HELSINKI UNIVERSITY OF TECHNOLOGY
Department of Electrical and Communications Engineering
Communications Laboratory

Niko Isotalo

Proactive fault management in mobile core network

Master's Thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Technology.

Espoo, February 19, 2008

Supervisor: Docent Timo Korhonen

Instructor: Ira Antikainen, FM

HELSINKI UNIVERSITY OF TECHNOLOGY

ABSTRACT OF THE MASTER'S THESIS

Author:	Niko Isotalo	
Name of the Thesis:	Proactive fault management in mobile core network	
Date:	November 12, 2007	**Number of pages:** 74 + 15
Department:	Department of Electrical and Communications Engineering	
Professorship:	S-72 Laboratory of Communications	
Supervisor:	Docent Timo Korhonen	
Instructor:	FM Ira Antikainen	

Fault management is one of the most important aspect of mobile network management. As the networks are coming larger and more complex single fault can affect to the large part of the network.

The purpose of this study is to improve fault management, ease up network operators' everyday maintenance activities and make root cause analysis of more difficult problems. Root cause analysis is about finding out the root cause of the problem. This analysis is done by the network vendor.

This is achieved by designing an equipment which will audit customer network elements and execute there predefined tasks. If a problem is found from the network element, one should collect required logs in order to perform root cause analysis as soon as possible.

Design process of this equipment is gone through in this study. The process starts by defining the tasks that should be done in order to improve fault management and network maintenance. Second step in the design process is to plan hardware and software configurations. This was made with SWOT-analysis and Force Field Analysis models. The third step was to design user interface. Last step in the design process was to test user interface with the user centred design methods.

This thesis concludes that by using necessary tools and databases about known network problems efficiently, a lot of time can be saved in the support organization and in the network operators' side. Following this strategy network operators' will have the support they need. Also time spent collecting the logs and information will be smaller. Using this equipment results to more healthy and stable network.

Keywords: Error detection, Core network,mobile communications systems, fault management, proactive maintenance, usability, user centric design

TEKNILLINEN KORKEAKOULU DIPLOMITYÖN TIIVISTELMÄ

Tekijä:	Niko Isotalo
Työn nimi:	Ennaltaehkäisevä vianhallinta matkapuhelinverkkojen runkoverkossa
Päivämäärä:	1.5.2016 **Sivuja:** 74 + 15
Osasto:	Sähkö- ja tietoliikennetekniikan osasto
Professuuri:	S-72 Tietoliikennetekniikka
Työn valvoja:	Dosentti Timo Korhonen
Työn ohjaaja:	FM Ira Antikainen

Virheiden hallinta ja paikallistaminen on yksi tärkeimmistä asioista matkapuhelinverkkojen hallinnassa. Kun matkapuhelinverkoista on tullut laajempia ja monimutkaisempia yksi onglma voi vaikuttaa verkon toimintaan suurella alueella.

Työn tarkoituksena on parantaa vikojen paikallistamismenetelmiä matkapuhelinverkkojen runkoverkossa, helpottaa verkko-operaattoreiden päivittäisiä verkonhallintatoimia ja tehdä vaikeimpien ongelmien ratkaisemisen nopeammaksi.

Tämä tavoite saavutetaan laitteella, joka tarkkailee asiakkaan verkkoelementtejä suorittamalla siellä ennaltamääriteltyjä komentoja ja tehtäviä. Jos ongelmia havaitaan, laite kerää tarvittavat logit, jotta virheen perussyyn etsiminen on mahdollista.

Tässä tutkimuksessa kyseinen laite suunnitellaan. Laiteen suunnitteluprosessi aloitetaan kartoittamalla tehtävät, jotka laitteen pitää tehdä, jotta virheiden- ja verkonhallinta paranee. Toisessa vaiheessa laitteisto ja ohjelmistovaihtoehtoja analysoidaan käyttämällä SWOT- ja Force Field analyysiä. Tämän jälkeen suunnitellaan ja testataan laitteen käyttöliittymä.

Työn johtopäätöksissä todetaan, että käyttämällä olemassa olevia työkaluja ja tietokantoja tehokkaasti, aikaa ja rahaa voidaan säästää tukiorganisaatioissa. Käyttämällä näitä työkaluja oikein mahdollistetaan, että tarvittava tuki asiakkaan sitä tarvitessa saadaan heille mahdollisimman nopeasti. Tätä laitetta käyttämällä verkko-operaattorin verkko on paremmassa kunnossa, mikä lisää myös verkon toimintavarmuttaa.

Avainsanat: Virheiden etsintä, runkoverkko, matkapuhelinjärjestelmät, virheiden hallinta, ennaltaehkäisevä kunnossapito, käytettävyys, käyttäjäkeskeinen suunnittelu

Acknowledgements

This Master's thesis has been the most difficult journey in my life. It toke much more effort and time than I had expected. Now that I know that it is behind it is huge relief both for mentally and workload wise. This thesis is a result of hard work and great support from many instances.

Thanks to Professor Timo Korhonen for supervising and giving help in finding information and getting started on the right track. Christopher Mosley for the opportunity to work and do this thesis in Global Delivery Center at Ericsson. For my instructor Ira Antikainen for her opinions during the the project. To my colleagues Markus, Mika, Mikko, Nagarajan and Sebastian for the ideas and help during the user interface development process and testing it.

Thanks to all my friends from the University with whom I have made project works, study for exams and spend time together. Special thanks to my family and to my girl friend Satu to their valuable support.

Finally, I would like to thank the most precious thing in the world and the main reason why this thesis was ever done, my little baby girl Aava, who with her happiness and joy helped me to proceed during the difficult moments.

Espoo, November 13, 2007

Niko Isotalo

List of key concepts

Core network

Core network carries out switching functions and manages the communications between mobile phones and the Public Switched Telephone Network. Core network consists several nodes like MSC, HLR and Media gateway.

Error detection

Finding out system malfunctioning

Fault

In the context of network management a fault is defined as a cause of malfunctioning. Faults are responsible for making it difficult of preventing the normal functioning of a system and they manifest as themselves through errors.

Fault management

Fault management is about detection, isolation and correction of faults

Proactive maintenance

Proactive maintenance is about reducing the errors and faults in the network and also to prepare for the situations where there is no fast or easy way to recover the network.

Usability

The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use

(ISO 9241-11)

User centric design

User Centered Design is about creating services that are appealing to users and that they find intuitive, easy to use and which adds value to the intended users.

Contents

Abbreviations

A	
AMPS	Advanced Mobile Phone Service
APG	Adjunct Processor Group
ARP	Auto Radio Phone
AuC	Authentication Centre
B	
BDC	Backup Domain Controller
BSC	Base Station Controller
BSS	Base Station Subsystem
BTS	Base Transceiver Station
C	
CEPT	Conference of Postal and Telecommunications Administrations
D	
DAG	Directed Acyclic Graph
E	
EIR	Equipment Identity Register
F	
FCC	Federal Communications Commission
FDD	Frequency Division Duplex
G	
GPRS	General Packet Radio Service
GSM	Global system for Mobile Communications

H

HCI	Human Computer interaction
HLR	Home Location Register
HSDPA	High Speed Downlink Packet Access
HSUPA	High Speed Uplink Packet Access

I

IMEI	International Mobile Equipment Identity
IMSI	International Mobile Subscriber Identity

M

MS	Mobile Station
MSC	Mobile Switching Centre

N

NMT	Nordisk Mobiltelefon

P

PDA	Personal Digital Assistants
PDC	Personal Digital Cellular
PDC	Primary Domain Controller
PLMN	Public Land Mobile Network
PSTN	Public Switched Telephone Network

Q

QoS	Quality of Service

R

RAB	Radio Access Bearer
RNC	Radio Network Controller
RNS	Radio Network Subsystem

S

SGSN	Serving GPRS Support Node
SMS	Short Message Service
SWOT	Strength, Weaknesses, Opportunities, Threats

T

TACS	Total Access Communication System
TRX	Transreceiver
U	
UCD	User Centered Design
UE	User Equipment
UMTS	Universal Mobile Telecommunications System
V	
VLR	Visitor Location Register
W	
WCDMA	Wideband Code Division Multiple Access

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

Field of telecommunications is changing all the time. New techniques and technologies arrive to the market and new services are developed. This leads to the situation where the end customer needs more and more training and support activities as the new products are usually more complex than the old ones, not maybe to use but there is much more new functionalities at least. This Master's thesis has been written in the network vendor point of view which means that the customer, which is mentioned through the thesis is a network operator.
Importance of customer support is increasing all the time. When the competition is tough, every field have to be in good shape in order to be proficient or to be the market leader.
Nowadays there exists many types of support activities. Some customers want that their whole network is operated and maintained by the network vendor, some customers want to have full maintenance from the vendor for example to the some parts of the network for example to the core network side only and some customers just want support and troubleshooting activities when they have a problem they can't solve by themselves.
Support has been very important business and it still is. A problem in the support activities is that they usually help only when the fault has already occurred or alarm raised and the focus is not in the proactive error detection and correction side. Wouldn't it be great from the customer point of view if the network vendor could not only fix the problems but also efficiently predict and fix errors before those even cause any visible problems. This could be considered to be even more valuable service than the traditional support activities (support that focuses on problems that already exists). Of course we can't forget the traditional support because it is impossible to predict all the faults that might happen.

1.2 Objectives

The purpose of this Master's thesis is to develop a fault diagnostic equipment which will audit the mobile networks' core nodes and that way find out faults or error behaviors that are known by the network vendor. This equipment will also help the network operator in their everyday maintenance activities and network vendor when problem can not be solved by the network operator. This same idea can be applied to the radio network side as well.
In the support point of view the main purpose of this study is to reduce high and emergency priority customer service requests, reduce the time needed for the recovery actions when emergency priority customer service request arrives. In the software supply point of view the main purpose of this equipment is to ensure that software deliveries can be delivered on time and network rollout (software upgrade on whole customer network)schedules can be planned more exactly, because nodes are in better shape.
At the same time when the amount of high and emergency priority customer service requests will decrease, it can be assumed that the amount of medium and low priority customer services requests will increase. This leads to a situation where engineers can spend more time on a root cause analysis of the problems and smaller configuration problems and less time for the problems that needs immediate recovery actions like complete loss of charging or even traffic. Even though amount of lower priority customer service requests increases it can be expected that total amount of support request will decrease.
From the customer point of view this new equipment means more stable and reliable network, with relatively less amount of maintenance work than before.

Here are four the key challenges in this thesis:

1. Error detection of the equipment should be effective
   Equipment should be able to find out known issues that happens in the node, also updating the new problems should be done effectively.

2. What to do when potential error is detected ?
   When the equipment finds out the fault should it try to fix it by itself, escalate the problem to the customers engineers or escalate the problem directly to the network vendor.

3. Equipment should be available for as many customer as possible
   In ideal situation this equipment is available for every customer, but in general that can be very challenging, due to for example different end customer security policies.

4. Equipment should be easy to use and maintain
   This equipment shouldn't cause a headache to the customer. Usability of the equipment should be logical, which means that no manuals or help is needed and maintenance needs are as low as possible.

1.3 Structure of the Thesis

In this section we will go through the structure of this thesis. Chapters from two to five are based on theory and previous research results in the area of mobile communications, usability, computer security and error detection in computer network. These five chapters are also preparations for the chapter six, where the theory is applied when designing the equipment starts. The chapters six and partly chapter seven are authors own contribution to this thesis. Eight chapter is dedicated to the conclusions and summary.
In the second chapter we will go through the history of mobile communications systems from the 1947, when the cellular concept was first introduced till today. We will also get familiar with the most common mobile communications networks like GSM and UMTS. Network structures and functional parts of these two communications systems are introduced. This second chapter acts as a base to understand the communications systems and this thesis.
Third chapter is about usability. Main focus in this chapter is to emphasize the importance of usability. The reason why importance of usability is emphasized is that if usability fails or if it is not in a satisfied level, most probably launch of the product fails, even though product might be technically useful. This chapter start with the discussion about the usability in general. The second section of this chapter is dedicated to the importance of usability. In the Third section we will go through the key principles of the user centric design. In the last section we will investigate the ways how we can measure usability.
In the fourth chapter focus will be in security. Security requirements and how to build good security is also described. How to estimate risks the system is suffering and what are the threats of today's communications system. In the end of the fourth section we will go through the reasons why security fails and what mistakes can be done when estimating the risks.
Fifth chapter is about error detection in mobile communications network. This section starts with concept definitions. After that we will go through several fault management methods. In the last section of this chapter proactive maintenance is investigated.
In the Sixth chapter we will start planning the equipment itself. In the first section system/equipment specifications and functionalities will be decided, which network nodes it is applicable to is also gone through. In the second subsection of this chapter hardware and software requirements for the equipment will be decided. To be as through as possible, all the possibilities will be compared and strength and weaknesses, opportunities and threats will be considered by following the SWOT model. When all the possible options have been taken into account and the best one for this purpose is chosen, system security needs to be thought. This is done in the third section of this chapter, In this section Security policies and goas of the system are defined, when this is done building the security will start according to the theory described in the chapter 4. In this section threat scenario and most probable threats are gone through. After creating the threat model, next thing is to start planning

the user interface. This is done in the fourth section. User interface is designed according to the principles defined in the chapter 3, section user centred design, this can be considered as the most important part of this thesis. When user interface has been designed, its usability has to be tested. This is done in the chapter 6. If there is something in the user interface that needs to be improved it will be done. Seventh section will go through the further development possibilities of the equipment. How the error detection can be done more effectively and more accurately. Also updating the know fault is important and it is considered as one development phase. Interesting new way to do the proactive error detection is also introduced. Eight and last chapter is about the results and conclusions. How this equipment affects to the network vendor and how it affects to the network operator

Chapter 2

Mobile Communications Systems

2.1 Introduction

In this chapter we will go through mobile communications systems from the early days till future. Some telecommunications systems are just mentioned and the most important ones will be described in more detail. The purpose of this chapter is to get readers familiar with the mobile telecommunications systems in general, to know some historical background and also to know detailed core- and radio access network structures of GSM and UMTS networks, which will be needed further in the thesis.
This chapter will start with a brief history review from the mid 20th century till today, from simple cellular communications network to complex 3rd generation mobile communications systems. After the review the most common mobile communications systems are introduced. In the next two section we will go through both GSM and UMTS core and radio network structures. In the UMTS section, also differences and new features between different releases are mentioned, but not gone through deeply. This chapter gives a good base level to understand the rest of this thesis for those who are not that familiar with the mobile communications systems and a good revision for those who are familiar with the mobile communications, but still need to refresh their memory.
After this chapter readers should have basic knowledge about history of mobile telecommunications. Detailed knowledge about mobile telecommunications systems their network architecture and basic functionality of the network that are used in today's mobile communications.

2.2 History of cellular mobile communications systems

2.2.1 Early days of cellular networks

The idea of the first cellular network was made in December 1947 by Bell Labs engineer D H Ring[16]. It was also proposed by the other Bell Labs engineer Phil Porter, that the cell towers will be at the corners of the hexagons rather than in the centers and directional antennas would transmit and receive signal in three directions. Only problem to construct this kind of system was that there was no technology available until the 1960s.
In 1967 mobile industry had gone so far that it was possible to make and receive calls from the mobile cellular network, but each mobile phone had to stay within the cell area serviced by one base station throughout the phone call. In 1970 Bell Labs invented a system to allow mobile phones to move to another cell area during a single call without loss of conversation. This is called handoff or handover.
In December 1971 AT&T submitted a proposal for cellular service to the Federal Communications Commission (FCC). FCC approved the proposal in 1982 for Advanced Mobile Phone Service (AMPS) and allocated frequencies in the 824-894 MHz band.
At the same time in year 1971 the first successful public cellular mobile phone network was opened in Finland. It was called ARP (Auto Radio Phone). The ARP phones operated at 145 MHz band, and many consider it as a 0G cellular network.

2.2.2 Towards automatic mobile communications

In the 1970s development of first full automatic mobile communications system was started. It was called Nordic Mobile Telephone system (NMT) and it was developed by Ericsson. It used 450 MHz frequency band in the beginning and later on 900 MHz frequency band was used, because of higher capacity. First commercial NMT network was launched in the first of October 1981 in Saudi-Arabia. New features in NMT was that roaming was possible, also mobile equipments were smaller. NMT was considered as a first generation mobile communications system. Other important 1G systems were AMPS, which was in use in north America and TACS, used in Japan.

2.2.3 Digital mobile communications systems

In 1982 European Conference of Postal and Telecommunications Administrations (CEPT) planned to create Groupe Spécial Mobile with the objective of developing a standard for a mobile telephone system that could be used across Europe. In 1987 European operators agreed to launch first live GSM network at first of July 1991. In the same meeting original French name was changed to Global System for Mobile communications. In 1989 standardization responsibility was transferred to European Telecommunications Standards Institute. In the result of this standardization work,

in 1991 the first GSM network was launched in Finland as planned. GSM had digital signaling and speech channels and it was considered as a first 2G mobile communications system. In 1994 also digital AMPS and PDC (used in Japan) were launched.

2.2.4 3G mobile communications systems

Mobile data communications is evolving quickly because of Internet, Intranet, Laptops, PDAs and increased requirements of workforce mobility. To satisfied this demand mobile communications system with higher data transfer capability was needed. 3G mobile networks were answer for this. First 3G specification was ready December 99 and it was called R99. It was also planned that the first 3G network will be launched 1st of January 2001. The first commercial 3G network was launched by NTT DoCoMo in June 2001 in Japan. The main difference to GSM was higher data rates (384kbit/s). Next big improvement in 3G network came in Release 5, standardization took longer than expected and was ready in March 2002. In Release 5 HSDPA was introduced. It increased the downlink data rate in first phase to 1 Mbit/s and in the future up to 14 Mbits/s. At the moment over 100 HSDPA networks have been commercially launched. The standardization of release 6 was ready in December 2004 and Release 6 standard HSUPA was introduced. HSUPA makes the same to the uplink that HSDPA did for downlink. Live implementation of HSUPA is ongoing at the moment.

2.3 Global system for Mobile Communications

GSM (Global System for Mobile communications) is an open, digital cellular and most popular technology used for transmitting mobile voice and data services. GSM service is used by over 2,6 billion people across more than 212 countries and territories which is over 30% worldwide penetration. The ubiquity of the GSM standard makes international roaming very common between mobile phone operators, enabling subscribers to use their phones in many parts of the world.
GSM differs from first generation wireless systems in that it uses digital technology and time division multiple access transmission methods. Digital technology means that both signaling and speech channels are digital, that is why GSM is considered as a second generation (2G) mobile phone system.
GSM is a circuit-switched system that divides each 200kHz channel into eight 25kHz time-slots. GSM operates in the 900MHz and 1.8GHz bands. GSM supports data speeds of up to 9,6 kbit/s, allowing of basic data services such as SMS. Another major benefit is its international roaming capability. It means that you can use your mobile equipment as you where at home when you are abroad. GSM network architecture is described in figure 2.1

2.3.1 GSM core network

GSM core network or network switching system carries out switching functions and manages the communications between mobile phones and the Public Switched Telephone Network (PSTN). GSM core network consists several network nodes, with different tasks to make mobility possible. There are many services possible in the basic GSM network, for example voice, data and fax services. In the basic GSM network all the traffic is circuit switched. There is also an overlay architecture on the GSM core network to provide packet-switched data services and it is known as the GPRS core network. GPRS nodes are not described in this document. In the next five subsections GSM network nodes and their functionalities are described.

Mobile Switching Center

The Mobile-services Switching Center is an exchange which performs all the switching and signalling functions for mobile stations located in a geographical area designated as the MSC area. Mobile switching center is the heart of the GSM network. The main difference between the MSC and an exchange in a fixed network is that the MSC has to take into account the impact of the allocation of radio resources and the mobile nature of the subscribers. MSC constitutes the interface between the radio system and the fixed network. MSC performs all the necessary functions in order to handle the calls to and from the mobile stations.

Home Location Register

The Home Location Register (HLR) is a data base in charge of the management of mobile subscribers. In the HLR following kinds of information is stored: subscription information, location information enabling the charging and routing of calls towards the MSC where the MS is registered (e.g. the MS Roaming Number, the VLR Number, the MSC Number, the Local MS Identity). A PLMN may contain one or several HLRs, it depends on the number of mobile subscribers, on the capacity of the equipment and on the organization of the network[1].

Visitor Location Register

The Visitor Location Register (VLR) is the location register, other than the HLR, used by an MSC to retrieve information for, e.g. handling of calls to or from a roaming mobile station currently located in its area. A mobile station roaming in an MSC area is controlled by the Visitor Location Register in charge of this area. When a Mobile Station enters a new location area it starts a registration procedure. The MSC in charge of that area notices this registration and transfers to the Visitor Location Register the identity of the location area where the MS is situated. If this MS is not yet registered, the VLR and the HLR exchange information to allow the proper handling of calls involving the MS. A VLR may be in charge of one or

several MSC areas. The VLR contains also the information needed to handle the calls set-up or received by the MSs registered in its data base[1].

Authentication Center

The Authentication Center (AuC) is an entity which stores data for each mobile subscriber to allow the International Mobile Subscriber Identity (IMSI) to be authenticated and to allow communication over the radio path between the mobile station and the network to be ciphered. The AuC transmits the data needed for authentication and ciphering via the HLR to the VLR and MSC which need to authenticate a mobile station.The Authentication Center (AuC) is associated with an HLR, and stores an identity key for each mobile subscriber registered with the associated HLR[1].

Equipment Identity Register

The Equipment Identity Register (EIR) is the logical entity, which is responsible for storing in the International Mobile Equipment Identities (IMEIs). The mobile equipment may be classified as "white listed", "grey listed" and "black listed" and therefore may be stored in three separate lists. An IMEI may also be unknown to the EIR. According to GSM standard EIR shall as a minimum contain a "white list" (Equipment classified as "white listed")[1].

2.3.2 GSM radio access network

GSM radio access network consists of two different network elements Base Station Controllers and Base Transceiver Stations, which together are referred to as Base Station Subsystem. Packet control unit is a late addition to the GSM standard and it is assumed to be integrated to the BSC and not be as a separate unit in the BSS. BSS' purpose is to manage the radio link between the mobile phones and core network. GSM Radio Access Network connects to the core network with two interfaces A and Gb.

The A interface is used to carry traffic channels and BSSAP user part of the SS7 stack between BSC and MSC. Gb interface connects the BSS to the SGSN in the GPRS core network.

Base Station Controller

The BSC is the functional entity within the GSM architecture that is responsible for RR (Radio Resource) allocation to a MS (Mobile Station), frequency administration and handover between BTS (Base Transceiver Station) controlled by the BSC. Typically BSC has 10s or even 100s of Base Transceiver Stations under its control. The BSC function may be physically located with the BTS.

Base Transceiver Station

In cellular system the Base Transceiver Station terminates the radio interface. It allows transmission of traffic and signaling across the air interface. The BTS includes the baseband processing, radio equipment, and the antenna. Each BTS may consist of a number of TRX (Transceiver), typically between 1 and 16. In the GSM system the BTS is also responsible for encrypting and decrypting communications with the Base Station Controller.

2.4 Universal Mobile Telecommunications System

Wideband code division multiple access is one of the main technologies for the implementation of third generation cellular system. It is based on radio access technique proposed by the ETSI Alpha group and the first release of the specification was finalized in 1999 and it is called R99. This specification defines the basic things that are needed. Frequency band in WCDMA networks was decided to be 1920Mhz

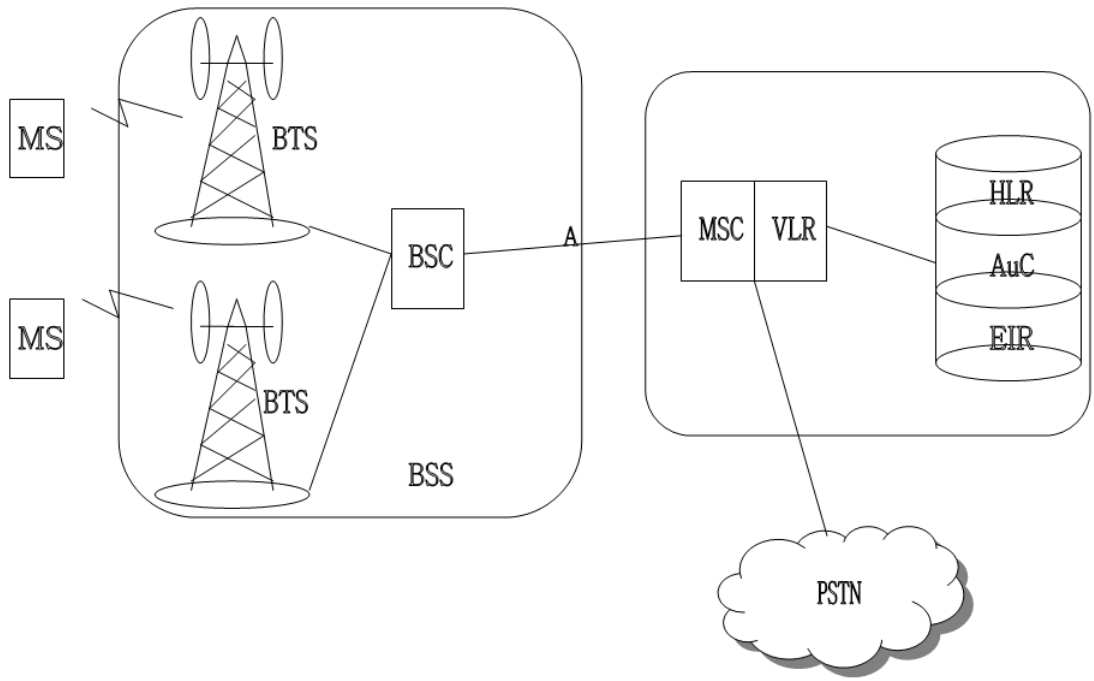


Figure 2.1: GSM network architecture [2]

- 1980Mhz for uplink, 2110Mhz - 2170Mhz for downlink and the used duplex method used is FDD. Frequency reuse factor is 1. The reuse factor can be one instead of being the same as in the GSM network because in WCDMA network can operate with only one frequency in the uplink and in the downlink because of the spreading codes which are different for different users and because of rake reception, where many receiver can receive the same signal in the different time delay. That is one thing why WCDMA technique is much more spectral efficient.

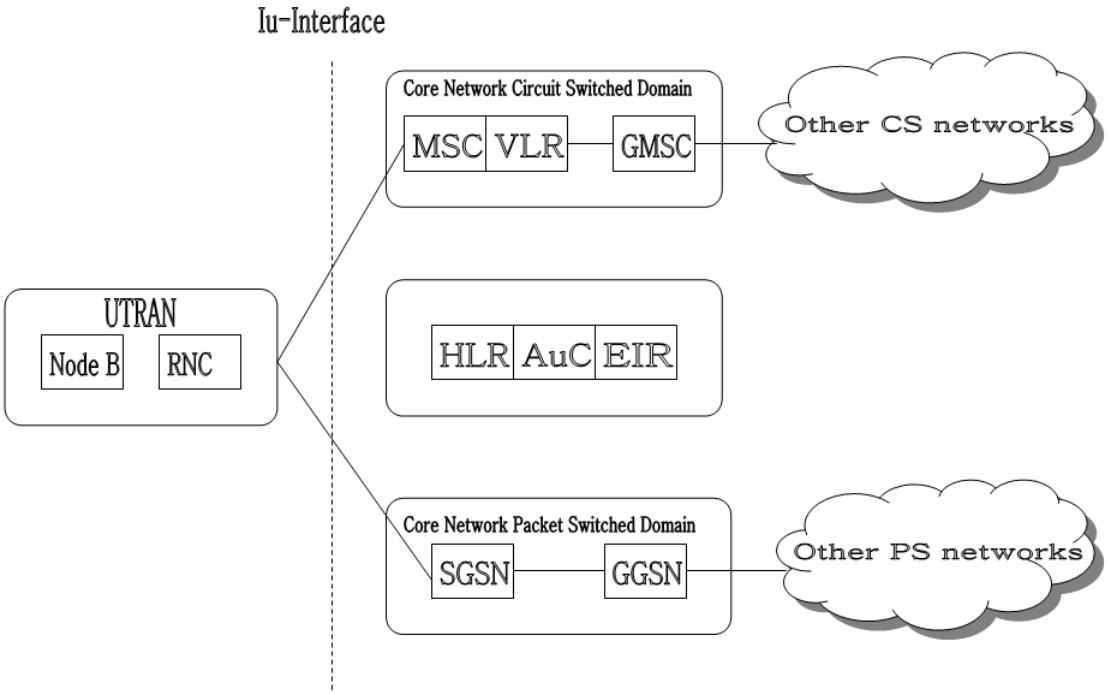


Figure 2.2: 3G R99 core network architecture.

2.4.1 UMTS core network

Core network consists of control nodes, where the logic of call and service setup, maintenance and realize is implemented, and the connectivity layer nodes that realize the physical connections. The Core Network domain elements are able to handle both 2G and 3G subscribers. The core network nodes in the 3GPP R99 figure (2.2) are the same as in the 2G network, but for example security mechanisms during connection set-up are different in 2G and 3G and their task is now to handel both 2G and 3G traffic.

The Iu interface connects WCDMA RAN to the Core Network. The Iu interface is divided into two instances, the Iu Circuit Switched (CS) to connect WCDMA RAN to the MSC server and the Iu Packet Switched (PS) to connect to the Serving GPRS support node (SGSN). The Iu interface is divided into a control plane and a user plane for handling signaling respective user data.

2.4.2 UMTS radio access network

The Radio Access Network is divided into a user plane and a control plane. The user plane is used for sending user data while the control plane is used for signaling required to establish, maintain and release a connection. A Radio Access Bearer (RAB) is the connection segment between the UE and the Core Network to support Quality of Service (QoS) for UMTS bearer services. Each of the radio access bearers are mapped onto one or more Radio Bearers. Each Radio Bearer is mapped onto one Radio Link Control entity. Each Radio Link Control entity communicates (UE-RNC) with its peer entity using one or more logical channels. The conversational speech Radio Access Bearer is decided to be 12.2 kbps Adaptive Multi Rate (AMR) speech and will also be used to carry emergency calls.

WCDMA RAN Architecture is described in figure 2.3. WCDMA system is a third generation network built on a horizontal layered structure consisting of the:

Service layer

Service Network is an IP-based network providing services to end users using WCDMA system resources.

Control layer and Connectivity layer

Certain interfaces have been defined and implemented to realize the connection between the nodes and the network parts. The Iub interface connects the Node B to the RNC. The Iur interface provides capability to support radio interface mobility between RNCs, of UEs having a connection with WCDMA RAN. This capability includes the support of handover, radio resource handling and synchronization between RNCs.

The interfaces used for Network management system

Mub is management interface towards Node B from Network management system. The Mub exposes all services in the service layer of the Node B and the IS, for the

purpose of accessing Node B information Mur. Management Interface towards RNC from Network management system. The Mur is the management interface provided by the RNC. It exposes all services in the service layer of an RNC and the IS, with the purpose of accessing RNS information.

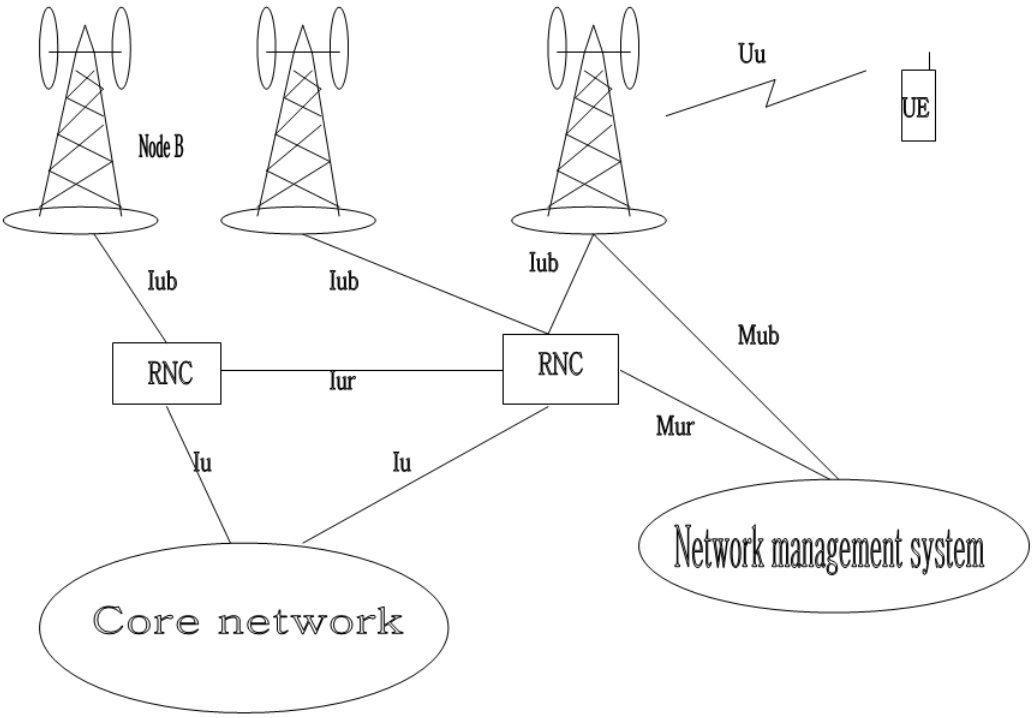


Figure 2.3: 3G R99 radio access network.

Chapter 3

Usability

3.1 Concept of usability

Usability is the quality of a product making it easy to operate and matching those more closely to user needs and requirements. International standard ISO 9241-11 defines usability as *The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use.*

Jakob Nielsen defines usability as a quality attribute that assesses how easy user interfaces are to use [13]. The word "usability" also refers to the methods for improving ease-of-use during the design process. Nielsens five quality components are:

1. Learnability: System should be easy to learn. When new users will start using the system, they would learn the logic of the system as fast as possible, so the real productive work could be started as early as possible.

2. Efficiency: It is an amount of resources and effort one has to use in order to be able to use the system. Once users have learned to use the system it should be fast to use and level of high productivity should be possible

3. Memorability: When users return to the system after a period of not using it, system should be created so that you don't have to learn everything all over again.

4. Errors: When new system is designed it should be done so that when a new user starts using the system, error rate should be as low as possible, if the user makes an error, user should be easily recovered from the error situations. Any major errors shouldn't happen.

5. Satisfaction: System should be pleasant to use, it means that users are satisfied to it and they actually like to work with the system

Usually in usability planning, the target of the usability study is a user interface either to some application or to the web page. It can be said that concept of usability is to help manufacturers to help consumers.
Usability shouldn't be confused with the functionality. Functionality is focusing to the products functions and features and doesn't speak out if the users are able to use the product or not. Better functionality doesn't mean better usability.

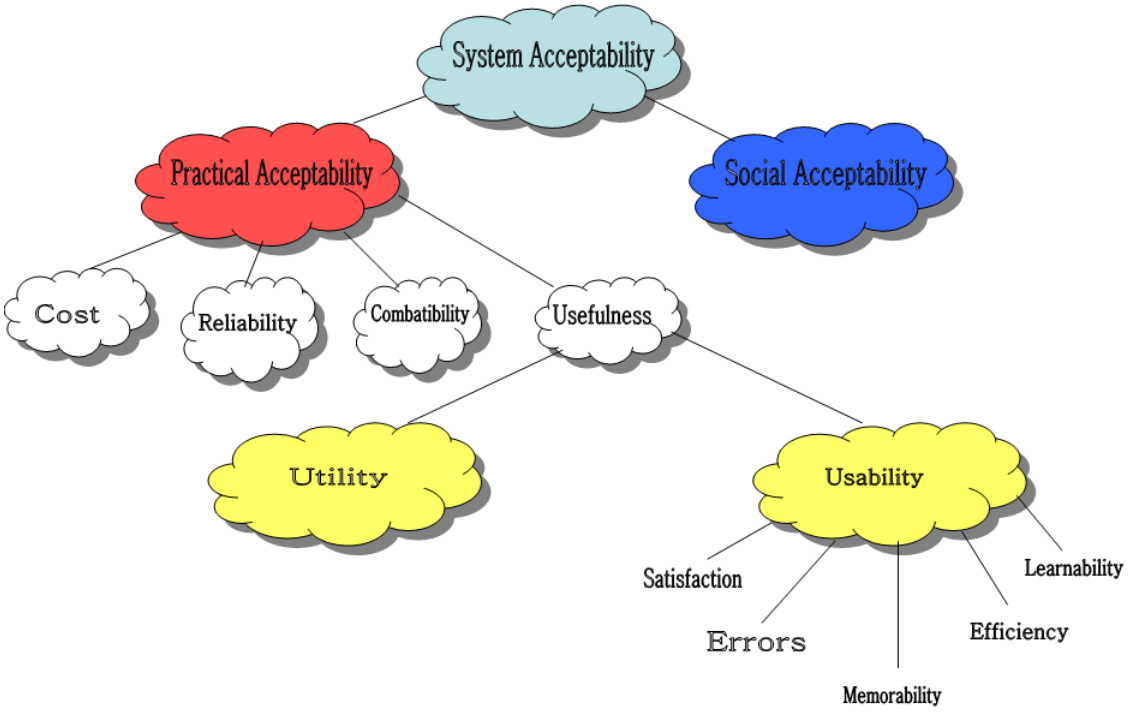


Figure 3.1: The context of usability in general[13]

3.2 Importance of usability

Usability is a very important part of the product development process. Limited usability is a problem that will endanger the technological progress of the product. Written solutions and manuals are not often a solution to the problem. Sometimes the space on a product is insufficient. Sometimes it is hard to decide on the language to be used ; and who really wants to study 200 pages of instructions for a rented car or tiny letters for a mobile phone usage? Many times even with the instructions consumers just do not succeed in getting their equipments into operation, even after trying for several hours. If the usability of a product is defective, even good instructions do not save the day. Because of the human nature many people struggle with the same problems and what is interesting is that many times they come to a same solution even though the solution might be completely wrong.

When we are talking about usability, users can be roughly divided to a four different groups. The importance of product usability varies between the user groups[9].

1. End users
2. Companies
3. Buyers
4. Society

For the end user importance of usability means attributes like easy to use, errorless to use, secure to use and fast to use. For the company usability is very important because it is a great sales argument, it is a part of a company's image and usable product is financially very profitable. If the company usability fails it might fail the whole company's income. For the buyers usability means better productivity and better personnel satisfactory. For society it also means better productivity, equal chances for the people and prevention of displacement.

3.2.1 Usability benefits

Usability of the products can create numerous benefits. The most important benefit from the company view is increase of sales and profit. Increased sales is a result of products better functionality and more accurate product specification, which makes the product more interesting from the customer point of view. The better profit is a sum of many different areas; Decreased R&D costs, reduced time spent on R&D, reduced maintenance costs, reduced human resource costs and reduced help desk costs.

3.3 User centric design

User Centered Design is about creating services that are appealing to users and that they find intuitive, easy to use and which adds value to the intended users. In order to do so, users have to be intimately involved with the design and thus an integral partner in the development process. Below there are nine important key principles of user centric design[6]:

1. Design for the users and their tasks
   Interactive computer systems do not operate in isolation, purpose of these systems is to support users to perform their tasks. A successful system meets business objectives through being user-centred and task-oriented. Always bear in mind the characteristics of the user population, their real-world tasks, and their working environment.

2. Involve users to the development process
   Representative users should actively participate in the entire development process from the early phase of system lifecycle till the end.

3. Be consistent
   To minimize the learning requirements for users, make the behavior of common interface elements and dialogue boxes as consistent as possible. This often means designing to be consistent with other existing components of the computer system. Although you may be able to design a new slick style of interaction, if it is inconsistent with the rest of the system, it will take users time and effort to learn it and get used to it.

4. Use simple and natural dialogue
   The dialogue between user and system should follow the natural sequence expected by the task. There should be no more information presented to the user than is necessary to complete the current task, because each item of irrelevant information adds complexity to the dialogue. Terminology should be defined so that the same term always has the same meaning.

5. Reduce unnecessary mental effort by the user
   Users should be able to concentrate on their task without worrying about the tool they are using. The more complicated the interaction with the tool, the more frustrated users become and the more distracted they are from their real task. If users have to invest too much mental effort in working out how to operate the system, they will be less efficient and make more errors. Simplify frequent tasks as much as possible. Users should not have to remember information from one part of the system in order to use another part. Instructions for the use of the system should be visible or clearly retrievable whenever appropriate.

6. Provide feedback to the user
   Users need to be confident that their actions have been successful. This is

usually evident from a distinct change in appearance of the window. If completion will take more than a second or so, a progress or working indicator should be displayed to give the user confidence that the computer is still operating. But avoid presenting the user with unnecessary or irrelevant diagnostic or status information about the internal state of the system. Feedback should be provided at several levels of interaction. At a low-level, users should receive confirmation that they have operated a control successfully – for example, a button immediately indicates when it has been operated by appearing momentarily pressed in. Users should also be informed when a longer sequence of operations has been completed.

7. Use prototypes to evaluate ideas
   Use multiple sketches and prototypes to support creative process, elicit requirements and visualize ideas and solutions. The prototypes should be evaluated with the real users.

8. Usability goals should control the development
   Pre-evaluated usability goals and specifications should be the main driver in the design process. Evaluate design against the goals in co-operation with the users. In the early phase, analyze users reactions to the prototypes.

9. User centric attitude
   User centric design requires user-centered attitude throughout the project team, development organization and the client organization. All people involved in the project must know the importance of usability and must be committed to work according to that.

3.4 Measuring usability

Usability measurements are performed by usability tests, which should cover aspects of the product for example, the interaction between the users and the graphical user interfaces, software and hardware. It is easy to specify usability metrics, but hard to collect them. Typically, usability is measured relative to users' performance on a given set of test tasks. In order to measure the usability of the product following information is needed[10]:

- The intended goals
- The context
- The expected results

3.4.1 Intended goals

To be able to improve usability it is important to specify the intended goals. The goal could, for example, be to achieve a certain task within a specified time with

some error rate. If usability measurements are to be performed, it is important that time is also allocated for redesign.

3.4.2 Context

The context is specified with reference to the users knowledge, education and experience. The context should also specify the surrounding as the equipment and tools that should be available.

3.4.3 Expected results

In order to determine any usability measurement, some sort of expected target of effectiveness, efficiency and satisfaction based on the goals and context must be defined. Examples of expected results are based on the answers to the following questions:

- Percentage of successfully completed tasks
- Time needed to complete task
- Number of errors made
- Users subjective satisfaction

Once you've gathered the results, you can use the numbers to formulate an overall conclusion about your design's usability, if it doesn't fill in the goals defined earlier, you have to do some redesign. According to Nielsen "It is enough if the testing is done to the five users and that the best results come from testing no more than 5 users" [12].

Chapter 4

Security

The growing number of malicious attacks against computer systems is forcing organizations to improve and rethink their security systems and policies. In many cases you can't even know who the enemy is, so preparing against the attack is even harder. Possible threats are for example that valuable data is exposed to unauthorized entities, invalid data is accepted as a genuine or purpose of attack is denial of service. One thing is sure, there is certainly no time at which security does not matter.

When organization wants to secure its systems, it must first determine what requirements to meet. The university needs to protect the integrity and confidentiality of the grades and data on its systems, hospital needs to protect the patient records from unauthorized entities and network operators need to secure the availability of the internet access they are providing to the customers. Some data integrity is so secret that it would rather be deleted than read by unauthorized people (top secret military information).

Even though organization has determined the security requirements, security may still fail. The reason is that threat scenario has been changed or there has been mistakes in the risk estimation for example designer has underestimated or overestimated the risks.

In this chapter we will first go through security requirements of a secure system using case studies. In the second section building security is described. Third section will evaluate security mechanisms which answers for example to following questions: "What assets are you trying to protect?, What are the risks to these assets?". In the last section we will go through the risk estimation, how to create a threat scenario, how the risk estimation can fail and why security fails. When planning the security you have to remember that the system is as secure as the weakest link of it.

4.1 Security requirements of a secure system

The security of a system is a combination of its ability to support data confidentiality, data integrity and system availability[5]. A failure of a system to protect any of these characteristics leads into a security violation or weakness.

4.1.1 Confidentiality

Confidentiality means that the information or resources stored on a system are protected against unintended or unauthorized access. Since systems are used to manage sensitive information, confidentiality is often a measure of the ability of the system to protect its data or resources. Confidentiality is achieved by access control. If it is the information (for example sensitive data or passwords) we try to protect, cryptography is the method to be used. If it is resource we try to protect, for example on site access to the main computer, physical locks are effective way to achieve confidentiality.

4.1.2 Integrity

Integrity means that the information stored on a system, resources or the source of the information is reliable and can be trusted. Since systems are used to manage information, Integrity is a measure of the quality of that information. To achieve integrity, prevention of unauthorized change of data has to be achieved.

4.1.3 Availability

System availability means how often a system is available for use by its intended users, since downtime is usually very costly. If the system is working 99.9% of time it means almost 9 hours outage (system is not available) time/year.
Availability is a system design principle usually against hardware or software failures. If system is wanted to be highly reliable there may be multiple independent software implementations running on different hardware that vote for right action. Possible attack would be overloading the server for certificate revocation lists which means that users cannot check for revoked certificates and would accept a compromised certificate.
Other cause of downtime is Denial of Service attacks. The purpose of this type of attack is to prevent authorized access to a system resource or delaying the system operations and functions. This leads to a situation that system is either unavailable (crashed) or unusable (response is slow).

4.2 Building system security

Building the security is important part of system design. In order to have good security many things have to be taken into consideration. When start building security one has to remember that security has three components: requirements, policy and mechanisms. Security requirements define security goals and they answers to the question "What do you expect security to do for you?" In order to find out the requirements, you should be familiar with the possible threats that could influence your system.
Security policy defines the meaning of the security and top management must be committed to follow the security policy. Without management support there cannot be real security. Policy answers to the question "What steps do you take in order to reach the goal set in the requirements?". In order to effectively follow the security policies, the ones that need to follow the policies should know why these policies are created.
Security mechanisms enforce policy and it will answer to the question "What are the procedures to ensure that the security policies and requirements are met?". The goal of security mechanisms is to ensure that the system never enters the disallowed state. Disallowed state is defined in the security requirements. Security mechanisms can be either technical or operational. For example some organization has a normal and secret documents. Users that are not allowed to read secret documents can not read those, because access control mechanisms prohibits it. This control is technical mechanism.
Though security policies may have been defined, real security requirements are heavily influenced by the threats that are expected for the given system. Informal approaches that have been shown useful in practice are based on risk analysis and threat identification. In this risk analysis, system and its environment are investigated in detail, in order to find out the possible attacks, their probability and possible loss in case of the attack, this way critical system components, where risks cannot be tolerated are found and security requirements can be met. [4]

4.3 Evaluation of security mechanism

As stated in the earlier section "Security mechanisms enforce policy" and it will answer to the question "What are the procedures to ensure that the earlier conditions are met?" In order to enforce security policy we have to know the security requirements as well. It depends on a security policy how much money company is willing to use to enforce the policy, what assets they are tried to protect or what are the risks of these assets. Security is about trade-offs. If you place a lock to you home, you have always risk that you will loose or forget the keys, but if you don't have lock in your door, it is sure that your property is not in safe. To effectively evaluate the security mechanisms following questions should be answered[14]:

1. What assets are you trying to protect?
2. What are the risks to these assets?
3. How well does the security solution reduce those risks?
4. What other risks does the security solution cause?
5. What costs and trade-off does the security solution impose?

4.4 Estimate risks

Risk estimation is a process of defining and analyzing dangers. Risk estimation is a cornerstone of information security. In general, purpose of the new equipment is to bring greater benefit than risks, if the risk estimation is not properly done, situation is certainly not like that. Risk estimation can be either quantitative or qualitative. In quantitative risk estimation, purpose is to numerically estimate probabilities of various attacks and amount of losses those will cause. When doing qualitative risk estimation numerical prediction of loss is not involved. Instead, it involves defining the threats, extent of vulnerabilities and countermeasures when attack occurs.
Common threats are easier to evaluate. Good examples about common threats are for example email viruses and internet worms. Issues that might not happen frequently or almost ever are much more difficult to evaluate. Threat is a potential way to subvert security and risk is probability and serious of a threat[14].

4.4.1 Threat Scenario

In order to have good security one has to create a threat scenario. It means that possible attacks can be prevented before those even happen. One of the most potential threat in the communications today is disruption. The purpose of disruption is to cause systems fail, cause long delays or even achieve denial of service of a certain server or service. Depending on a information that you are trying to protect, one very possible attack could be that someone is trying to have unauthorized access to the sensitive data. Usually in this kinds of attacks sensitive information is only copied and not deleted. Other possible treat scenario could be that someone unauthorized tries to have an access (username and password) to your company computers or servers by telling a lie about who he is and what are his purposes.

4.4.2 Mistakes in Risk Estimation

How much is enough security? Is the security being created too complicated? These are the questions that companies are facing when planning security and estimating risks. Sometimes the mistakes are done during the estimation process. One common

mistake what is done is that one underestimates the risk which has been identified. For example one thinks that having certain complexity in the passwords when logging to a company computers is useless and users can have any password they want. When the situation like this happens passwords that ones are using are something that is probably easy to remember and means something to the person, for example own birthday. Other serious mistake in the risk estimation is that the threat model is wrong. For example system is secured when the attack comes outside the company's network, but the biggest threat to the system is the users itself. Third common mistake that is done is that risk is overestimated. It basically means that some minor issues that is even very unlikely to even happen is considered as a major threat.

4.4.3 Why security fails?

When threat model is created certain assumptions about environment has to be done. When the time goes by also environment changes might cause security policy to fail. Environment changes are not that fast process so this shouldn't be very common threat, but there are examples about this as well [3]. Sometimes security can fail when the threat model is wrong, example about this was mentioned earlier in this section. When certain security policy is applied in a company, everyone should know purpose of that policy. When new security policy is applied it might feel nonsense and it is not followed if the meaning of that policy is not explained properly. Security should not be seen as a problem, but as a way of protecting information and property that has value to us.

Chapter 5

Fault management and proactive maintenance

5.1 Introduction

Greater demands and the increasing dependency of people in the mobile communications networks are the main driver of creating the better error detection mechanisms for mobile networks. Nowadays it is a matter of a course that you can use your mobile phone anytime almost everywhere. This makes fault management one of the most important aspect of network management.
Modern mobile communications networks may produce hundreds of alarms during one day. Fault situations can arise for example from hardware and software failures or from operational errors. Because of the number of the alarms, real time surveillance and fault management is difficult. When the large amount of alarms appear in a short period of time, it is very common that those alarms are overlooked or misinterpreted. One solution for this problem is that all the alarms are divided into classes so that for example A1 alarm is the most critical alarm class and A2 is the second class of alarms and so on. When the alarm is then raised it will tell the problem and also in which class it belongs. Of course this doesn't decrease the number of alarms but this way it is less likely to ignore the urgent alarms.
Other problem is that all the faults don't create an alarm. There can be several either configuration or other faults in the equipment and still there is no alarm available. These faults can be very critical and in the long run might even cause the system to crash or make your software upgrades or updates to fail. As you can imagine this problem is much harder to solve than the first one. One can think that this kinds of faults should have been found during the initial systems, software and inter operability testing. The most of the problems are, but not all. The reason for this is limited amount of time used in the testing and also the fact that these systems will be used in the several different network environments, with different vendors equipments connected to the nodes.

This chapter is divided into a three different sections, introduction and concept definition, fault management and proactive maintenance. In the beginning of the first section concepts fault and alarm are defined. After the definition we will get familiar with the fault management. Definition and methods of fault management are introduced. In the proactive maintenance section main focus are the methods how we can keep the systems in better shape and what preceding tasks should be done preparing oneself to a worst case.

5.1.1 Concept Definitions

Before going further in this section some basic concepts needs to be defined. After the definitions these concepts will be used throughout this thesis in this and further chapters.

Fault
: In the context of network management a fault is defined as a cause of malfunctioning. Faults are responsible for making it difficult of preventing the normal functioning of a system and they manifest as themselves through errors, that is, deviations relation to the normal operation of the system[11].

Alarm
: Alarms are specific types of notifications concerning detected faults or abnormal conditions, which may or may not represent an error. An alarm report is a kind of event report used in the transportation of alarm information[7].

Faults happen in the managed network or its components and alarm is an external indication of faults. Alarms that are defined by the vendors are observable by the network operator. In the ideal situation every fault (hardware, software or configuration) and abnormal situations happening in the network would cause an alarm, and the alarm text would indicate unambiguously where the problem is. Unfortunately this is not the case. There are many situations when there is a fault in the system, but no alarms raised or alarm doesn't tell much about the problem or where the problem actually is. There is a method for decreasing these kinds of problems and it is called alarm correlation. The purpose of the alarm correlation is that when the alarm will raise it would actually describe where the problem might be. Alarm correlation will also combine faults so that not every fault raise an new different alarm. Alarm correlation will be gone through in more detail in the following section.

5.2 Fault management

One of the most important areas in the telecommunications network management is the management of fault occurred during the normal operation of these networks. According to ITU-T network management can be grouped into a five different functional areas and fault management is one of those. In the ITU-T definition fault management is about detection, isolation and correction of faults[7].
The main requirement to perform fault management in an integrated way is the existence of information on the network's real time functioning. The abnormalities that occur during the operation of the network cause alarms, which are received by the network operation center. From that information people that are working in the operations center must attempt to identify and fix the fault occurred. If necessary, they raise a customer support request to the network vendor until the problem is solved.

5.2.1 Fault diagnosis

Fault diagnosis
: is a stage in the fault management process which consists of finding out the original cause for the received alarms. Before getting to the original cause it may be necessary to formulate a set of fault hypotheses, which will be needed to possibly reproduce and validate the problem.

Finding the root cause of a problem is essential for effective fault management. It is desirable that a system used for fault diagnosis has a model of the managed configuration, processes the flow of alarms in real time and is able to work on incomplete data. Added to this, system is expected to be able to identify changes in the appearance and in the importance of problems in function of time to separate cause from effects and to solve the problems according to their seriousness. The most difficult task in finding the root cause of the fault is to localize the fault into a right place in the system or in the large networks that are interacting with each others to the right system. When single fault happens it sometimes creates a several alarms. Alarm correlation may be used for network fault isolation and diagnosis, selecting corrective actions, proactive maintenance, and trend analysis, it will also improve telecommunications network surveillance and fault management. Alarm correlation plays essential role in localizing the faults into a right place. In the next subsection we will go through the basics about alarm correlation.

5.2.2 Alarm correlation

Alarm correlation is a conceptual interpretation of multiple alarms, leading to the attribution of new meaning of the original alarm. There exists several correlation

types. Those can be identified according to the operations executed on the alarms. In the below, most important ones are listed[11]:

Compression
: The purpose of a compression is to reduce multiple occurrences of an alarm into a single alarm and possibly indicating how many times the event occurred during the observation period.

Counting
: Counting consists of generating a new alarm each time the number of occurrences of a certain type of event surpasses a previously established threshold [11].

Suppression
: Suppression is a temporary inhibition of alarms to a given events according to predefined criteria which is linked to the presence of other alarms. For example when high priority alarm of given event is presence, low-priority alarm of same event is inhibited.

Scaling
: Scaling is an operation, in which alarm is canceled and another is created when new alarm has some parameter that has higher value than in the old one.

Filtering
: Filtering is about suppressing a given alarm, depending the values of previously specified parameters. Filtering takes into account only parameters of the alarm which is being filtered. There exists an intelligent filtering as well, this concept consists several other types of operations as well such as compression and suppression.

Generalization
: Generalization consists of replacing an alarm by the alarm corresponding to its super-class[1].

Specialization
: Specialization is an operation which is reverse of generalization, and consists of substituting an alarm for another corresponding its sub-class[2].

5.3 Proactive maintenance

Purpose of the proactive maintenance is to reduce the errors and faults in the network and also to prepare for the situations where there is no fast or easy way to fix the fault. With the task, that will be mentioned in this section you will reduce the possibility of the system to go down and also prepare for the worst case scenarios.

[1] superclass is a class from which other classes are derived

[2] The classes that are derived from a superclass are known as sub-classes.

5.3.1 Health check

In order to reduce the amount of faults in the network it is important that health checks of the network elements are done regularly, for example once a week. In below there is a definition of health check:

Health check
: Predefined set of commands to be executed in the system to find out if the system is functioning correctly or not.

Health check script
: Scripted command file which consists of commands to be executed in the health check

Commands that are executed in the health check are usually commands that reveal know problems which will not raise an alarm, but still will cause problems. You can't necessarily find out these problems in the normal operation, but when doing software upgrade or update problem might cause the upgrade or update to fail. New faults happen all the time and when the root causes are found out, proper checks are added to the health check script as well. In the list below most common problems that health check should check are listed.

Alarms
: The First thing what is usually done in the health checks is that it will check if there are any alarms in the system. If there are alarms health check script will write an entity to the result file.

Configuration errors
: Configuration errors are the most common problems that systems have. They cause problems during both normal operation and during upgrade. One example about a configuration error is that there might be some file in the system which has wrong access rights - authorized users have limited and unauthorized people have extended access rights.

Status of the Software level
: Current software level should be checked and it is compared to the baseline software. It should be also checked if there is any newer software level available.

Software comparison
: If the system has redundancy, softwares between the nodes are compared to verify that those are exactly the same.

5.3.2 Backup plans

In case there is a major fault which needs an immediate recovery actions, backups are very essential. Sometimes it is possible that there is such a major software fault

or fault that is very hard to localize or fix. In these kinds of situations only way for fast recovery is to use existing backups and restore the system from those. The backups should be taken often and should be stored to a place which is remotely accessible. The more important the device the more often the backups should be taken. After the backups are taken, verification of integrity of the data should be checked in order to be sure that backup is actually valid.
Other urgent issue that can happen is hardware fault. To be prepared to these kinds of situations, spare parts should be immediately available and already transferred on site.

Chapter 6

Designing the equipment

6.1 Introduction

In this thesis the main objective is to design an equipment which will audit the network nodes, its alarms and possible faults and also help customers maintain their network. Why this kind of equipment is designed? The main reason for this is, that there is certainly need for this kind of equipment in the market. The amount of customer service requests raised by the end customer that are related to the know issues or configuration faults is big. This creates a lot of work to fix these things. This equipment that is designed in this thesis will tackle to that problem.
In this chapter the equipment is designed. The first section gives some background why this equipment is designed. This section also introduces Ericsson network element called APG40. The system that we design in this chapter, is designed to help in maintenance of core network nodes. All examples that are shown in this thesis are based on the APG40 and this is the reason why it is introduced in this section.
In the second section functionalities of this equipment are defined. The second section also tackles to the first two key challenges defined in the introduction chapter.
The third section will focus on the software requirements and also some hardware requirements are gone through. Different functionalities and software possibilities are compared with the SWOT model to find out the most suitable for this purpose. Some of the software decisions are done based on the Force Field analysis as well.
Fourth section tackles to the equipment security challenges. This is one of the most difficult sections because different customers have different security requirements and security policies. In the introduction chapter of this thesis, third key challenge was, that this equipment should be available for every customer, it means that equipment security needs to be flexible but strict. Threat scenarios are also covered in this section and those are pretty much same for all customers.
In the fifth section design process of the user interface is described and also preliminary user interface for this equipment is introduced.
In the last section of this chapter usability aspects of this equipment are considered. In this section usability of the user interface is measured. After the measurements

results will be analyzed and if necessary user interface is improved. In this section we will also find out if the requirements of fourth key challenge is met.
After this chapter there should be well designed equipment waiting for the implementation. Predefined key challenges should have been taken into a consideration so that all the challenges have been answered in a proper way.

6.2 Background

The equipment that is developed in this thesis will monitor nodes in the mobile core network. In the first phase of this development process focus is on a developing mechanisms to the Ericsson core network equipments. This same idea is also applicable to the other vendors network elements and Ericsson radio network elements like BSC and RNC. Because of limited professional knowledge about the other core network elements all the examples are from the APG40.

6.2.1 Overview to the APG40

APG40 is an IO-device which acts as a user interface to the MSC, BSC and HLR. Main task for the APG is to act as a device which collects data from the MSC, BSC and HLR and processes it. After processing the data it is sent forward, for example to the billing gateway to be analyzed. APG also acts as a temporary storage place for the MSC, BSC and HLR backups. APG is a Microsoft Windows based machine with Windows NT or Windows server 2003 operating system in it. In order to get every functionality within APG40 working, it is not enough just to install operation system. On the top of the OS, Ericsson's own software is installed. This software takes care of collection and transformation of the data, where it should end up. Backing up and restoration of the system is handled by the Ericsson software as well as many other tasks.
There are three different hardware versions available for this device, but the basic functionality is the same in each of those hardware's.
APG40 consists of two computers (called nodes), node A and node B. Both nodes are identical and are acting as domain controllers[1]. These two computers form a cluster where the other node is acting as an active node and the other node as a passive node. The active node handles all the work and the passive node is listening. When active node reboots or goes down passive node takes the active nodes' role. It means that the nodes are redundant.
Both nodes have their own system disks where the operating system files are. Both nodes also have three data disks, where data is mirrored between the nodes. The data disk type is RAID-1. This means that every byte of the data written in the active nodes data disk is written also to the passive nodes disks. The reason for this

[1]In Windows NT domain controllers are called as a primary or backup domain controllers, PDC or BDC. In Windows server 2003 those are just domain controllers

kind of implementation is that it is more reliable and if the other node goes down no billing or any other data is lost.

6.3 System Functionalities

This section describes the functionalities of this system. Before we start we will define name for the system to be designed. From now on we will call this system as a PNeMaS, which comes from the words Proactive Network Management System. In this section functionalities of PNeMaS are introduced. The functionalities are designed so that challenges "Error detection of the equipment should be effective and What to do when possible error is detected?" from the introduction chapter are answered and met. Effective error detection in this thesis means that all the know faults and configuration errors are found. The second question is answered in the Summarize subsection.
The purpose of this section is to create set of functionalities which will be an integral entity. These functionalities will take into consideration every part of the network element in order to keep it as error free as possible. In case serious error happens you have up to date backups available, from where you can restore the system. In the below there is a list of PNeMaS's functionalities.

- Scheduled health checks
- Scheduled backups
- Software comparison
- Software auditor
- Alarm watch
- Summary

Purposes of these functionalities is to create benefit both for the end customer and for the network vendor. First two functionalities are mainly designed for the end customer and the purpose of these two functionalities is to help customer to fix know problems and configuration errors and also keep the nodes backups up to date. The third and fourth are functionalities that are mainly designed for the network vendor in case the end customer is not able to solve the problem. The last two properties are designed for the both, for the customer and for the network vendor. More detailed description about the functionalities are gone through in the next few subsections.

6.3.1 Scheduled Health check

Purpose of the scheduled health check is to keep the network elements in the better shape.Health check is done by running the predefined set of commands. This set

of commands consists of know problems, faults or configuration errors that do not necessarily raise any alarm, but are effecting or have chance to effect nodes normal behavior. Some examples are described later in this section.
Script itself does not fix any faults, it only tells what should be done and how. The responsibility to execute those fixes is left to the end customer. The reason for this is that some of the problems does not necessarily cause any harm now or those has to be fixed during maintenance window.
The health check script is run on both APG nodes at once. After the health check script is run on both nodes, we have two separate files one from the each node. These files will be then analyzed. This procedure results a file which includes the known problems found from the nodes, detailed information how the particular problems can be solved and the severity of the problem. If the solution of the problem is very sensitive and needs an expert help or fixing the problem should be done during maintenance window it is mentioned in the result file.
In this system we are designing, customer can select the time when he wants to run the health check. Default time period is once a week. When there are any updates to the health check script, system will automatically download updated versions of the script from the network vendors' server.

Examples about checks and solutions

Problem 1:

Severity: medium
The following service(s) has the wrong startup state
1. Cluster server

Solution 1:
Change startup type to automatic with following command
sc config clussvc start= auto

Problem 2:

Severity: medium
DCHP database could be corrupted

Solution 2:

Fix corruption with following commands:
dhcpcmd.exe 127.0.0.1 checkdb 192.168.169.0 fix
dhcpcmd.exe 127.0.0.1 checkdb 192.168.170.0 fix

6.3.2 Regular backups

The reason why this functionality is added to this system is that there are several situations per year when valid backups of the system are needed but are not available. In order to be sure that there always exists valid and up to date backups, those should be taken on a regular basis. In the PNeMaS customer can schedule regular backups to be done. Recommended time interval is every two weeks. Of course just before update or upgrade backups should be taken.
After PNeMaS has taken the backups it will verify if the backups are usable and transfer those to the predefined FTP area. The one reason why those backups are transferred out from the node is that for example if the APG is not able to boot, you still have backups and you can do Disaster recover to recover the IO device. It is not mandatory to select the FTP transfer, but it is highly recommended. Below there are couple of more reasons, why you should use FTP area to store the backups:

- APG has very limited disk space in the data disks
- Fetching the backups from the FTP area to the APG is easy and fast
- In case of data disk failure, you are still able to restore the node from the backups

6.3.3 Software comparison

Purpose of the software comparison functionality is to ensure that customers have baseline software in the nodes. When customers have baseline softwares in the node, it is much easier to determine what upgrade path will be used and also that no software modules are missing or newer version of that software module is not available. This is how it works; The software print is fetched out from the APG or from the any network element and after that it will be compared to the baseline software released from the product line. As a baseline software, the most suitable is selected (the one that is the most equal). Softwares are also compared between the nodes if redundancy exists. If there are any differences, report will be created and sent to the customers Network Operation Center. The current software level is also mentioned in the report.
Software comparison can be scheduled to be run how ever customer likes it. Nevertheless it is recommended to be run on a same time when the health check is done. As a stand alone functionality this is not the most important one, but this functionality is definitely needed in order to fill in challenges created in the introduction chapter.

6.3.4 Software auditor

Software auditor is a support tool for finding out problems that are already known and are related to the software itself for example some software module doesn't work properly in certain environments. This functionality checks the software print and compares it to a database, where all the trouble reports for the Ericsson software modules exists. If it finds a match it will send the problem description report to the customer. Only reports that have proper technical answer and solution will be sent. Any trouble reports that are under investigation are not reported and also trouble reports that do not have a valid solution are not reported. Valid solution in this case means that root cause of the problem has been found and it is corrected in the newer version of this particular module.
Trouble reports are loaded every day to the PNeMaS data base. When customer has selected this option, PNeMaS will fetch the software print from the node and compares it to the entities in the data base. If match is found, then it will be reported to the customer. Customer is able to schedule this task as well, but it is not mandatory.
Reason why this functionality is selected into this system is that it will ease up the network vendor problem investigation, if all the information about existing product faults are already found out. It will also prevent duplicate trouble reports to be created. This functionality does not give more possibilities to the end customer to solve the problems, but helps network vendor quite a lot.

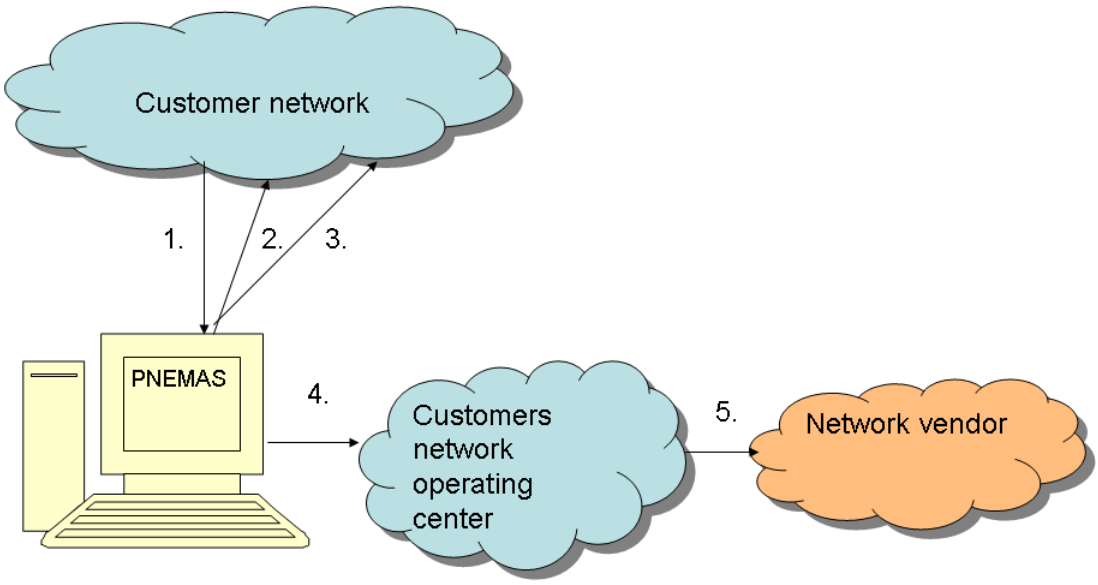
6.3.5 Alarm watch

The purpose of the alarm watch is to check if there are any new alarms in the node and act according to procedure described below.
When the alarm is raised PNeMaS will have the information that there is new alarm/alarms available. When this happens alarm watch will run the health check, take logs related to the statistics and take mktrs. Mktr is a list of logs that are created by the operating system and the Ericsson software, which helps engineers in trouble shooting and makes it possible to find out the root cause. In the below there are the most important logs that are included in the mktrs:

- Windows operating system event logs (application, security and system log)
- Microsoft Diagnostics Report
- Dr Watson logs[2]
- Software records (Ericsson and Windows)
- Logs about Ericsson software parameters

[2]Dr. Watson is a software troubleshooting tool, which collects information from the computer when there is an error in the program http://support.microsoft.com/kb/308538/fi

There are several other logs included in the mktr, but usually the root cause can be found from the logs mentioned above.
The purpose of the alarm watch is to collect as much information as possible and deliver that information further as soon as possible. When error happens and alarm is raised all the above logs will be taken. After that PNeMaS creates a report about the occasion and sends it to the Network Operating Center which will proceed with the necessary actions. Process flow is described in the figure 6.1
The purpose of the alarm watch is not to fix anything automatically, but help the operator to have all the necessary information to find the root cause of the alarm and depending about the problem give an indication what might be the problem. If the network operator is not able to solve the problem, network vendor has much better chances to find out the root cause as it has all the sufficient and fresh information taken right after the problem occurrence.
The reason behind this characteristic is that sometimes when end customer is facing the problem and wants to know the reason behind the erroneous behavior, it is impossible to find it out because lack of sufficient information. Alarm watch makes finding the root causes more probable and even faster, because of the amount of necessary information. Every time new fault with root cause is found it will be fixed and that makes the network element more reliable.



1. Alarm is raised in the customer network
2. Commands to take all the logs are sent to the network element
3. Logs are fetched from the network element
4. Logs are sent to the Network Operating center
5. Logs are sent to network vendor if necessary

Figure 6.1: Alarm watch task flow

6.3.6 Summarize

All the health check files are stored 6 months in the system. The reason behind this is that you can see the history of the node if it is needed. If the customer doesn't know when some fault appeared, you can see it from the history of the node. It is expected that 6 month is enough time to store the files.
If all the above functionalities are designed and same time period for all the checks is selected, customer will have a single report once in a selected time period. Recommended time period is once a week. In below there is an example about the report.

Time and date: 28.09.2007
Switch ID: MSC32
Node type: APG40
Errors:

1. Severity:High
   Following folders have corrupted access lists G:ftpvol Missing - Administrators:F
2. Severity:Medium
   The following services have the wrong startup state Windows Installer Cluster server
3. Severity: Low
   There are Core.xxx files present

Fixes:

1. echo Y—cacls G:ftpvol T Administrators:F
2. sc config clussvc start= auto
   sc config "windows installer" start= demand
3. Delete core files if those are old del F:acs data core Core.xxx

Software comparison:

Software in the APG matches to the baseline software CM013

Software auditor:

There are no new versions or trouble reports of any module available.
Next upgrade package will be AGM015, release date 1.11.2007

6.4 System requirements

The main focus in this section is to select all the important components of the PNe-MaS so that it can be easily built and also design the equipment that way that there are no barriers, which will limit possibilities for the network operators to purchase this equipment. All the possibilities will be evaluated and compared with the SWOT analysis or with the Force Field Analysis method and the most suitable's are selected. The purpose of the SWOT-analysis is to compare **S**trength, **W**eaknesses, **O**pportunities and **T**hreats of the possible solutions and select the one that has most advantages[8]. The aim of any SWOT analysis is to identify the key internal and external factors that are important to achieving the objective. The key in the SWOT analysis is to define clear objective. If clear objectives have not been defined, it might be that the whole analysis does not create any value and it is useless. SWOT analysis may be used in any decision-making situation when a desired objective has been defined. SWOT analysis has its own weaknesses. It presents the resulting lists uncritically and without clear prioritization so that weak opportunities may appear to balance strong threat. In this thesis we will put weight to every characteristic. Weight scale is from 1 to 5, where five is very important and one is less important.
In this thesis we have clear objectives. All the objectives of this thesis were defined as a key challenges in the introduction chapter. In this section our objective is to define system requirements so that "Equipment is available for as many customers as possible". Of course in the first place the primary target is to design equipment the way, it is possible for everyone to order it and the way it will create value to the every network operator. In this case SWOTs are:

- **S**trengths: features of the system or environment that are helpful to achieving the objective
- **W**eaknesses: features of the system or environment that are harmful to achieving the objective
- **O**pportunities: external conditions that are helpful for achieving the objective
- **T**hreats: external conditions that are harmful to achieving the objective

The other method used in this chapter to evaluate choices is Force Field Analysis. Force Field Analysis provides a framework for looking at the factors (forces) that influence a situation, originally social situations. It looks at forces that are either driving movement toward a goal or blocking movement toward a goal. Force Field analysis were first used in the social science, psychology and change management, but can as well be used in the selection of technical characteristics.
Each subsection of this section consists of one system characteristic selection. All the decisions are made based on the the results of the SWOT- or Force Field Analysis. In the end of this section all the data will be gathered and summary is made to find out if the objective was achieved.

6.4.1 Where to place the equipment

When designing the equipment it should be known where it will be put after the design process is over. During that decision making we have to have some kind of a objective where we are headed. In the first Chapter of this thesis four key challenges where defined. This section should answer to the key challenge "Equipment should be available for as many customer as possible".

Here in this situation there are two options. These options are listed below.

1. Customer network
2. Vendor network

These both options have advantages and disadvantages because we have clear objectives. In this case the best method to validate these options is to use SWOT analysis. In below both options are considered as a different solutions and SWOT-analysis is done to the both options. The option which score better will be selected as a solution.

Equipment in the customer network

Strength	**Weaknesses**
-No remote connection needed to the customer network (+4) -No Vendor maintenance of equipment is needed (+1) -Every customer has their own equipment (+1) -Configuration of the equipment is easier (+1) -Every customer has possibility to purchase product (+5)	-Updates has to be loaded from centralized database (-2) -Every equipment needs a connection to the Vendor network (-3) -Customer is doing maintenance to the equipment (-1) -Courses has to be arranged to the customer (-2)
+12	-8

Opportunities	**Threats**
-more customers will be able to purchase the product (+5)	-no database update are done by the customer (-3) Customers are not interested to install new devise to their network (-2)
+5	-5

Equipment in the vendor network

Strength	**Weaknesses**
-Centralized maintenance of the system (+3)	-All the customers are not able to purchase this equipment due to security policies (-5)
-Database updates are easy (+3)	-Huge physical space for the devices is needed (-1)
-Getting needed information for the customer service requests easily (+4)	-Maintenance of the equipment is done by the vendor and it creates costs (-4)
+10	-10

Opportunities	**Threats**
-Customers might by the product or service, because they get proactive maintenance with no work (+4)	-Customer specific information easier to sneak out. (-3) Direct access to the customer network (-5)
+4	-8

Results

As seen in the tables above, both options would have been very good as the final solution. Both options had their advantages and disadvantages. As the objective was that this equipment should be available for as many customers as possible, it meant that placing this equipment in the customer network would be better idea. The main reason for this is that some customers might have stricter security policies which does not allow to take remote connection to the customer network. As a result the equipment will be designed that way, that it will placed into the customers own network.

6.4.2 Selecting user interface

In this subsection user interface selection is done. Because we have already defined that there is no single software that is installed in the each user computer, but a one separate device which is used by the ones that are working in the network maintenance, the best possible choice is to use web based user interface. There are two possible programming languages that are possible.

1. PHP
2. Javascript

The only requirement for the programming language is that it has support to the database access, because all the information about the nodes and their tasks are stored in the database. In below there is a SWOT analysis from both options.

PHP

Strength	**Weaknesses**
-Widely used programming language (+3) -Good documentation easily available (+3) -support for the databases (+5) -Suited especially for web development (+3)	-current version of PHP has no namespace support (-2) - PHP does not have complete native support for Unicode or multibyte strings (-3) -The standard function library lacks internal consistency -2)
+14	-7

Opportunities	**Threats**
-Developing language (+2)	-New better programming language is developed (-1)
+2	-1

Javascript

Strength	**Weaknesses**
-Independent about operating system (+1) -Light to use (+3) -simple syntax (+3) -Good documentation easily available (+3) -Widely used programming language (+3)	-more difficult database access than in PHP(-4) -Poor security compared to PHP (-3)
+13	-7

As the both options have same opportunities and threats, it is not added here again.

Results

Both of these options have same kind of possibilities. Both have good documentation available and both languages are widely used. The most crucial difference between these two languages is that PHP has simpler database access and it supports many different databases. This is the reason why PHP is selected as programming language for the user interface

6.4.3 Selecting database

Because in the previous subsection we decided that the programming language we are going to use in the user interface of the system is PHP, we have to select the data base from those that are supported by the PHP. This in not a problem because there are several databases available. The most usable are listed in below.

- MySQL
- Oracle
- Microsoft SQL server

Because we want to have database which supports more than one operating system, it decreases our options to two, Oracle and MySQL. Force field analysis is made in order to find out the most suitable solution as a database. The one that has more forces for its side will be the selected solution.

Force Field Analysis for MySQL

Forces for MySQL	**Forces against MySQL**
- Available with free licence (5) - Cross-platform capability (2) - Documentation and manuals easily available (5) - One of the top performers in a group of databases (2) - Very commonly used (2) - Open source software which can be developed more by anyone (3)	- Some features that are available in the other data bases are not available in MySQL (-3)
+19	-3

Force Field Analysis for Oracle

Forces for Oracle	**Forces against Oracle**
- Has support to the several platforms (5) - good security (3) - Documentation and manuals easily available (5) - High performance (2) - widely used (2) - more features than in MySQL (3)	- No free licence available (5) - Not an open source software (2)
+20	-7

Results

Both options were very even, but after taking cons away from the pros MySQL has little better values. Even though Oracle has more features than in the MySQL, all the necessary features and functionalities are available in MySQL.

6.4.4 Selecting the Operating system

Now that we have already programming language for the user interface and data base for the equipment it is time to select the operating system for the device. Because user interface for this system is web based, operating system is invisible for the end customer and there is nothing that they basically has to do with the operating system, when the PNeMaS is in normal operation.
There are two possible solutions as an operating system for the PNeMaS:

- Windows Server 2003
- Linux Ubuntu Server edition

Force field analysis for Windows server 2003

Forces for Windows server 2003	**Forces against Windows server**
- User management is easy (3) - acquainted user interface and functionalities (4) - good support pages on the internet (4) - Clear and easily manageable File system permissions (5) - More hardware works on windows (3)	- Relatively high price (2) -source code not open (3)
+19	-5

Force field analysis for Linux Ubuntu Server edition

Forces for Linux	**Forces against Linux**
- freely available (2) - includes all the properties that were selected earlier (5)) - user guides easily available (4) - good security (4) - Includes Apache HTTP server (4) - updates possible without reboots (3)	
+22	0

Results

Even though Windows server 2003 would have been good option as an operating system, Linux has better characteristics for our purposes. Main issue is not that it is freely available, but the fact that this Linux Ubuntu server edition already includes everything we need and everything we already had defined as system requirements, like MySQL database and support for the PHP programming. It also includes Apache HTTP server. This is something we would have also needed.

6.4.5 Summary

Now almost all the decisions about system requirements and system software has been defined. There is still one more thing to be considered. In the system functionalities it was defined that all the data is stored in the PNeMaS hard drive for six months. This means that there should be large storage capacity available. If the checks are done on a weekly basis and customer has 200 nodes in their network, it means that needed storage capacity is 10Mbx26x200 $\approx$ 500 GB. Log sizes of 10 Mb is overestimated but sometimes those can be 10 Mb
In the figure 6.2 can be seen where the PNeMaS is being placed. From the customer network PNeMaS takes a connection to the Ericsson network from where it will get the latest information about the health checks, software releases, software corrections and update/upgrade packages. This data is then transferred to the PNeMaS database and it will be used when scheduled tasks are run.

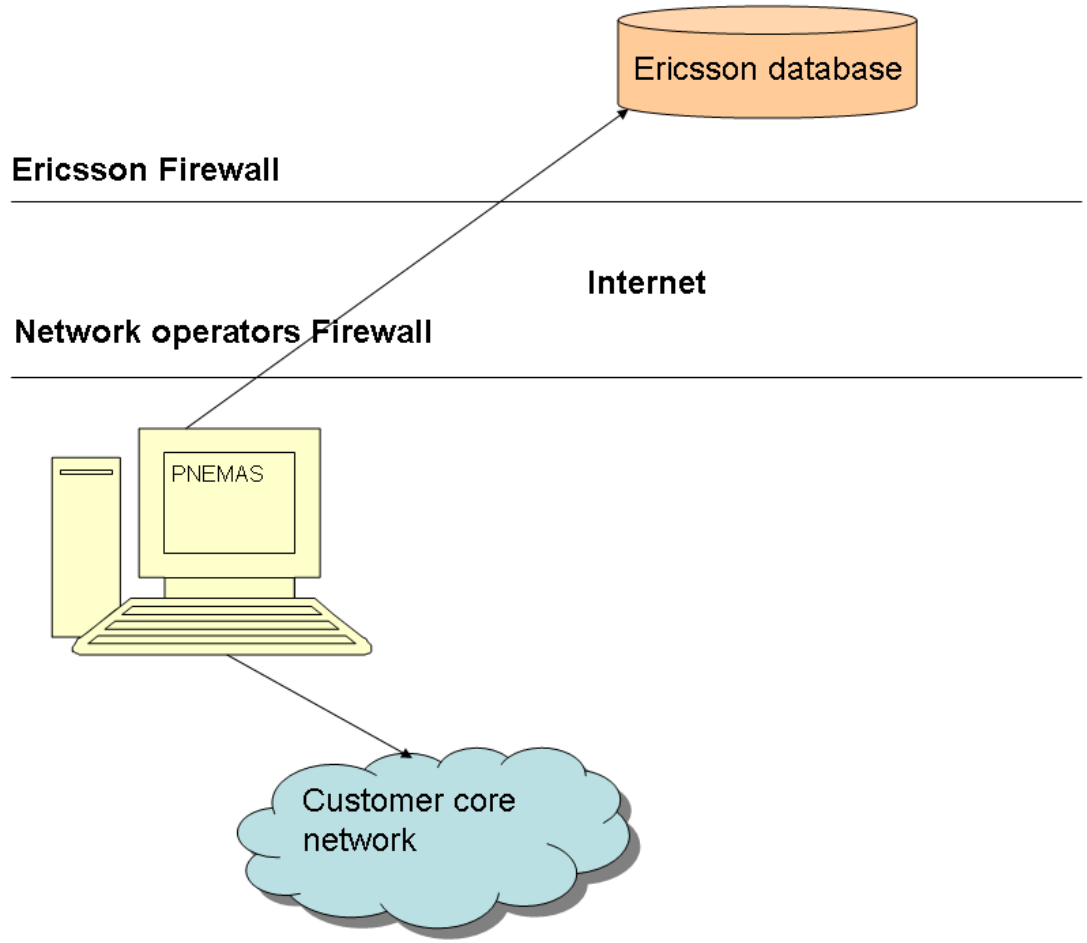


Figure 6.2: Where PNeMaS is placed

Other selected characteristics were:

1. Operating system: Linux Ubuntu server edition
2. Database: MySQL
3. User interface programming language: PHP

6.5 Equipment security

In this section we will go through the systems security. Many customers have different security policies. In this thesis we will create rather tight security policy so that it is suitable for every customer. End customer can later ease the policy if they like.
When starting defining the security we have to know what assets we are trying to protect. In other words we could say that what do you want that security does to me? In order to find out the security requirements we have to have some kind of threat scenario. Threat scenario is created in the first subsection. After creating the threat scenario, security requirements for the equipment will be created. Security requirements will define the equipments disallowed states, if some of the security requirements fail, system enters to the disallowed state.
Second subsection defines the security policies and mechanisms. This subsection creates the mechanisms that will enforce the system state to be allowed.

6.5.1 Threat scenario and security requirements

As stated earlier in the chapter 4 "The security of a system is a combination of its ability to support data confidentiality, data integrity and system availability[5]". In this case it means that we need to support confidentiality of the customer Switch data and passwords, customers to be able to count on a reports sent to them and ensure that system has high availability. The possible threats that can cause some of this to fail are discussed in the below:

Threats affecting to the high availability

Because this designed equipment doesn't directly affect to the operation of the customer network it is very unlikely that someone tries to attack against the system as a purpose to cause denial of service or slow down the normal operation of this equipment. If someone tries to attack to the system the most probable attempt is made from the inside of the customer network and the target is the physical hardware. The reason for this is that the equipment is not directly accessible from the public internet.
Other problem could be that someone adds huge amount of nodes to the database, which does not exist in the customer network at all and try to cause problems to the nodes normal operation.

Threats affecting to the data confidentiality

Most probable threat this equipment is suffering is that someone tries to find out information that he not allowed to have. One example about this is the switch user-

names and passwords to gain access in the switch and fetch out sensitive information about the customer and the mobile equipment users for example end user locations and to whom end user calls.
Other attractive targets are the logs and reports this equipment holds. These logs and all the history data of these node could be deleted or sent to the unauthorized people. In the worst and most probable case these logs will be deleted.

Threats affecting to the data integrity

One threat that can happen is that someone modifies the data in the reports that are sent to the customer. The data can be modified either in the reports or in the root files. As an example: Someone changes the data in the database which consists the procedure done when the fault happens to some other procedure which causes harms to the switch. This problem is not very likely to happen but must be taken into consideration anyway because if this kind of thing happens, customer can't trust on a report results anymore.

Other threats

One type of a threat is that some with too high access rights access to the node and deletes either node information or logs from the database by accident. The probability of this kind of issue in the long run is high and should be taken seriously.

Security requirements

From the threat scenario we could see what are the most probable threats that are affecting to the device. According to these we have to define our security requirements or goals. Because the most probable threat was that someone tries to have information not belonging to him, our first goal is to protect unauthorized access of any data in the system.
As there are several smaller threats as well we have defined other security requirements so that system can be considered secure and all the information sent to or from the equipment are safe. In order to consider system as a safe we need three different security requirements and those are introduced in the below.

- All the customer data (node history, node data, passwords and etc.) are protected from unauthorized access
- All the reports sent to the customer are reliable and unmodified.
- System availability is high

6.5.2 Security mechanisms

Security mechanisms are the methods that are used to meet the security requirements. The first thing we have to ensure is that unauthorized people can't access to the system. For this purposes every user has their own username and password. The password has minimum length of eight characters and in the password there has to be numbers and uppercase and lowercase letters. Password has to be also complex enough.

To fulfill the first and second security requirements we have to do following:

- All the password are encrypted in the system
- Folders where the Node logs are located, have limited access rights
- Health check files and reports exists in the directory that needs a separate password
- Not all users have administrator rights in the system

When all the passwords are encrypted it makes it much harder to find out the clear language passwords.
Node logs are not so attractive target so there is no need to encrypt or password protect those logs. It is enough that only limited group has access to these files. This is also the reason why there is no need to send the logs to the customer via secure email.
In order to protect the integrity of the health check files and reports sent to the customer, those should be placed in a folder protected by the password or place in a folder where only administrators have access. Not every user can have administrators rights in the system. If this would be the case, creating just a generic user would be enough and threat of any kind that comes from inside the customer network would have been highly underestimated. When users have proper access rights according to their tasks and skills to the system, the possibility that someone harms the system by accident is decreased.
To fulfill the third security requirement, one has to take care that this system is physically placed in the place where there is only limited access.

6.6 User interface design

In this section we will design user interface for the PNeMaS. During the design procedure we will use the nine key principles of user centric design, described in the usability chapter, as a base for our design process. These key principles will help during our design process to find out the key characteristic of a successful user-interface design.
In this section we will also focus on fulfilling the key challenge number four, introduced in the chapter one. In the chapter one it was stated that *Equipment should be easy to use and maintain.* It means that usage of the equipment is logical and PNeMaS guides the user to automatically select the tasks he/she wants to do. Main purpose is that PNeMaS could be used efficiently without any user manual or expert help.

6.6.1 Defining goals to the design

User centric design process starts with evaluating usability goals and specifications. This should also be the main driver in the design process. During the design process suggestions and results are evaluated against the goals.
What we need to fulfill user requirements and our requirements introduced in the chapter one?
Our own goal was that the equipment should be easy to use. In general it means that the usage of the equipment is logical and system automatically guides you towards wanted action.
User requirements can vary a lot, but main requirements are that the equipment is efficient to use, which means that user effort is as low as possible. Visual expression should be neutral so that it doesn't evoke bad feelings.

Technical goals

In order to satisfy security requirements we need a login screen which will ask username and password. This is very essential, because otherwise unauthorized access is possible. This is where we will start.
Second, we have to have page which will include the possible tasks we are able to do, for example adding network elements, removing network elements and etc.
In the system functionality section we defined what tasks the system is able to do. Now we have to have a way to select the tasks and select the nodes where we want these tasks to be executed. We also have to have way to change the assigned tasks or network element information. In the below there is a list of mandatory tasks that PNeMaS has to be able to handle.

- Add network element
- Remove network element

- Configure tasks done in the network element
- Change network element configuration

To keep the system as simple as possible, but still practical there might be a need for other functionalities as well. This will be found out later in this chapter.

6.6.2 Design process

We start the design process by defining the target audience of this equipment. The target audience is the people who are working in the network operating centers and their knowledge about technical equipment is considered as very good. In our case it means that designed system can be little more complex. It still doesn't necessary mean that resulted user interface is complex at all. To meet the business objectives system has to be user-centred and well designed to the target audience.
During the design process ideas from the users were taken into account. The users that participated to the design process had very different backgrounds and the purpose was to have feedback, both from the practical and visual implementation.
Most challenging task in the user-interface design is definitely the fact that this system itself should guide you to do right decisions in order to complete some task. This design process results a prototype which will be then evaluated by the users and also usability measurements with this prototype will be done. The results of this part of the project will be gone through in the next section.

6.6.3 Results of the design process

Sign in page

As stated earlier in this section, the design process was started from the sign in page design. The requirements for the sign in page was that it should authenticate the user who tries to login to the system. It was also decided that sign in page is as simple as possible, but not boring. Only necessary information will be presented. From the figure 6.3 you can see the prototype picture.

Front page

The main focus in the front page design was to add only necessary information and help to the page. In the front page there should also be the tools available so that those can be selected from the tools menu. Earlier in this section it was stated that at least this system should be able to add network elements, remove network elements and change network element configurations. This would have been enough, but it was decided that possibility of listing network elements could be important in some situations.

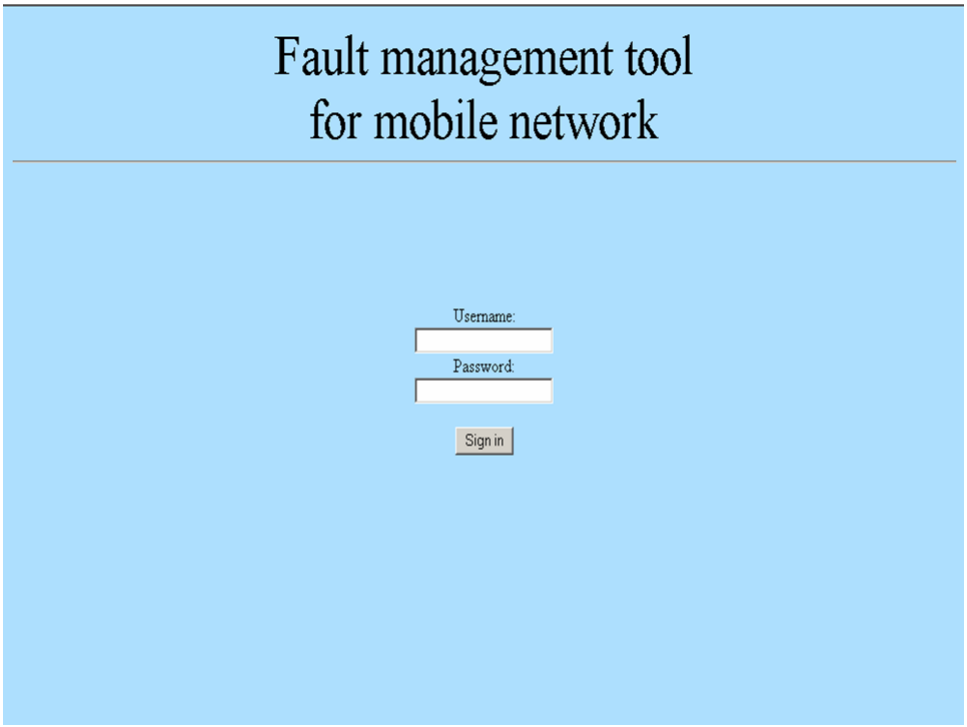


Figure 6.3: Sign in page

To make this system more simple to use it was decided that from everywhere in the designed user interface you can select which task or functionality you like. Because of this there is small column in the left side of page. It is similar on each page. This characteristic makes the user interface much simpler to user and outward is more stable. In the figure 6.4 front page of this system is shown.

Adding network element

Adding network element with this user interface is made very easy. In the first step user has to select the network element and after that only thing that is expected from the end user is to fill in the node information and the tasks that are wanted to be executed in the node. When the information is submitted, user will be given a summary about the information saved in the data base and necessary information if it was successful or not. This was first tried without feedback to the end user, but it was found to be confusing. In the figure 6.5 network element selection is shown.

In this example APG is selected as a network element and in the figure 6.6 example about element configuration and tasks is shown.

In the figure 6.7 printout after successful element adding to the data base is shown.

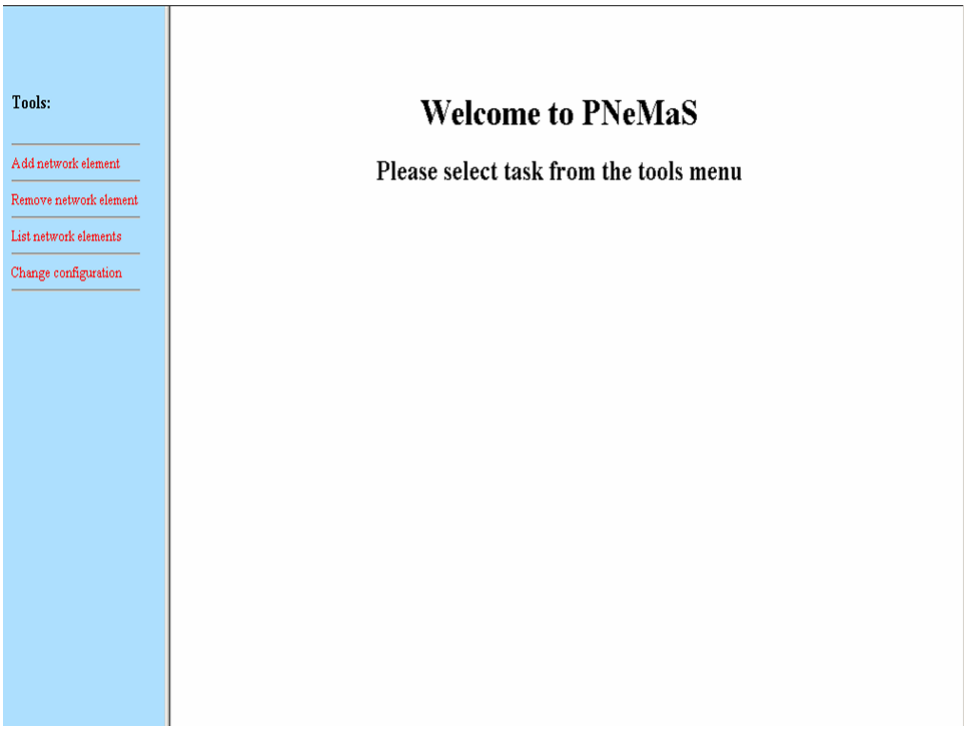


Figure 6.4: Front page of PNeMas

Removing network element

If the network element doesn't exist any more it has to be removed from the data base. This feature was created so that data base contains only valid information and old information can be removed. In this feature all the nodes are printed and in the end of every node there is possibility to delete the node (see figure 6.8).

Listing network elements

This feature was made because there might be situations that user wants to see what nodes are already in the data base and what tasks are assigned to this node. It was decided that all the nodes, properties and tasks of the nodes are printed in this feature (see figure 6.9).

Change network element configuration

Changing the network element configuration is the most important feature in order to describe this system as usable. This is because many customers add new network elements and remove older elements from their networks. There are several situations why you need this tool. Customer might change its network configuration and change for example IP-addresses of a certain network element. It might be

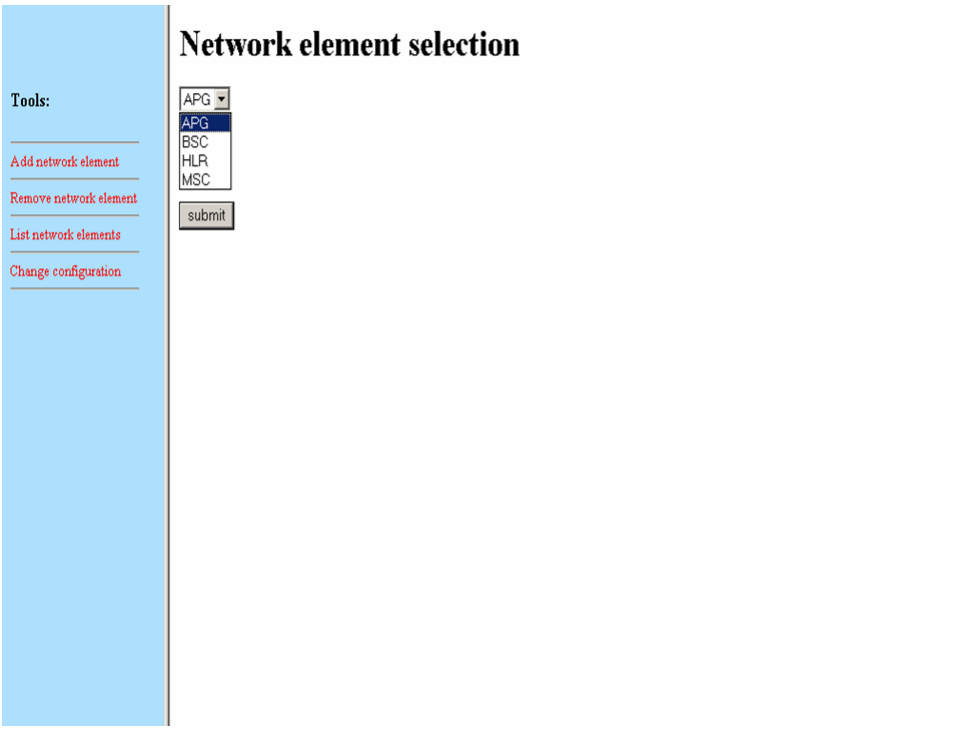


Figure 6.5: Element selection with PNeMas

also that some tasks are added or removed to the network element. Without this tool configuration changes should be made by first removing the element and then adding the element, which takes much more time than just changing one task or IP address.
When you start modifying the node information, old information is present in below of the page (figure 6.10), so you can see it all the time. When the information is changed, feedback if the action was successful or not is given.

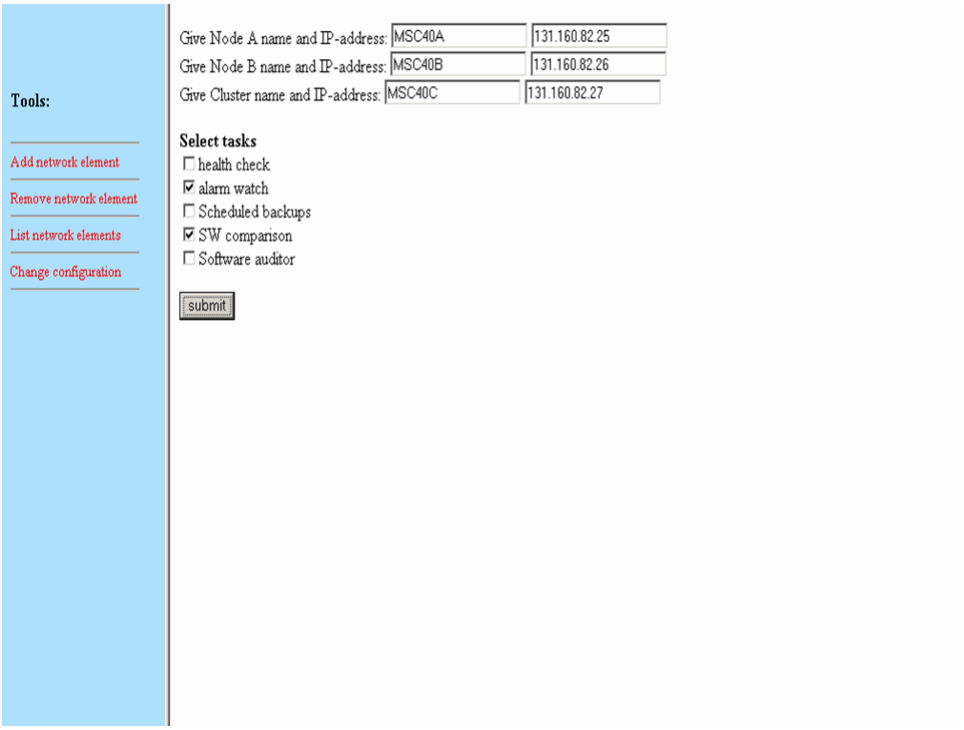


Figure 6.6: Task selection with PNeMas

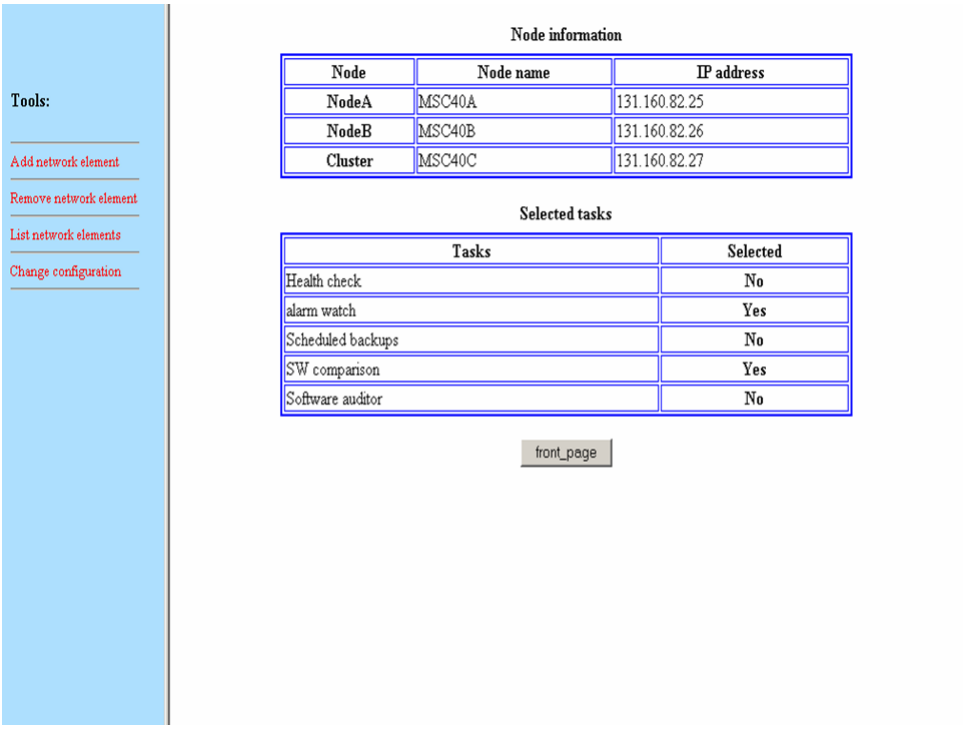


Figure 6.7: Feedback after successful element adding

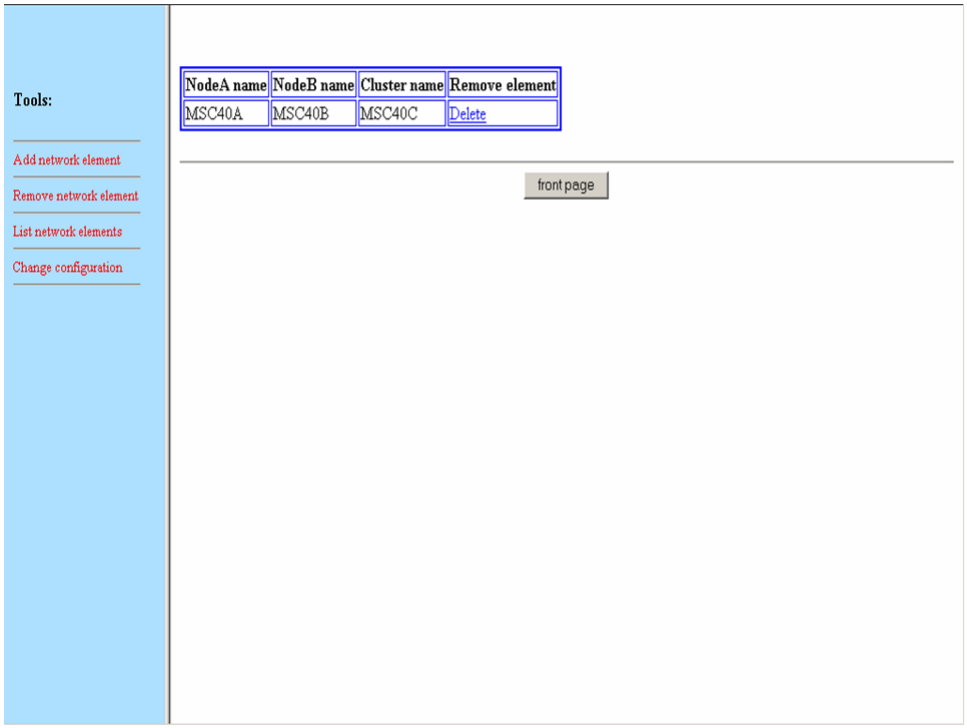


Figure 6.8: Removing network elements

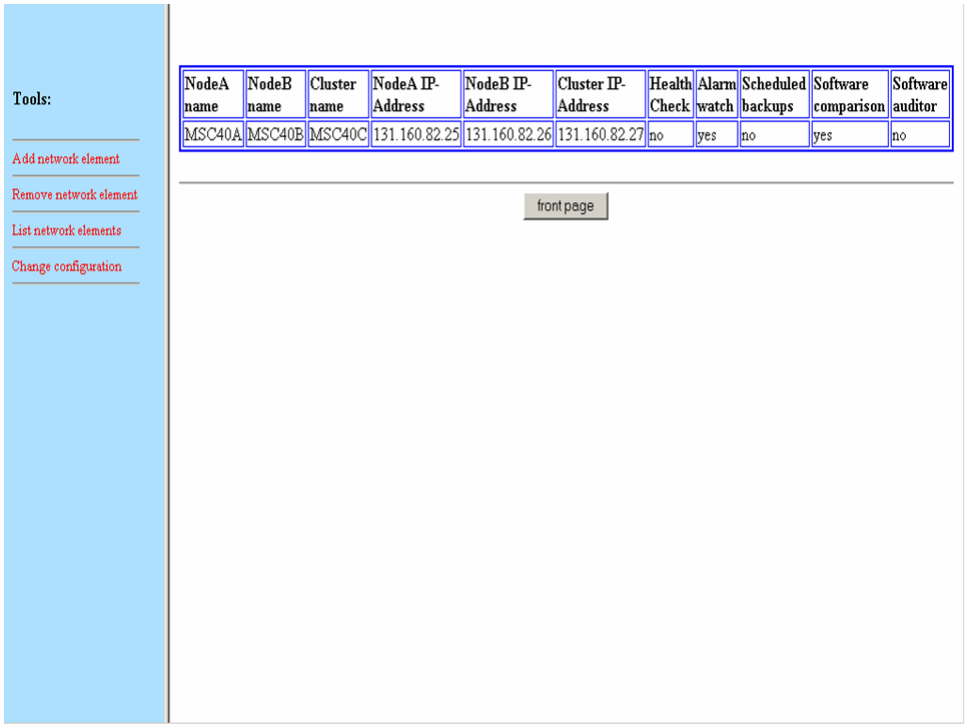


Figure 6.9: Listing network elements

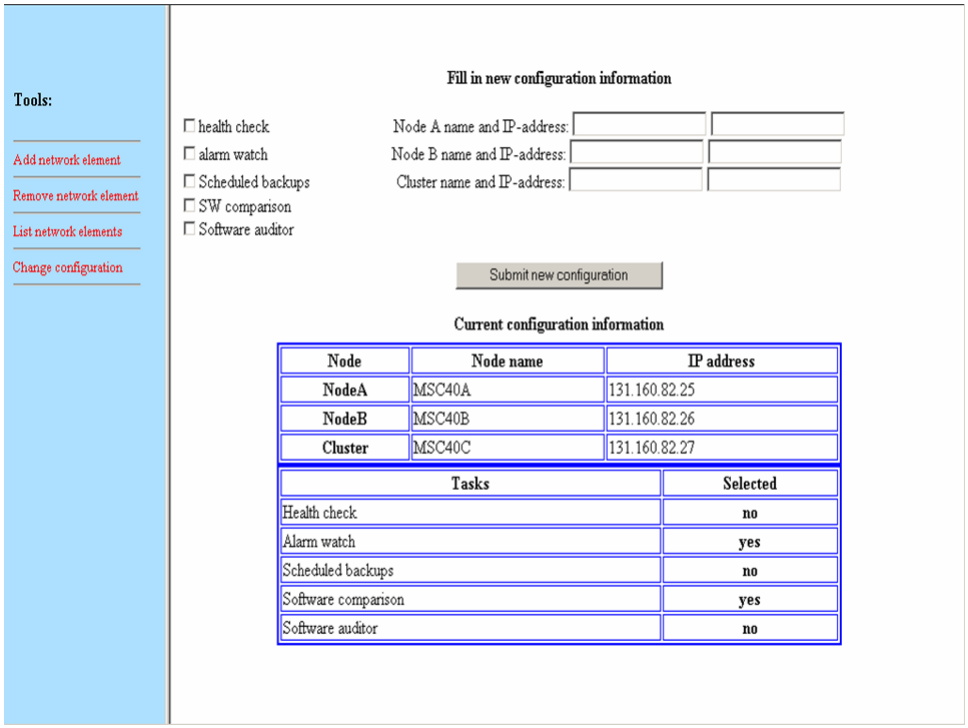


Figure 6.10: Change configuration of network elements

6.7 Measuring usability of the equipment

After the design process, usability of this PNeMas was measured. Measurements are performed with usability tests. In this case usability is measured relative to users' performance on a given set of test tasks. Test tasks are introduced in the Test Scenario subsection. In order to measure usability of the product following information is needed:

- Intended goals
- The context
- The expected results

Intended goals

Intended goal is to perform all the tasks specified in the "Test Scenario subsection" within 10 minutes with error rate less than 5%. This means that every person who takes in part of this measurement can do two mistakes and still the goal is reached. Every mistake in this test is graded equally so it doesn't matter what mistake you do it is still counted as an important.

Context

The context is specified with reference to the users knowledge, education. The context should also specify the surrounding as the equipment and tools that should be available. In this case everyone should be able to use this equipment despite their background knowledge and only tool that is available is computer.

Expected results

In order to determine any usability measurement, some sort of expected target of effectiveness, efficiency and satisfaction based on the goals and context must be defined. Expected results are based on the answers to the following questions:

- percentage of successfully completed tasks ≥ 95
- Time needed to complete task ≤ 10 minutes
- Number of errors made ≤ 2
- Users subjective satisfaction on scale 1 to 5

As stated in the Usability chapter, it is enough to test usability of a product with five people. Because of that PNeMaS is tested with five people. Three of those people are using computers in their daily work, while the other two are not.

6.7.1 Test Scenario

You have installed five new APGs to your network, four to the live network and one to the test plant. You would like to use PNeMaS to help to keep the nodes in good shape and also help in your daily maintenance work. In order to do that, all the nodes and their tasks will be added to the PNeMaS database. In the table 6.1 nodes and task that are added to the database are described.

Table 6.1: Node and task information

NodeA name	**NodeB name**	**Cluster name**	**NodeA IP**	**NodeB IP**	**Cluster IP**	**Tasks**
MSCAPG03A	MSCAPG03B	MSCAPG03C	131.160.80.181	131.160.80.182	131.160.80.183	all
MSCAPG04A	MSCAPG04B	MSCAPG04C	131.160.82.3	131.160.82.4	131.160.82.5	all
HLRAPG02A	HLRAPG02B	HLRAPG02C	131.160.81.154	131.160.81.155	131.160.81.156	all
BSCAPG03A	BSCAPG03B	BSCAPG03C	131.160.80.72	131.160.80.73	131.160.80.74	all
TSTAPG02A	TSTAPG02B	TSTAPG02C	131.160.80.163	131.160.80.169	131.160.80.165	all

After a while you notice that test plant APG TSTAPG02 has wrong IP address for the nodeB in the database. The right IP address should be 131.160.80.164. You have to change that, you also want to change the nodes task list so that it will only do health check and scheduled backups.
After some time you have forgotten which network elements are in the database and you will list all the elements. You noticed that there exists one network element (MSCAPG03) in the list that you have already removed from your live network and your test APG does not exist any more either. Because of this you will decide to remove those from the PNeMaS database.
After the tasks each test person was asked to give feedback from the overcome of this user interface

6.7.2 Test results

Each user had to make 61 tasks described above Each user was satisfied with the outcome of this user interface. Average value for subjective satisfaction was 4. All the people that were participating in the usability measurement tests were able to complete all the tasks with less that 5% error rate. Everyone was also able to complete all the tasks in less than 10 minutes. Time was calculated from the moment each test person got the task instructions. Test results are introduced in table 6.2

Table 6.2: Performance by each test user

Number of errors (%)	**Time Spent**	**Number of errors**	**satisfactory level**
3,3	7min 40 sec	2	4
0	8min 50sec	0	5
1,6	7min 20sec	1	3
0	9min 10sec	0	5
3,3	6min 45sec	2	3

6.7.3 Summary

It can be said that the user interface design process was successful. All the test users were able to complete all the task within the limits which was set before. Main reason for this result is that nine key principles of user centric design introduced in chapter 3 were followed very strictly. Target group was specified early in the design process. Feedback from the people were taken into account during the design process. Very simple dialog was used in the user interface, necessary amount of feedback was given to the end users and usability goals of this user interface were defined in the early phase of the design process.

Chapter 7

Further development

New faults and error behaviors are found out all the time. In order to fulfill the requirements of network operators in the all market situations and develop towards more automated maintenance, further development of the designed equipment is needed. Purpose of this chapter is to see the opportunities where this system could be developed in the future and how it could help the network vendor as well.
At the moment as such the designed system saves time in the network maintenance, finds faults and eases up the root cause analysis. It is possible that error detection and network maintenance could be more efficient. One way to improve the network maintenance is to update the list of already known faults and procedures how the faults are fixed. This approach is dealt in the first section of this chapter. This type of error detection and network maintenance is based on the fact that problem or error behavior is already know.
There are also methods available which not only detect errors, but can predict errors from the equipment behavior. In this approach there is no need to specify errors beforehand. This approach is gone through in the second section of this chapter.

7.1 Updating the known faults

In order to keep the equipment attractive from the customer point of view, it has to satisfy customer needs. As stated earlier in this chapter new faults are found all the time. In order to stay in touch, known fault descriptions have to be updated all the time. Also fixes to the problem have to be updated all the time.
This is a good way to stay in touch, also in the long run because all the newest fault descriptions and fixes are available.

7.2 Proactive fault detection

In the previous chapters and sections we have talked fault detection in theory and in practice. The problem in the approach we have used so far is that we have used a methods that are only applicable when the faults that should be detected are specified beforehand.
In this section we will go through an other fault detection method, which is even more efficient than the ones we have used. The difference between this and the previous ones is that in this method the system, that monitors the fault is adaptive. It means that system learns the normal behavior of each measurement variable. Successful prediction of errors in a computer system results to the significantly improved system management. At least there are following benefits:

1. Unknown faults can be detected
2. Subtle changes before the failure are detected
3. Faults can be detected when they are developing

These benefits will lead to a situation where the network management has more time to take corrective actions and reduce or prevent the system downtime or possibly steer some jobs away from the failing nodes. These predictions will reduce scheduled system maintenances and will avoid unplanned outages. Even if the prediction of an error happens too late to allow proactive action, the same analysis can serve as a foundation for more effective error isolation.
There are several number of prediction algorithms that has been proposed to help with the proactive error detection such as time-series algorithms, rule-based classification techniques and Bayes network models. Based on the analysis made in [15], different classes of algorithms are effective at predicting different kinds of system events.
It is feasible to predict system performance related parameters with high degree of accuracy using time series models[15]. Similarly rule-base classification algorithms can predict critical events with 70% accuracy. Bayesian network based algorithms can be successfully used to build dependency graphs to isolate the root cause of the problem.
In this section system performance and root cause isolation as such are off the interest and those are not gone through here. In the next subsection Rule-base algorithms are gone through in more detail.

7.2.1 Rule-based classification Algorithms

Learning to recognize rare events is a difficult task. The difficulty stems from several reasons: Very few examples support the target class; events are described by categorical features that display even inter-arrival times, such as those occurring in computer-network logs and transaction logs. Predicting target events involves the following steps:

- Finding all event types which frequently precede target events within a fixed time window. This set of events is called eventset
- Validating eventsets that uniquely characterize target events, and do not occur far from the time arrival of target events
- Combine validated eventsets to build a probabilistic rule-based system prediction.

Searching for frequent events

We are interested in finding sets of event types which occur frequently before event within a a time window W see figure 7.1. On every occurrence of a target event, all event types within the window are stored as a new transaction. Once all events have been analyzed, it is straightforward to apply an association-rule algorithm to find all eventsets above a minimum user-defined support[15].
Ordering of events and the inter-arrival times between events within each time window are not relevant. This is useful when an eventset occurs under different permutations, and when inter-arrival times exhibit high variation. These characteristics are present in many domains, including real production networks.

Accurate eventsets

There is another data filtering process conducted within the algorithm based on the rules establishing the confidence levels. The general idea is to look at the number of times each of the frequent eventsets occurs outside the time windows preceding target events. Such an information enables us to compute the confidence of each frequent eventset and to eliminate those below a minimum threshold.
Validation phase ensures that the probability of an eventset Z appearing before a target event is significantly larger that the probability of Z not appearing before target events. The validation phase discards any negative correlation between Z and the occurrence of target events. This phase serves as a filtering step to reduce the number of candidate patterns used to build a rule based model for prediction.

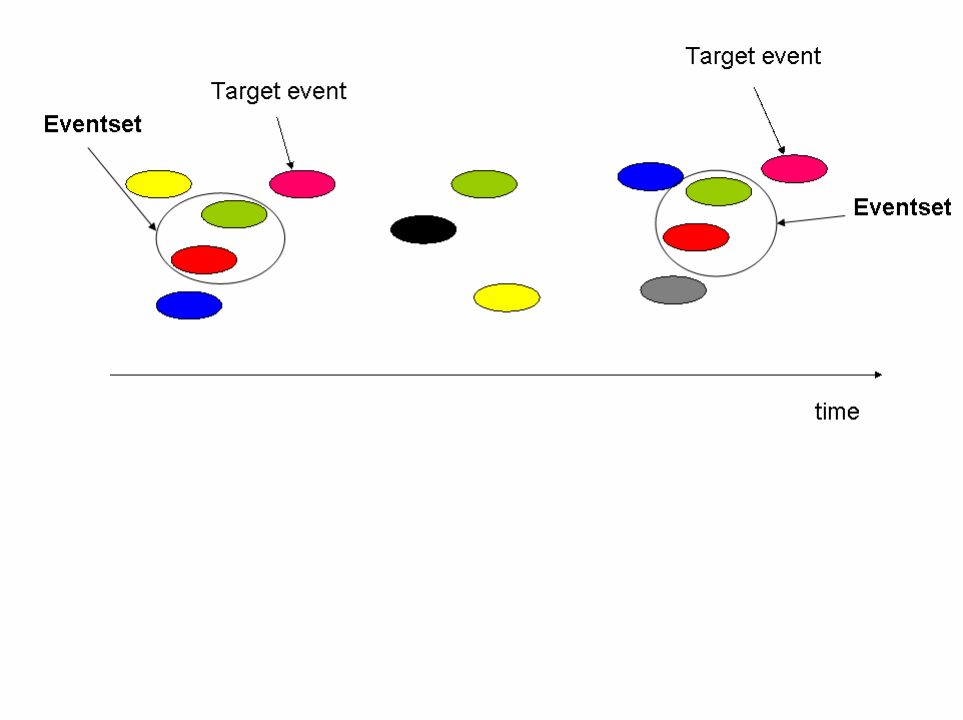


Figure 7.1: Searching target events with help of eventsets

Rule-based model

Frequent and accurate sets of eventsets are combined into a rule based model. The idea behind rule-based system is to find the most accurate and specific rules first. Let $\Im$ be the set of large and validated eventsets. The first thing is to sort all eventsets according to their confidence. In the next step predefined algorithm selects the next best eventset ℓ and removes all other eventsets $\wp$ in $\Im$ more general than ℓ. The resulting rule is of the form $\ell \longrightarrow$ target event. The search then continues for all eventsets capturing different patterns preceding the occurrence of target events. The final rule-based system $\Re$ can be used for prediction by checking for the occurrence of any of the eventsets in $\Re$. This model predicts finding a target event within a time window of size W after any such of eventset is detected. According to the experimental results rule-based classification can predict critical events with up to 70 percent accuracy[15].

Chapter 8

Conclusions

This chapter has two objectives. The first one is to provide short summary about the creation of this thesis, where this idea came from, what were the goals of this thesis and how well these goals were achieved. The second objective is to summarize the benefits that this thesis creates, both to the end customer and for the network vendor.

8.1 Thesis Evolution

The Purpose of this thesis was to design a system or software, which will help the end customer in their everyday network maintenance and fault detection. This system will help network vendors to solve the problems that customers are not able to solve by itself much faster. In case new fault is found it could be identified faster with all the information and logs available, so that the root cause analysis of the problem could start as fast as possible.
The idea of this raised because there were so many customer service requests from the problems, that had been already solved by someone else inside the organization -already known faults, but still those problems were escalated to the higher level in the organization.
The second main driver for designing this system was that very often when the root cause analysis of the problem had to be done, some essential information was missing and the finding out the root cause turned out to be very difficult or even impossible.
The third driver was that there were some situations that customer had software corruption in the system drive, but no valid backup available.
All the information for the effective problem solving are already available, at least in some extent, but efficient use of this information is a problem. This thesis is about combining the existing information in a way that all the resources, time and information in the area of network management, fault detection and root cause analysis are used efficiently. This leads to the situation where less time is spent

with the known problems and more time can be spent finding and fixing new faults.

8.1.1 Creating and meeting the challenges

To be able to fulfill the intended goals, we have to find out the most difficult parts in designing this equipment. These part are considered as a key challenges of this thesis. There were four key challenges in this thesis and those are introduced in below.

1. Error detection of the equipment should be effective
2. What to do when possible error is detected?
3. Equipment should be available for as many customer as possible
4. Equipment should be easy to use and maintain

In this chapter we will go trough those challenges and take a look how we were able to meet these challenges.
In order we can say that the designed system could be successfully delivered to the end customer, the properties it has, has to give some benefits to both to the network operator and to the network vendor. This means that the key challenges created earlier has to be met.
While designing the equipment we have been forced to take many decisions and according those decision, we will see how well the key challenges were met and what possible disadvantages there is in the designed equipment.

Error detection of the equipment

Error detection in this system is made by monitoring known problems. If the fault is found report and the way fault should be fixed is created and sent to the customer. This is very effective way to monitor the faults, only problem in this approach is that the fault has to be already known. In the beginning of this thesis our goal was to create effective way to find the fault from the customers equipment. In that time it was stated that all the known faults should be detected. Because the list of the known faults are updated in a monthly basis this method is very efficient.

Procedures when error is detected

When error is detected or alarm is raised in the node all the possible logs and full health check will be run immediately. The log files are saved in to the equipments hard drive and those logs are sent to the end customer as well. If there are any problems mentioned in the report customer can fix those by itself. If it doesn't help, escalation to the network vendor is needed. If fixing the error needs special

attention or system expert to fix it, it is mentioned in the result file. This procedure ensures that there are sufficient logs available in case root cause analysis is needed. This procedure also makes it easier for the customer to fix the problem if it is not a new one. Procedures what to do when error or alarm is detected are very straight forward.

1. Run the complete health check
2. Collect all the necessary logs
3. Save the logs and reports to the data base
4. Send the logs and reports to the end customer

System applicable to the all the customers

One of the key challenges was that every customer should have possibility to purchase this equipment and it should be applicable to every customer as well. Because some customers might not have remote connection to their network or it is allowed to take that only in emergency situations, it was essential that this equipment would be placed inside that customer network.
Other problem was that different customers have different security policies. Because of that the equipment security is flexible, but certain guidelines are given about what information should be protected and how.

Easy to use and maintain

According to the measurements made in the 6 designed system is very usable and the user interface is clear and easy to use. The maintenance need of this equipment is low. Old information is deleted when its half year old so disks are not filling up. Only thing is possible software updates.

8.2 Benefits of the study

This section summarizes the benefits this equipment will create both to the customer and to the network vendor.

8.2.1 Benefits to the end customer

The purpose of this thesis was to design a system which helps customer in their every day network maintenance. This objective was met. This means to the network operator that the basic network maintenance tasks are more efficiently done. In case any faults in the nodes are found, those can be fixed more efficiently. This is because

large amount of known configuration and other faults and how to fix those problems are available in the designed systems database.
On the other hand this means that customer will have faster and more accurate support from the network vendor. In the reports that are sent to the customer in case a of a failure or in a prescheduled basis, customer can see what actions are required in order to solve the problem and how difficult the task will be -does it need an expert help to fix the problem.

8.2.2 Benefits to the Ericsson

For the Ericsson this means that less customer service requirements are coming in, but the ones that are coming are more demanding. This leads to a situation where more people can focus on finding the solution to the new problems and less people are focusing in the problems that are already know and have a valid solution.
On the other hand, with the help of this equipment all the necessary logs from the nodes are available, which leads to a situation where root cause analysis is easier to do and fault or problem is more likely to be found. This also saves time when you what all the possible information available immediately.
In the both situations Ericsson is saving money by saving time and by moving easier parts of problem solving to the end customer and handled only more demanding operation and maintenance tasks, root cause analysis and new software faults itself.

Bibliography

[1] TS 23.002. Technical specification group services and systems aspects; network architecture. Technical report, 3GPP, 1999.

[2] TS 100 522. Digital cellular telecommunications system (phase 2+); network architecture. Technical report, 3GPP, 2000.

[3] Ross Andersson. *Security Engineering: A Guide to Building Dependable Distributed Systems*. Wiley & sons, 2001.

[4] Matt Bishop. What is computer security? *Security Privacy Magazine, IEEE*, pages 67–69.

[5] Matt Bishop. *Introduction to Computer Security*. Addison Wesley Professional, 2004.

[6] Jan Gulliksen. Key principles for user-centred systems design. *BIT*, pages 5–7, 2003.

[7] ITU-T. Recommendation x.733, information technology - open systems interconnection - systems management: Alarm reporting function, 1992.

[8] Jan-Erik Lindroos ja Kari Lohivesi. *Onnistu strategiassa*. WSOY, 2004.

[9] Maria Köykkä. Käyttäjäkeskeinen suunnittelu organisaation näkökulmasta. Lecture notes, S-72.2510 Tietoliikennepalveluiden käyttäjäkeskeinen suunnittelu, 2006.

[10] Nigel Bevan; Miles MacLeod. Usability measurement in context. *National Physical Laboratory, Teddington, Middlesex*, 13, 1994.

[11] Dilmar Malheiros Meira. A model for alarm correlation in telecommunications networks.

[12] Jakob Nielsen. Why you only need to test with 5 users.

[13] Jakob Nielsen. *Usability Engineering*. Morgan Kaufmann, 1994.

[14] Markus Peuhkuri. Introduction to communications security.

[15] Ramendra K. Sahoo, Adam J. Oliner, Irina Rish, Manish Gupta, José E. Moreira, Sheng Ma, Ricardo Vilalta, and Anand Sivasubramaniam. Critical event prediction for proactive management in large-scale computer clusters. In Lise Getoor, Ted E. Senator, Pedro Domingos, and Christos Faloutsos, editors, *KDD*, pages 426–435. ACM, 2003.

[16] George I. Zysman, Joseph A. Tarallo, Richard E. Howard, John Freidenfelds, Reinaldo A. Valenzuela, and Paul M. Mankiewich. Technology evolution for mobile and personal communications. *Bell Labs Technical Journal*, January 2000.