs. Since the initial load of the system is set to have full load i.e. 1, therefore the available power is equally divided among all the available PRBs. After the simulation has run, the same process is applied to find the users which cause interference, as mentioned in the section 3.2.1. And finally, the symptom average user equipment SINR downlink (AV_UE_SINR_DL) is analyzed for all the users which experience interference.

Interference is caused at both uplink and downlink at the same time, therefore in the first simulation, after checking out the coordinates and analyzing symptom plots of each user, the same users were detected to have caused downlink interference: 3, 4, 5, 9, 10, 20, 25, 26, 27, 29, 33, 35, 39, 40, 46, 55, 71, 81, 85, 89.

The users 35 and 55 experiences interference caused at uplink during the frames 600-900 and 300-600, which can be seen in figure 4.7 and figure 4.8 respectively. The neighbouring users also experience interference while they reach the boundary and bounce back. The user 35 (cell 1) reaches the boundary and then bounces back, during the same frames user 6 from adjacent cell (cell 4) reaches the boundary during frames 600-800. Therefore both experience the interference, which can be seen in figure 4.9.

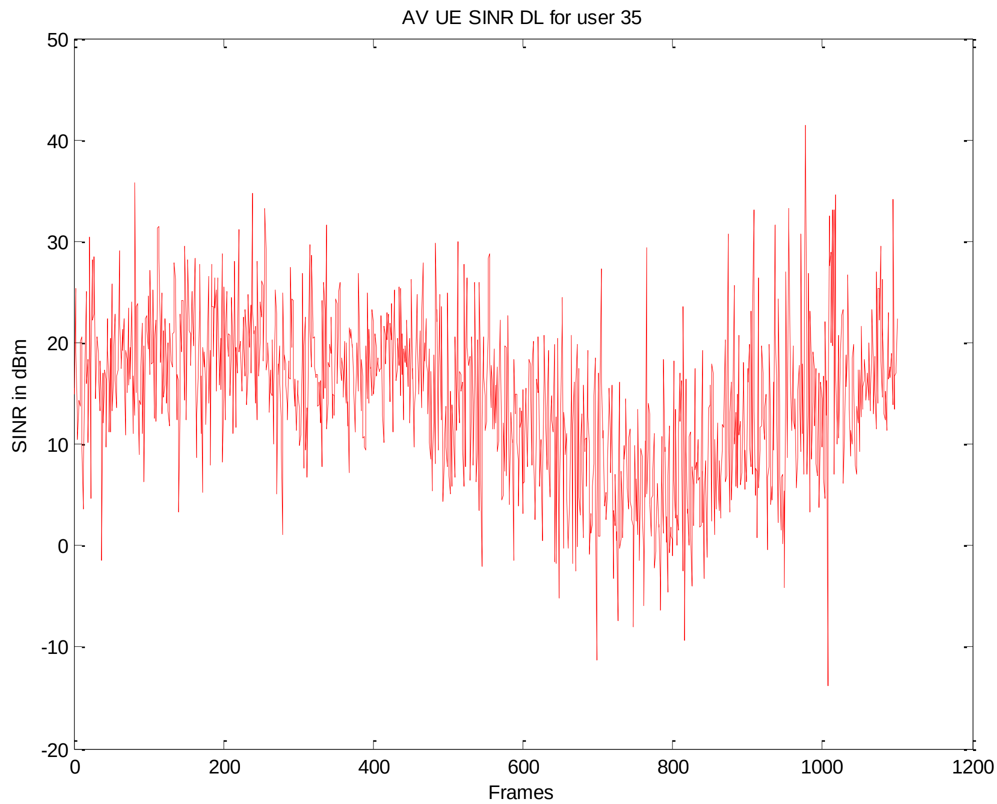

Figure 4.7: Interference caused at downlink for user 53

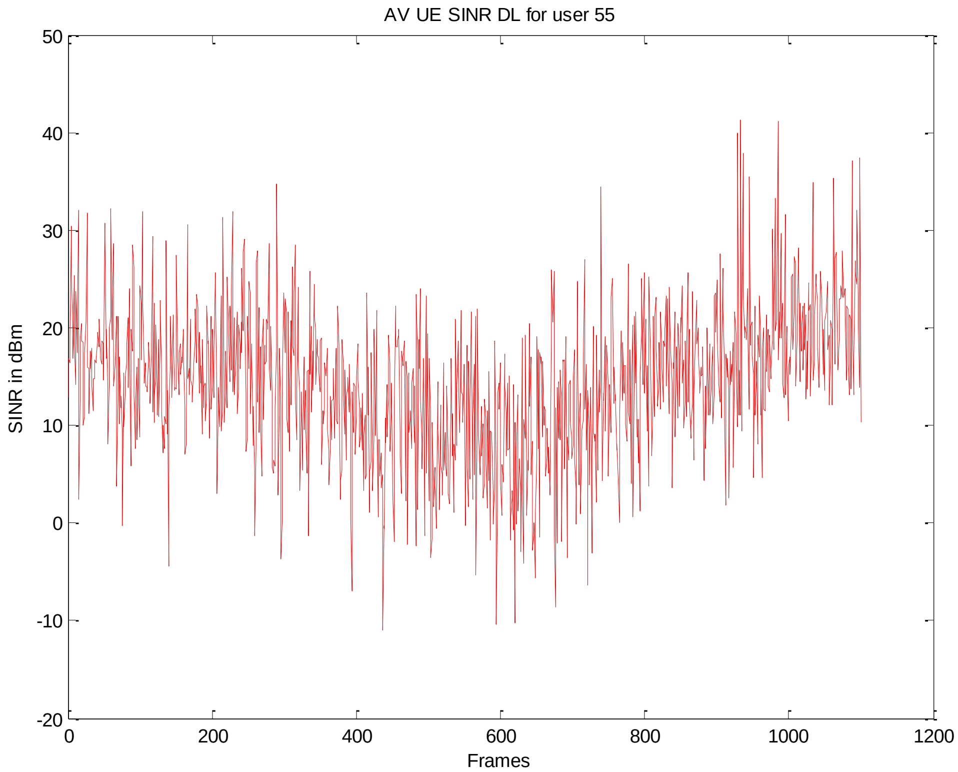

Figure 4.8: Interference caused at downlink for user 55

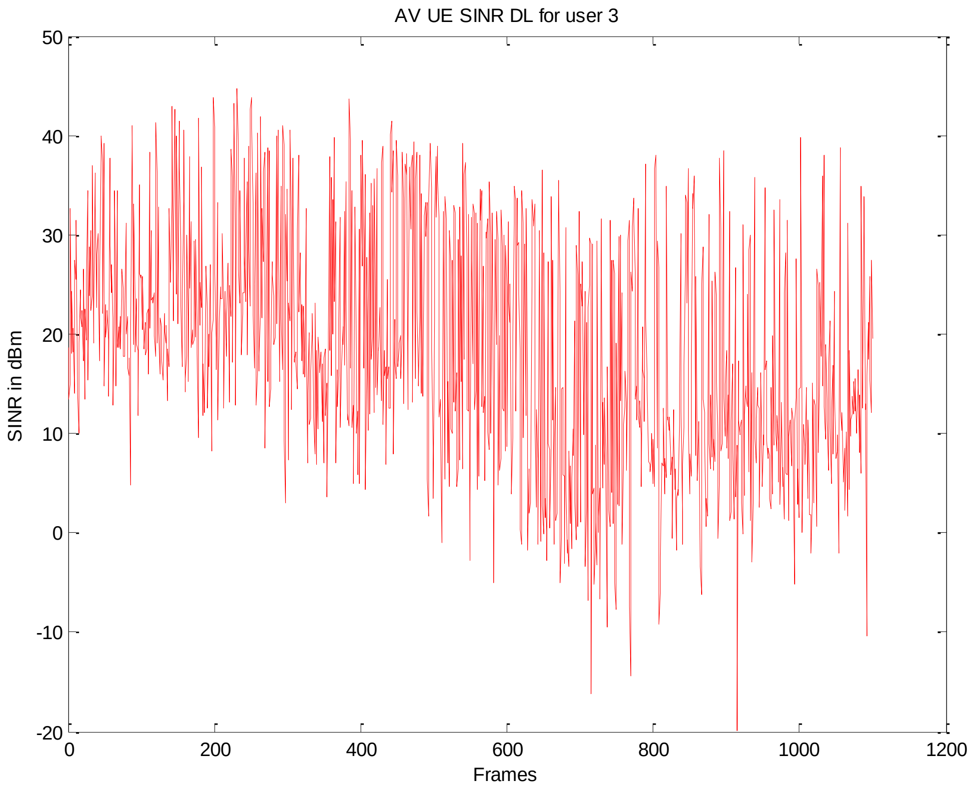

Figure 4.9: Interference caused at downlink for user 3

4.3 Coverage at Downlink

The other reason for call drops in a network may be due to coverage problem occurred at downlink. The possible reasons for coverage problem could be shadowing and fading occurred in the propagation path. The fading has been randomly generated and fading losses have been taken into account in the each simulation. The cause for coverage scenario considered here for the simulation is the shadowing. When the test drive or radio network planning process was being carried out upon the time of deployment of network, the shadowing losses were estimated to a value. After the network deployment, some obstacles may be raised such as trees or high rise buildings or residency in the non-occupied area. The addition of such obstacles creates a shadowing loss due to multipath propagation, and the users in that specific area face shadowing losses, and finally the SINR at downlink is decreased.

The area in which obstacles have been created that cause shadowing losses is specified in table 4.1, and the view of network having obstacles can be seen in figure 4.10 below:

Cell number	**Obstacles area x axis**		**Obstacles area y axis**	
	Start	End	Start	End
Cell 1	50	90	510	550
Cell 2	350	390	520	560
Cell 3	500	540	200	240
Cell 4	0	40	100	140

Table 4.1: Area for obstacles added to the cells

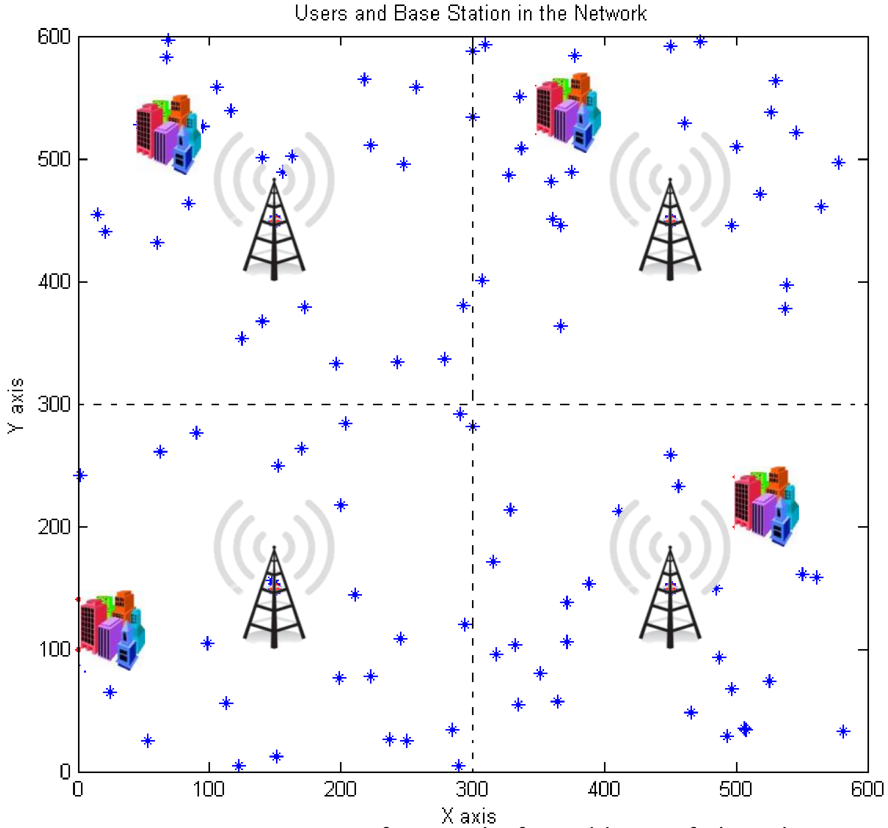

Figure 4.10: A view of network after addition of obstacles

The simulation parameters are similar to those in the interference scenario simulation. The power control has been turned off, so that shadowing losses can be visualized easily. If the power control would not have been turned off, the shadowing losses could overcome by power control mechanism and could not be seen in the symptom average user equipment SINR downlink (AV_UE_SINR_DL). Any user while moving in the cell area comes to the obstacles area, then shadowing losses are added to the path loss for the duration of those frames, as a result downlink SINR is reduced.

When the simulations were run for the same users coordinates and same motion path, the following users the following users were detected to have caused shadowing losses at downlink:
23, 37, 38, 48, 58, 65, 73, 77, 78, 83, 88, 95, 97.

The above users are detected first by analyzing the coordinates and then analyzing the symptom, in the simulation there is no any user that have also experienced interference losses. If there would have been the same user, then the identification process is subject to coordinates check up. In the simulations done till this stage, this process has not been utilized, but it will be utilized in the next stage for detection of thresholds.

The users 58 and 73 experiences shadowing losses caused at downlink during the frames 525-750 and 200-380, which can be seen in figure 4.11 and figure 4.12 respectively.

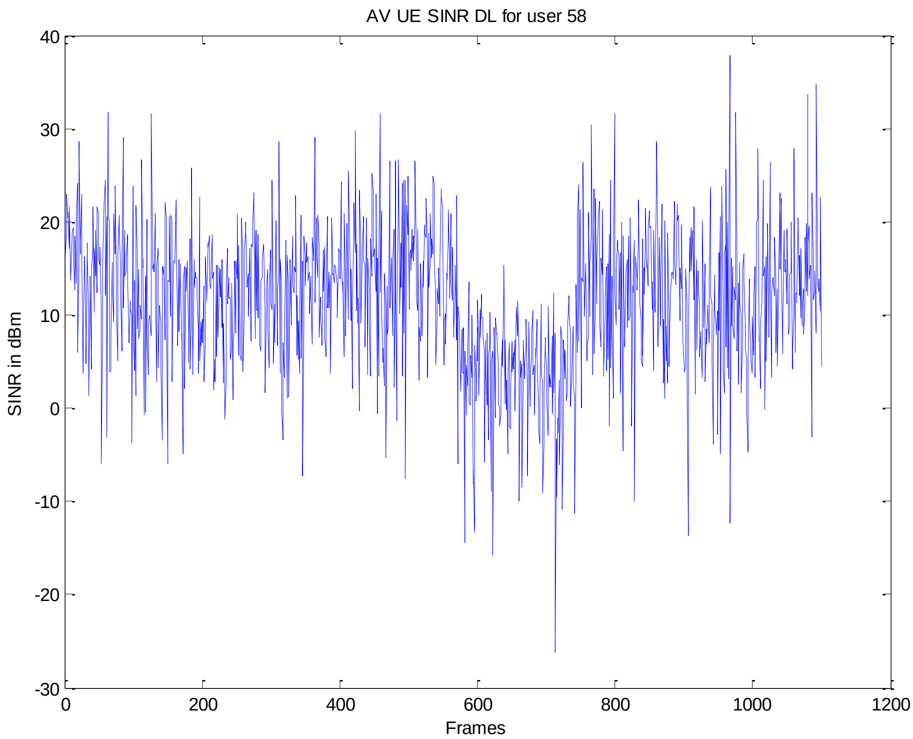

Figure 4.11: Shadowing losses caused at downlink for user 58

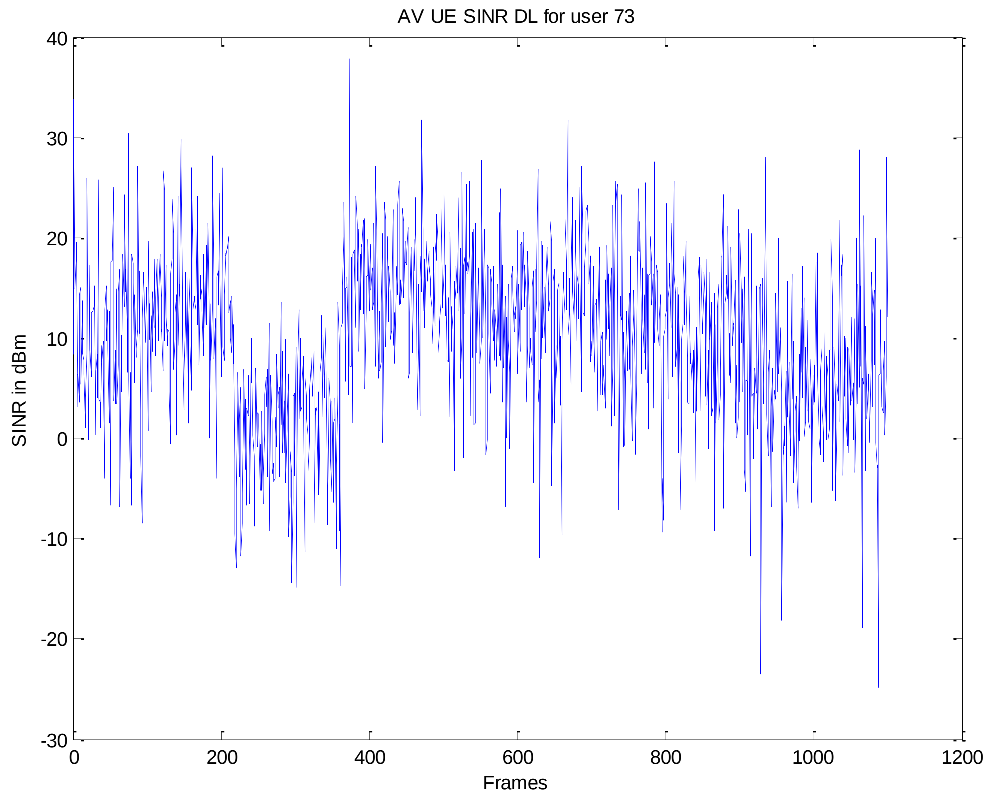

Figure 4.12: Shadowing losses caused at downlink for user 73

4.4 Setting up threshold for symptoms

After the symptoms have been analysed, the next step is to define the threshold for each user for the respective user. The thresholds upon computation are stored in the database, when the simulation is re-run, the self-healing framework will detect the symptom values for each user and will compare it with the threshold lying in the database. If the symptom value is less than the threshold in the database, then it will detect the problem or fault and will raise an alarm as a notification of fault.

According to [1], the probability density function (PDF) of the symptoms can be approximated using beta pdf. In order to compute beta pdf for a symptom for any user, the parameters 'a' and 'b' are required, which are estimated by a step by step process. At first, the symptom value or training set of data will be plotted as histogram, then the training set of data will be plotted using different values of parameters 'a' and 'b' for beta pdf. This iterative process will continue as long as the minimum difference between the histogram plot and beta pdf plot is found, and those minimum values will be declared as the beta pdf parameter values, an example can be seen in figure 4.13 on the next page.

For every user that experiences interferences or shadowing loss, the beta pdf is approximated. In order to find the threshold, we will consider another training set of data which is not related to causes anymore i.e.

a data which is free from interference as well as shadowing losses. Therefore, a training data has been generated in which the boundary limits for movement of users have been reduced, and users don't any experience any interference or shadowing losses in that training set of data. An overview of the movement of users in such situation can be seen in figure 4.14.

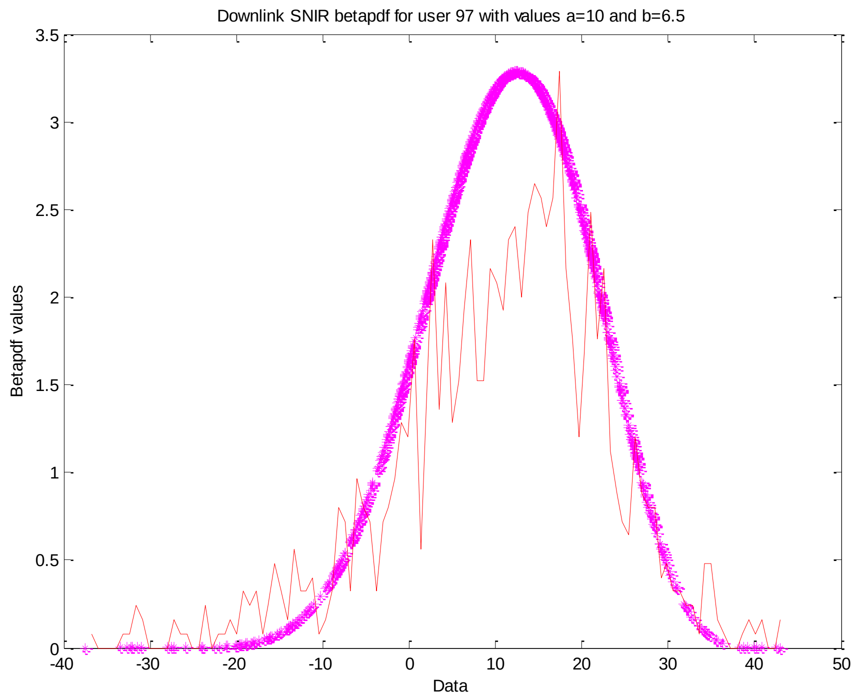

Figure 4.13: Approximation of beta pdf parameter values 'a' and 'b'

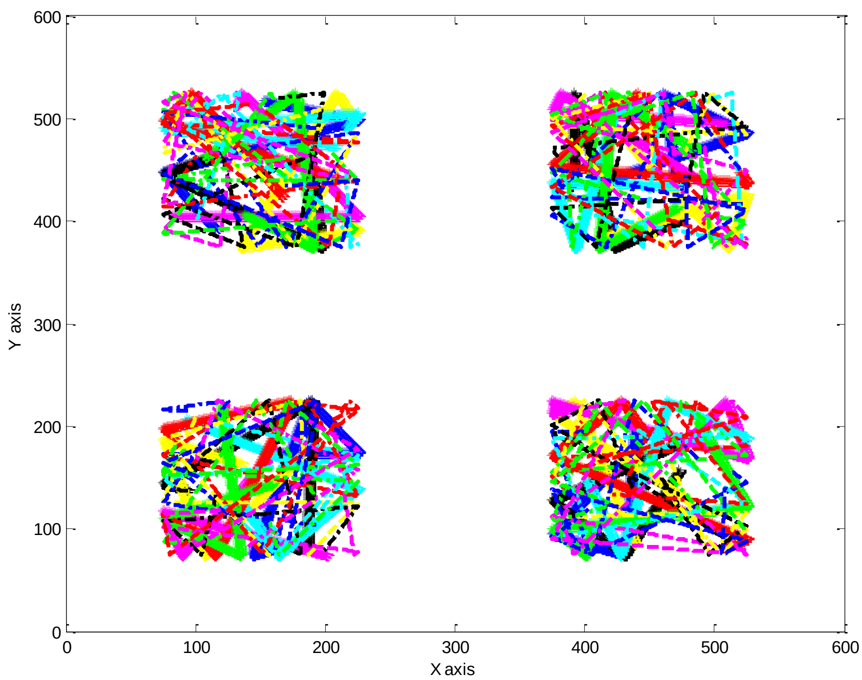

Figure 4.14: An overview of users' movement in the training data not related to any cause

The beta pdf parameters 'a' and 'b' will be estimated for the training data not related to any cause in the same manner. After the beta pdf parameters 'a' and 'b' for both the training set of data (one related to causes and other not related to any cause) have been estimated for each users, then the beta pdf for both the training data sets will be plotted in the same window. The point/area at which both the beta pdf plots overlap or meet each other, the value of symptom at that point is declared as the threshold for that user. Moreover, from these beta pdf plots it can be verified that whether the user has faced any interference or shadowing losses or not, the beta pdf plot of a user facing any interference or shadowing losses would have lower SNIR values than the other beta pdf plot not facing any losses. The beta pdf plots for threshold estimation for user 95 and 97 can be seen in figures 3.15 and 3.16 below.

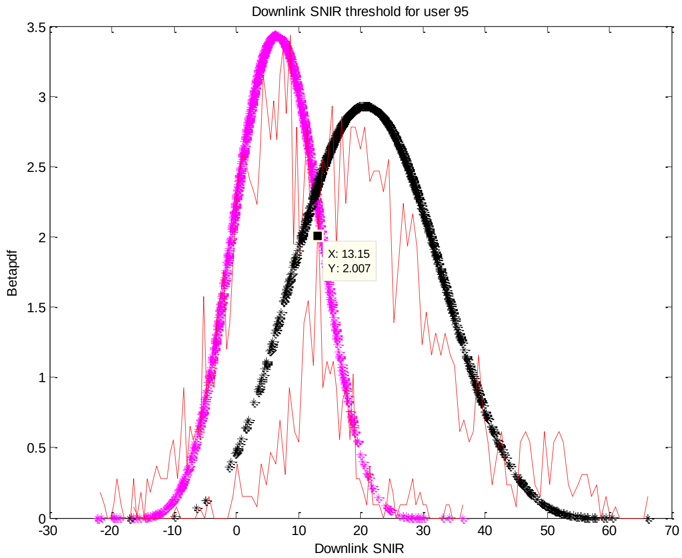

Figure 4.15: Threshold for user 95

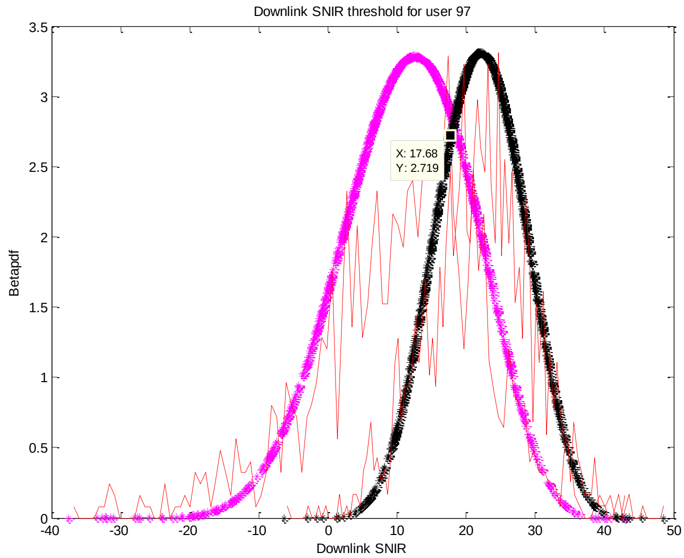

Figure 4.16: Threshold for user 97

The threshold results for the users which experience shadowing losses are specified in table 4.2 below:

User No.	**Downlink**				
	Without shadowing		With shadowing		Threshold (dBm)
	a	b	a	b	
23	3.6	5.3	4.7	3.9	19.04
37	9.5	7.8	10	5.7	19.13
38	10	8	7.9	5.4	16.97
48	10	6.9	10	6.1	18.22
58	6.9	7	10	7	19.2
65	10	6.6	7.4	5.2	18.2
73	7.9	8.8	10	7.9	17.06
77	10	8.4	10	7.9	13.81
78	5.1	9.4	8.5	7.3	13.36

83	10	8.2	8.4	6.9	16.24
88	6.9	6.5	10	9.7	13.87
95	6.4	7.5	9.3	9.7	13.15
97	9.1	8.6	10	6.5	17.68

Table 4.2: Thresholds for users experiencing shadowing losses

The threshold results for the users which experience interference losses are specified in table 4.3 below:

User No.	**Downlink**					**Uplink**				
	Without Interference		With Interference		Threshold (dBm)	Without Interference		With Interference		Threshold (dBm)
	a	b	a	b		a	b	a	b	
3	7.1	5.3	4.9	3.3	23.58	6	5.1	6.7	4.1	23
4	6.3	7.3	3.1	3.1	17.93	6.6	8.6	4.6	3.8	16.15
5	3.8	3.7	4.8	5.3	32.63	5.7	4.4	4.7	4.9	25.41
9	4.2	4.7	5.3	3.7	21.48	5.3	5	4.6	4.6	19.39
10	4.1	4.6	4.3	4.4	23.06	5.2	5.3	5.3	5.3	23.06
20	3.6	3	5.2	3.5	26.66	5.5	4	4.7	3.3	22
25	6.7	4.7	3.1	2.8	16.2	7.7	5.5	5.7	3.4	20.56
26	6.3	7.4	10	7.3	15	7.4	6.3	8.2	7.1	17.2
27	9.4	6.7	8.7	9.2	18	8.1	5.7	8.1	7.4	19.94
29	10	6.4	4.4	6.6	12.32	10	5.2	6.6	5.6	15
33	8.7	6.7	6.3	6.4	18.16	10	6	6.2	4.9	15.52
35	7.2	5.8	8.6	8.1	20.99	9.8	7	9.9	7.3	20.01
39	8.7	9.2	8.7	8.5	24.79	5.6	5.7	5.1	5	17
40	8.1	8.3	10	7.9	21.72	7.9	6	9.8	6.4	21.31
46	10	10	11.5	12.7	15.44	10	9.9	9.6	8.9	18.21
55	10	6.8	9.1	9.8	18.78	7.2	6.9	10	6.4	19.85
71	5.9	6.3	9.7	8.7	21.01	5.9	5.3	8.2	6.3	21.49
81	6.7	10.5	4.4	6	13.93	6.9	9.8	4.2	6.3	15.93
85	10	8	8.2	8.3	16.18	8.6	7.5	9.1	6.6	17.62
89	7.7	8.7	8.5	9.7	15.17	7.7	7	10	5.7	18.51

Table 4.3: Thresholds for users experiencing interference losses

4.5 Probability computations

After running the simulations, the probability of an occurrence of cause and the conditional probability i.e. probability of a cause given the symptom are calculated. These probabilities are further utilized to calculate the detection accuracy of the self-healing framework.

Observing the movement of user in the cells, the number of frames for each user during which the user is nearby or at the boundary of neighbouring interfering cell, are calculated. The interference will only occur, when one or more than one user is at the boundary of the respective cell as well. The results are specified in table 4.4.

User No.	No. of frames	User No.	No. of frames	User No.	No. of frames
1	0	35	171	68	162
2	1015	36	0	69	0
3	227	37	42	70	164
4	235	38	202	71	186
5	17	39	0	72	163
6	0	40	94	73	0
7	0	41	212	74	0
8	106	42	1100	75	0
9	250	43	0	76	0
10	286	44	0	77	184
11	180	45	0	78	187
12	0	46	351	79	0
13	0	47	0	80	0
14	205	48	0	81	0
15	199	49	0	82	0
16	34	50	170	83	118
17	226	51	142	84	189
18	22	52	0	85	327
19	191	53	0	86	0
20	189	54	0	87	0
21	568	55	184	88	169
22	210	56	0	89	502
23	166	57	225	90	0
24	153	58	0	91	0
25	351	59	0	92	0
26	229	60	121	93	0
27	475	61	0	94	0
28	162	62	0	95	181
29	207	63	163	96	0
30	10	64	0	97	132
31	220	65	0	98	0
32	0	66	236	99	0
33	173	67	0	100	167
34	205				

Table 4.4: The number of frames experiencing interference losses for each user

Observing the movement of user in the cells, the number of frames for each user during which the user is in the area of shadowing, are calculated. The shadowing losses are independent of the interference losses, the results are specified in table 4.5.

User No.	No. of frames	User No.	No. of frames
23	145	73	43
37	115	77	146
38	144	78	169
48	163	83	149
58	173	88	98
65	143	95	169
70	11	97	162

Table 4.5: The number of frames experiencing shadowing losses for each user

The correct cause probability is the probability of a cause being detected correctly, and is denoted by $P_{corr}(Cause)$. Assume that the interference cause is denoted by C_1 , and the coverage cause is denoted by C_2. The detection accuracy is the percentage of all the cases correctly detected.

From these results in table 4.4 and table 4.5, some conclusions can be drawn for each cause.

- For the interference cause, from a total 100 users, 56 users have reached the interfering cell boundary and then have bounced back. The rest users have also reached the boundary, but they have reached the other two boundaries, as each cell has only two interfering boundaries with the adjacent cells.

  Therefore, $P\ (C_1) = \frac{56}{100} = 0.56$ for all users

  From these 56 users, only 20 users have experienced interference losses, therefore the correct cause probability is, $P_{corr}(C_1) = \frac{20}{56} = 0.35$ for all users

- For the coverage cause, from a total 100 users, 15 users have been through the area with obstacles.

  Therefore, $P\ (C_2) = \frac{14}{100} = 0.14$ for all users

  From these 14 users, all only 13 users have experienced shadowing losses, therefore the correct cause probability is, $P_{corr}(C_2) = \frac{13}{14} = 0.93$ for all users

- For both the causes, from a total 100 users, 40 users had been detected by the self-healing algorithm. When the each user was checked for verification purposes, then 33 users had been faced with causes.

  Therefore, $Detection\ accuracy = \frac{33}{40} = 82.5\ \%$

The simulations have been re-run and the results are summarized in table 4.6 below.

Simulation no.	$P\ (C_1)$	$P_{corr}(C_1)$	$P\ (C_2)$	$P_{corr}(C_2)$	$Detection\ accuracy$
1	$\frac{56}{100} = 0.56$	$\frac{20}{56} = 0.35$	$\frac{14}{100} = 0.14$	$\frac{13}{14} = 0.93$	$\frac{33}{40} = 82.5\ \%$
2	$\frac{68}{100} = 0.68$	$\frac{18}{68} = 0.27$	$\frac{14}{100} = 0.14$	$\frac{9}{14} = 0.64$	$\frac{27}{44} = 61.3\ \%$
3	$\frac{63}{100} = 0.63$	$\frac{16}{63} = 0.25$	$\frac{23}{100} = 0.23$	$\frac{13}{23} = 0.56$	$\frac{28}{52} = 53.8\ \%$

Table 4.6: Probability of cause and detection accuracy results

References for Chapter 4

[1]. Moreno, R. B. (2007). *Bayesian modelling of fault diagnosis in mobile communication networks.* Higher Technical School of Telecommunication Engineering. University of Malaga, Spain.

Chapter 5 . Conclusion and Future Work

5.1 Conclusion

In this thesis, a self-healing framework has been implemented, which detects the faults occurring in the network causing degradation to the overall network performance. Self-healing mechanisms are a part of the self-x cycle for heterogeneous wireless networks. The self-healing framework has been chosen to work with LTE technology, the dropped number of calls have been set as an objective to heal. The possible causes for the problem have been selected: interference and coverage, both of the causes are consequences of turning off the PC at both pathways i.e. uplink and downlink. The shadowing losses are subject to the addition of obstacles in the area and the interference is experienced from adjacent cells. All the simulations have been performed in Matlab.

There has been ongoing research in the self-healing area, the relevant work and architecture have been studied. In order to determine a threshold for a symptom, the threshold determination techniques have been studied and beta pdf method (BMAP) has been implemented. After the thresholds are determined, they are stored in the database, which are used for the detection purposes in the future simulations. When the simulation is re-run, then the database is checked for all the users. If the users have already obtained a threshold, the user is checked for threshold with the symptom, if the current symptom goes below the threshold level, an alarm is raised. Since only detection of the faults occurred have been covered, therefore the faults can’t be diagnosed or rectified. In order to rectify them, the fault diagnosis techniques will be applied.

The faults occurring are modeled using Bayesian modeling (Bayesian Networks modeling). Bayesian Networks is among the popular approaches for fault modeling. Based upon the simulations, training sets of data are collected and the probabilities are computed for each user, which help in determining the detection accuracy of the self-healing framework.

5.2 Future Work

The self-healing framework implemented in this project is limited only up to the detection of symptoms; it does not do any diagnosis the fault or symptom that has occurred in the network. Therefore, in the future work is needed in the area of fault diagnosis. The different methods of determining threshold for a symptom as well as different methods for fault modelling have been studied, but only one method from each has been implemented. Therefore the project can be guided in the direction of implementing all the studied methods and compare their performance metrics in the future.

Furthermore, in this thesis only two scenarios and symptoms have been considered namely interference and coverage. The different scenarios such as throughput can be added to test the self-healing framework. The antennas in the matlab simulation are omni-directional that radiates equal power in all the directions, where as practically in the cellular network directional sector antennas are mounted on the towers. Therefore, the antenna may be changed to directional antenna, so that more realistic situations can be considered and tested. Moreover, the directional antenna can be down tilted to create coverage problems, and can be added as symptom to the self-healing framework.

List of Acronyms

3G – 3rd Generation
3GPP – Third Generation Partnership Programme
4G – 4th Generation
ABR – Average Bit Rate
AV_UE_SINR_DL – Average User Equipment Signal to Noise and Interference Ratio in Downlink
AV_UE_SINR_UL – Average User Equipment Signal to Noise and Interference Ratio in Uplink
BCR – Block Call Rate
BS – Base Station
BMAP – Beta Maximum a Posteriori
CAPEX – Capital Expenditure
CAC – Call Admission Control
CBR – Case Based Reasoning
C/I – Carrier to Interference
DFTS-OFDMA – Digital Fourier Transform Spread Orthogonal Frequency Division Multiple Access
DL – Downlink
EMD – Entropy Minimization Discretization
eNB – e Node B
GGSN – Gateway GPRS Support Node
GPRS – General Packet Radio Service
GSM – Global System for Mobile communication
GW – Gateway
KA – Knowledge Acquisition
KAT – Knowledge Acquisition Tool
KPI – Key Performance Indicator
LA – Link Adaption
LR – Logistic Regression
LTE – Long Term Evolution
MCS – Modulation and Coding Scheme
MDL – Minimum Description Length
MEST – M-Estimate
MLE – Maximum Likelihood Estimation
MME – Mobile Management Entity
MS – Mobile Station
MU-MIMO – Multi User Multiple Input Multiple Output
O & M – Operation & Maintenance
OBF – Over Booking Factor
OFDMA – Orthogonal Frequency Division Multiple Access
OPEX – Operational Expenditure
PC – Power Control
PDF – Probability Density Function
PRB – Physical Resource Block
PSD – Power Spectral Density
QAM – Quadrature Amplitude Modulation

QPSK – Quadrature Phase Shift Keying
QoS – Quality of Service
RRM – Radio Resource Management
RSS – Really Simple Syndication
RSSI – Received Signal Strength Indicator
SAE – System Evolution Architecture
SC-FDMA – Single Carrier Frequency Division Multiple Access
SEMD – Selective Entropy Minimization Discretization
SGSN – Serving GPRS Serving Node
SINR – Signal to Noise and Interference Ratio
SLA – Service License Agreement
SLAH – Statistical Learning Automated Healing
SNR – Signal to Noise Ratio
SON – Self Organizing Network
UE – User Equipment
UL – Uplink
UMTS – Universal Mobile Telecommunication System
UTRA – Urban Terrestrial Radio Access
VoIP – Voice over Internet Protocol
TTI – Transmission Time Interval