**Πανεπιστήμιο Πειραιά**
Τμήμα **Ψηφιακών Συστημάτων**
Π.Μ.Σ. ***Τεχνοοικονομική Διοίκηση και Ασφάλεια Ψηφιακών Συστημάτων***
Κατεύθυνση ***Ασφάλειας Ψηφιακών Συστημάτων***

# A Risk Analysis Methodology for Modern Networks

**MSc Thesis by**
**Nikolaos Vavoulas**

**Supervisor**
**Dr. Christos Xenakis**

**Submitted to the Department of Digital Systems at the University of Piraeus in partial fulfillment of the requirements for the degree (MSc)**

**FEBRUARY 2011**

# Περίληψη

Τα τελευταία χρόνια, οργανισμοί και επιχειρήσεις ανά τον κόσμο αντιλαμβάνονται την ανάγκη υιοθέτησης κάποιου προγράμματος διαχείρισης επικινδυνότητας προκειμένου να ενισχύσουν την ασφάλεια των πληροφοριακών τους συστημάτων. Ωστόσο, η πλειοψηφία των ποιοτικών/εμπειρικών μεθόδων δεν είναι συμβατές με το πρότυπο ISO-27005, το οποίο περιγράφει όλες τις διαδικασίες που πρέπει να ακολουθούνται κατά την διαχείριση της επικινδυνότητας και μελετάνε τις απειλές μόνο σε υψηλό επίπεδο, αγνοώντας σημαντικές παραμέτρους σχετικές με την κάθε απειλή. Στην παρούσα εργασία, αφού κάνουμε εκτενή αναφορά στην σχετική με την ανάλυση επικινδυνότητας, αλλά και εκτίμηση ασφάλειας γενικότερα, βιβλιογραφία προτείνετε μια ποσοτική μεθοδολογία ανάλυσης επικινδυνότητας για εσκεμμένες απειλές. Η προτεινόμενη προσέγγιση ακολουθεί τα βήματα που προτείνονται από το πρότυπο ISO 27005, επεκτείνοντας τα προκειμένου να εστιάσουν στις εσκεμμένες απειλές κα τα διαφορετικά επεισόδια ασφάλειας που τις πραγματοποιούν. Η προσέγγιση που προτείνετε έχει τρία διακριτά επίπεδα: Το επίπεδο της εννοιολογικής θεμελίωσης, το επίπεδο των εργαλείων μοντελοποίησης και το επίπεδο της μαθηματικής θεμελίωσης. Το επίπεδο της εννοιολογικής θεμελίωσης ορίζει και αναλύει τις εμπλεκόμενες έννοιες χρησιμοποιώντας τα διαγράμματα κλάσεων της Unified Modeling Language (UML). Το επίπεδο των εργαλείων μοντελοποίησης εισάγει συγκεκριμένα εργαλεία, τα οποία βοηθάνε στην μαθηματική μοντελοποίηση των σχέσεων των διαφορετικών εννοιών. Το επίπεδο της μαθηματικής θεμελίωσης περιλαμβάνει όλες τις μαθηματικές συναρτήσεις και τεχνικές που χρησιμοποιούνται για τον υπολογισμό των τιμών επικινδυνότητας της κάθε απειλής. Επιπροσθέτως, εφαρμόζουμε την μεθοδολογία στο GPRS, το οποίο αποτελεί ένα καλά μελετημένο αλλά ταυτόχρονα περίπλοκο σύστημα, και πιο συγκεκριμένα μελετάμε την υποκλοπή των δεδομένων του χρήστη στο ασύρματο μέσο, η οποία θεωρείται μια από τις πιο σημαντικές απειλές της ασφάλειας του GPRS.

# Abstract

Recently, organizations around the world are becoming aware of the need to run risk management programs in order to enhance their information security. However, the majority of the existing qualitative/empirical methods fail to adhere to the terminology defined by ISO 27005 and study deliberate threats in a high level which could prove misleading for complex threats. In this work, after making a review of current work in risk analysis nd security assessment in general, a quantitative risk analysis methodology for deliberate threats is introduced. The proposed approach follows the steps suggested by the ISO 27005 standard for risk management, extending them in order to focus on deliberate threats and the different information security incidents that realize them. It is based on three-levels: the conceptual foundation level, the modeling tools level and the mathematical foundation level. The conceptual foundation level defines and analyzes the terminology involved, using unified modeling language (UML) class diagrams. The modeling tools level introduces certain tools that assist in modeling the relations among different concepts. The mathematical foundation level includes all the different mathematical formulas and techniques used to estimate risk values for each threat. Moreover, we implement this methodology on GPRS, which is a well-studied yet complex system, and more specifically on user data interception over the radio interface, which is considered one of the most important threats of GPRS security.

# Contents

# List of Figures



# List of Tables

# 1. Introduction

IT systems operate in a menacing environment that constantly changes. Thousands of new vulnerabilities are discovered every year. It is worth mentioning that CERT (*Computer Emergency Response Team*) has catalogued over 21000 vulnerabilities from 2006 to 2008 [1]. Moreover, older vulnerabilities may become easier to exploit through time, since technology advances and becomes cheaper or implemented controls are becoming inefficient, allowing attacks that where either highly unlikely or impossible to happen in the past. More and more organizations around the world are becoming aware of these facts and adopt appropriate security assessment methods, as part of a risk management program, in order to get an insight into what threatens their systems and apply appropriate controls to minimize or even eliminate the consequences that would suffer from a potential security incident.

Some of the most widely used security assessment methods use risk as a metric. Risk is the combination of the probability of an event and its consequence [2]. Currently, there are several qualitative risk assessment methods such as CRAMM [3], CORAS [4], OCTAVE [5], etc., some of which are widely used. However, while qualitative methods are simple, easy to understand and convey, and thus it is easy for different parties that are involved in decision-making to reach consensus, they suffer from some very important deficiencies many of which derive from the nature of qualitative values that are used [6][7]. For example: (i) they do not provide a basis for a cost-benefit analysis, (ii) there are insufficient differentiations between important risks, (iii) they introduce subjectivity and (iv) it is difficult for an analyst to track improvements. Moreover, different threat types are analyzed in a common level of details without taking into account special characteristics, especially in case of deliberate threats where a security incident may be realized through a series of attack events. Last but not least, qualitative methods tend to focus on business' perspective, ignoring completely attackers and are ISO-27005 incompatible reducing their applicability. However, the application of qualitative methods may prove a useful tool in order to obtain a general indication of risk and to reveal the major risks of the system or organization under examination, without too much effort or too many

details. An example of such a high-level, qualitative risk analysis for the GPRS technology is presented in [8].

Many of the above mentioned issues could be addressed through quantitative methods. However, quantifying security is a difficult task and there is a lot of on-going research into what metrics/measurements should be used, what calculations should be made, how these calculations correlate and how we can get reports out of these measurements. Recently, a few quantitative risk analysis methods have been proposed, that try to address some of the limitations of the aforementioned methods [9][10]. Although, both of these quantitative methods provide an in-depth level of risk analysis for deliberate threats, they are attacker driven only, focusing on the probability of a vulnerability being exploited rather than the consequences that an organization would suffer. Furthermore, there are ISO incompatible, limiting their use, and do not consider multiple attacker profiles.

In this work, we propose and analyze a quantitative risk analysis method suitable for in-depth analysis of complex networks by expanding, improving and optimizing some of the ideas proposed in the aforementioned methods, free of the deficiencies mentioned. The proposed approach follows the steps suggested by ISO 27005 standard for risk management [11], extending them in order to focus on deliberate threats and the different information security incidents that realize them. It is based on three-levels: the conceptual foundation level, the modeling tools level and the mathematical foundation level. The conceptual foundation level is achieved by using class diagrams of the unified modeling language (UML) [12] that follow the risk analysis terminology defined in ISO 27005 [11]. This level is further facilitated by the modeling tools and the mathematical foundation level. The proposed modeling tools help in modeling conceptually the relations among different concepts. Finally, the mathematical foundation level includes all the different mathematical formulas and techniques used to estimate risk values for each threat. Furthermore, we implement it on GPRS system [13] as a proof of concept. We chose GPRS as it is a complex, heterogeneous mobile network, but yet mature and thoroughly studied. This work can be used as an out-of-the-box solution from GPRS providers, but it can also used as a guide to implement the proposed method to any mobile or other telecommunication network. The analysis is focused on GPRS user data interception over the radio interface, which is perceived by both companies and analysts as one of the most severe threats that the network faces. Furthermore, we analyze the results of the risk

analysis process and try to draw conclusions about the main security incidents that might occur and the different threat-source (or attacker) profiles. Finally, we propose some cost-efficient, yet effective security controls that can be implemented in order to minimize the potential risk of GPRS user data interception.

# 2. Background and Previous Work

In this section, some information about current trends in security quantification and quantitative risk analysis are provided. Moreover, the GPRS Network and some of its security mechanisms related to the confidentiality of transmitted data over the radio interface are presented.

## 2.1 Security Quantification Background

The main question that led security researchers into security quantification is clear and simple: How can we improve something we cannot measure? Empirical/historical data related to security incidents could provide important information and enhance decision-making significantly. However, these data are scarce [14], since organizations are reluctant to publish information regarding the attacks on their systems, for fear that the same or similar vulnerability will be exploited by other attackers, or for fear of suffering reputational damage. Current trends in security quantification tend to split researchers into two major groups: the “modelers” and the “measurers” [15].

“Modelers” mostly focus in the mathematical formalization of the relations among the different concepts (i.e. threats, vulnerabilities, risk, probability, consequences etc.). These models get some parameter values as input and provide an output which gives an insight into a system’s or organization’s security posture. “Measurers” focus on the different ways to collect data related to security, what data should be collected and how should be analyzed in order to get information on security status. These approaches should be considered as different roads to a similar destination which is security metrics. At this point it should be pointed out what is the difference between metrics and measurements. Measurements are generated by counting while metrics are generated from analysis. In other words measurements are

objective raw data and metrics are either objective or subjective human interpretations of those data [16].

Currently there is hundreds of published research that follows either the one or the other approach. But what characterizes a good metric/measurement and what a good model? A good model according to [17] should:

- Describe a real system as accurately as possible and at the same time, it should be easy to use for system analysis. In other words, a good model should be realistic enough so that the output can give a fairly realistic description of how the system would behave under certain changes and easy to use at the same time.
- Be Verified/Cross-validated. This practically means that should be checked how the fits to empirical data and make the appropriate calibrations if necessary.

On the other hand, a good metric/measurement according to [15][16] should be:

- Consistently measured and not depend on subjective judgments. Different people should be able to apply a method to the same data and come up with the same or at least equivalent results.
- Cheap to gather, preferably in an automated way. Security metrics need to be computed frequently to help companies analyze their security effectiveness in regular basis and for this reason gathering them should be cost and time effective.
- Expressed as a number or percentage using at least one unit of measure. Cardinal numbers, which measures how many of something there, are considered the best metrics/measurements. Andrew Jaquith [6] considers ordinal numbers, which denotes the position that something is in, as bad metrics as he claims that they introduce subjectivity. This is true in many cases where high-low-medium scales are used, however ordinal numbers may not be always subjective. In many cases, the ordinal numbers are based on clearly defined criteria and different analysts can reproduce the same numbers.
- Contextually specific. Good metrics should be in context. Measurements should be meaningful in order to lead to results.

## 2.2 Previous Work on Security Quantification

In the following section we are going to critically review some of the models and metrics proposed recently and check if they comply if the aforementioned criteria for good metrics and measures.

### 2.1.1 Security Assessment Models

In 2007, Zaobin Gal et al. have presented a risk estimation methodology for information systems [9], which introduces a "through the eyes of the adversary" approach. The goal of this method is to identify, rank and report the most dangerous attack scenarios. In this work, they propose the use of an extended version of the Shneier's attack trees [18][19]. An attack tree represents the attacks that realize a threat as a tree structure. The root node represents the goal of the attack, while the leaf nodes represent the different attacks needed in order to accomplish the goal. The extended version of attack trees covers the case of attacks launched under the condition that other attack(s) preceded. This approach uses the risk metric for each attack which equals to the occurrence probability of the attack multiplied by the impact of the attack. After calculating risk for each attack involved in the realization of an attack goal, the total risk for a specific attack goal is aggregated using the extended attack tree and a set of predefined equations (i.e. one for each node type based on probability theory). However, the paper doesn't provide a way to calculate impact and in the practical example impact is not taken into account, considering risk equal to probability. Probability is defined as the weighted function of three parameters: cost, detectability and difficulty. The weight of each parameter is set to 0.4, 0.3 and 0.3 respectively without providing any reasoning for this decision. The probability parameters are set for each attack following the tables below:

**Table 1: Score Standard for Attack Cost**

| **Attack cost (AU$ 1000)** | **≥5** | **5-2** | **2-1** | **1-0.5** | **≤0.5** |
|---|---|---|---|---|---|
| **Grade** | 5 | 4 | 3 | 2 | 1 |

**Table 2: Scale of Attack Difficulty**

| **Attack Difficulty** | **Very hard** | **Hard** | **Moderate** | **Easy** | **Very easy** |
|---|---|---|---|---|---|
| **Difficulty level** | 5 | 4 | 3 | 2 | 1 |

**Table 3: Scale of Attack Detectability**

| **Probability of the attack detected** | **Very hard** | **Hard** | **Moderate** | **Easy** | **Very easy** |
|---|---|---|---|---|---|
| **Degree** | 1 | 2 | 3 | 4 | 5 |

While the authors characterize the method as quantitative, as one can notice from the table above this statement is partly true. While scoring for attack cost is based on measurements (i.e. the cost of the attack in Australian dollars), attack difficulty and detectability scoring is based on qualitative values which are not clearly defined (i.e. "very hard", "hard" etc.). This practically introduces subjectivity. Furthermore, the method is not validated with real world data.

A similar approach published in 2008 that focuses on the risk assessment of VoIP call interception [10], proposes a formal risk assessment method, which includes two modeling techniques: attack trees and vulnerability dependency graphs. While attack trees are used to model the threat under examination (i.e. VoIP call interception), the vulnerability dependency graphs present the dependencies among the identified vulnerabilities and how these vulnerabilities interact to each other. Two interesting enhancement compared to the aforementioned method are the different attack scenarios for different attacker profiles and the dependency graphs. In this paper the risk is considered a function of two parameters, the damage potential and the exploitability, but only the exploitability levels are studied for the specific case study (i.e. VoIP call interception). It should be noted that exploitability refers to the amount

of effort and expertise required in order to exploit a vulnerability. However, more than one security incidents may rise from a single vulnerability with different exploitability values and different potential loss values, so the authors should consider, instead of assigning exploitability and potential loss values to vulnerabilities, to assign the same values to security incident. After an exploitability value is assigned to each vulnerability it is checked through the vulnerability dependency graph whether or not a potential previously exploited vulnerability makes the exploitation of another vulnerability easier. If so, the exploitability value is changing accordingly. Then, the risk is aggregated through the attack tree to get the final risk value for the threat under examination. Moreover, just like the previous method there is a conversion of qualitative values for exploitability, which are not clearly defined (i.e. “very hard”, “hard” etc.), into quantitative, which introduces subjectivity. What is more, the method is not validated with empirical data.

### 2.1.2 Security Assessment Metrics

In 2005, two researchers from Carnegie Melon University proposed the use of attack surface security metric which helps determine if a software system is more secure than another [20][21]. The attack surface of a system indicates how easy is for someone to attack a system taking into account three abstract dimensions: methods (e.g. application programming interfaces), channels (e.g. sockets) and data (e.g. input strings). The bigger the attack surface of a system, the more likely the system will be attacked. The use of this metric is limited as it only calculates the relative security of two systems and not the absolute security of a single system. However, it can prove a very useful tool for choosing between two systems as regards to security.

Another attempt that is worth mentioning is the Common Vulnerability Scoring System (CVSS) [22]. The CVSS provides an open framework for communicating the characteristics and impacts of vulnerabilities. In other words CVSS offers a common language in vulnerability scoring that enables those who use it to know how important a given vulnerability is in relation to other vulnerabilities. The scoring is based in three aspects called groups: the base, the temporal and the environmental group. The base group includes the intrinsic qualities of a vulnerability (i.e. Access vector, access complexity, authentication, confidentiality, integrity and availability impact). The temporal group includes dynamically changing over time characteristics of a vulnerability (i.e. Exploitability, Redemption Level, Report Confidence). The

environmental group includes characteristics that change according to the environment that the vulnerability exists in (i.e. Collateral damage potential, target distribution, security requirements). Parameters in each group get a numeric value based on qualitative values. Qualitative values are clearly defined and in most cases they don't introduce subjectivity (you can see a sample in table 4). However, there are some cases where some qualitative value definitions use words such as "slight", "moderate", "significant" which can be interpreted differently by different people and thus introduce subjectivity (see table 5). CVSS scoring system allows skipping some of these values so that they cannot influence the total score. Following predefined equations a score from 0 to 10 is generated for each group and for the vulnerability as a whole.

**Table 4: Exploitability Scoring Evaluation in CVSS**

| Metric Value | Description |
|---|---|
| Unproven (U) | No exploit code is available, or an exploit is entirely theoretical. |
| Proof-of-Concept (POC) | Proof-of-concept exploit code or an attack demonstration that is not practical for most systems is available. The code or technique is not functional in all situations and may require substantial modification by a skilled attacker. |
| Functional (F) | Functional exploit code is available. The code works in most situations where the vulnerability exists. |
| High (H) | Either the vulnerability is exploitable by functional mobile autonomous code, or no exploit is required (manual trigger) and details are widely available. The code works in every situation, or is actively being delivered via a mobile autonomous agent (such as a worm or virus). |
| Not Defined (ND) | Assigning this value to the metric will not influence the score. It is a signal to the equation to skip this metric. |

**Table 5: Collateral Damage Potential Scoring Evaluation in CVSS**

| Metric Value | Description |
|---|---|
| None (N) | There is no potential for loss of life, physical assets, productivity or revenue. |
| Low (L) | A successful exploit of this vulnerability may result in slight physical or property damage. Or, there may be a slight loss of revenue or productivity to the organization. |
| Low-Medium (LM) | A successful exploit of this vulnerability may result in moderate physical or property damage. Or, there may be a moderate loss of revenue or productivity to the organization. |
| Medium-High (MH) | A successful exploit of this vulnerability may result in significant physical or property damage or loss. Or, there may be a significant loss of revenue or productivity. |
| High (H) | A successful exploit of this vulnerability may result in catastrophic physical or property damage and loss. Or, there may be a catastrophic loss of revenue or productivity. |
| Not Defined (ND) | Assigning this value to the metric will not influence the score. It is a signal to the equation to skip this metric. |

Andrew Jaquith in his book "Security Metrics" [15] collected over seventy-five (75) different metrics and measurements that organizations use to assess their security posture, diagnose security-related issues and measure security activities. The collected metrics are split into two main categories: Technical Metrics (i.e. metrics about perimeter defenses, coverage and control, availability and reliability, application risks) and Security Program metrics (i.e. metrics about planning and processing, acquisition and implementation, delivery and support, monitoring). He also proposes a list of common and more advance analysis techniques such as average, median, standard deviation, time series analysis, cross sectional analysis, quartile analysis and correlation matrices. While this collection of metrics complies with all the criteria we set for "good" metrics/measurements, analyzing so many metrics can be a daunting and time-consuming task. However, it can prove a great tool for analyzing specific areas of interest where a part of these metrics/measurements apply. An example of proposed technical security metrics that have to do with perimeter defense and more specifically E-mail system can be seen in table 6.

**Table 6: Perimeter Defenses Metrics – Email System**

| Metric (Unit of Measure) | Purpose | Sources |
|---|---|---|
| **E-mail** | | |
| Messages per day (number [#])<br>• Per organizational unit | Velocity of legitimate e-mail traffic; establishes baselines | E-mail system |
| Spam detected/filtered (#, percent [%]) | Indicator of e-mail "pollution" | Gateway e-mail content filtering software |
| Spam not detected/missed (#, %) | Effectiveness of content filtering software | Gateway e-mail content filtering software |
| Spam false positives (#, %) | Effectiveness of content filtering software | Gateway e-mail content filtering software |
| Spam detection failure rate (%)—not-detected plus false positives, divided by spam detected | Effectiveness of content filtering software | Gateway e-mail content filtering software |
| Viruses and spyware detected in e-mail messages (#, %) | Indicator of e-mail "pollution" | Gateway e-mail content filtering software<br>Workgroup e-mail content filtering software |

## 2.3 GPRS Overview

General Packet Radio Service (GPRS) [23] is a mobile wireless networking technology which has evolved from the Global System for Mobile Communication (GSM) to make high speed data transmission over mobile networks possible and realize the concept of mobile internet. Since GPRS is a GSM's overlay network, it adopts most of the security mechanisms of the latter. Thus, many of the security flaws discovered in GSM are also present in GPRS. GPRS' vulnerabilities and the related attacks have been adequately described and analyzed and in many cases specific countermeasures have been proposed [24][25][26]. Due to the severity of potential security incidents, often it is pointed out by researchers the need for a full transition to the more secure 3G networks. However, currently GSM/GPRS is the most widely used standard for mobile communications, holding the lion's share with approximately 3.5 billion connections worldwide [27] and network providers seems to be unwilling for a fast and probably costly transition to 3G for security's sake. The question that rises is *"At what cost?"*. Risk analysis provides the means to answer such questions.

### 2.3.1 The GPRS Network Architecture

The network architecture of GPRS [23] is presented in Figure 1. A GPRS user owns a Mobile Station (MS) that provides access to the wireless network. From the network side, the Base Station Subsystem (BSS) is a network part that is responsible for the control of the radio path. BSS consists of two types of nodes: the Base Station Controller (BSC) and the Base Transceiver Station (BTS). BTS is responsible for the radio coverage of a given geographical area, while BSC maintains radio connections towards MSs and terrestrial connections towards the fixed part of the network (core network).

The GPRS Core Network (CN) uses the network elements of GSM such as the Home Location Register (HLR), the Visitor Location Register (VLR), the Authentication Centre (AuC) and the Equipment Identity Register (EIR). HLR is a database used for the management of permanent data of mobile users. VLR is a database of the service area visited by an MS and contains all the related information required for the MS service handling. AuC maintains security information related to

subscribers' identity, while EIR maintains information related to mobile equipments identity. Finally, the Mobile Service Switching Centre (MSC) is a network element responsible for circuit-switched services (e.g., voice call) [23].

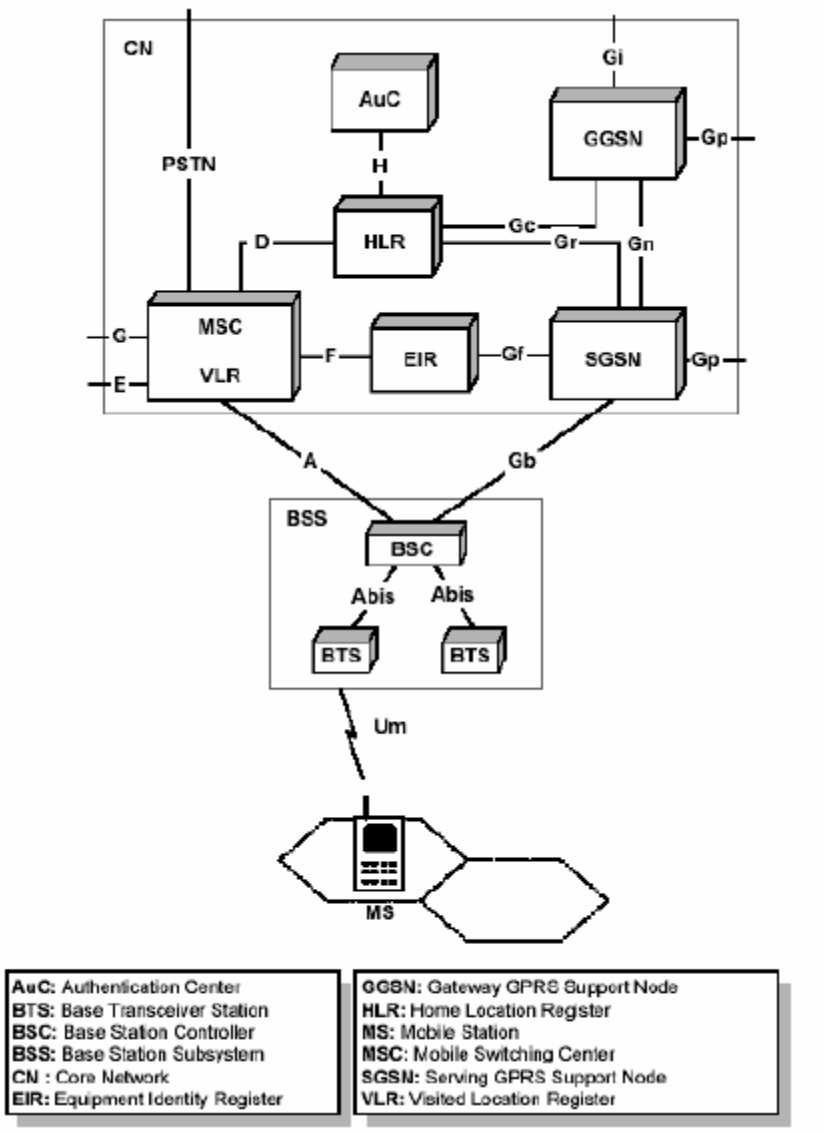


**Fig. 1: GPRS Network Architecture**

However, in order to build a packet-oriented mobile network some new network elements (nodes) are required, which handle packet-based traffic. The new class of nodes, called GPRS support nodes (GSN), is responsible for the delivery and routing of data packets between a MS and an external packet data network (PDN). More specifically, a Serving GSN (SGSN) is responsible for the delivery of data packets from, and to, a MS within its service area. Its tasks include packet routing and transfer, mobility management, logical link management, and authentication and charging functions. A Gateway GSN (GGSN) acts as an interface between the GPRS

backbone and an external PDN. It converts the GPRS packets coming from the SGSN into the appropriate packet data protocol (PDP) format (e.g., IP), and forwards them to the corresponding PDN. Similar is the functionality of GGSN in the opposite direction. The communication between GSNs (i.e., SGSN and GGSN) is based on IP tunnels through the use of the GPRS Tunneling Protocol (GTP) [28].

### 2.3.2 GPRS Mechanisms for Data Confidentiality

GPRS security is exposed into five (5) critical areas (see Figure 1) [26]: (i) the MS and the Subsriber Identity Module card (SIM-card), (ii) the interface between the MS and the SGSN, (iii) the GPRS backbone network (Gn interface), (iv) the packet network that connects different operators (Gp interface), and (v) the interface to the public Internet (Gi interface). This paper focuses on the interface between the MS and the SGSN and more specifically on the radio interface (i.e. area between MS and BTS). In the following, the security mechanisms that GPRS implements in order to protect the transmitted data over the radio interface data are described in details.

Each mobile user is personalized to the GPRS network through the use of a smart card named SIM-card [29]. The SIM-card contains a unique International Mobile Subscriber Identity (IMSI), which is the permanent identity of the user. In addition, it contains a secret key Ki (128 bit) that is used for subscriber authentication, an authentication algorithm (A3), a cipher key generating algorithm (A8), and a four digit code (Personal Identification Number – PIN) that is used to control user access to the SIM.

A mobile user that attempts to access the network must first prove his identity to it. User authentication [13] protects against fraudulent use and ensures correct billing. GPRS uses the authentication procedure already defined in GSM with the same algorithms for authentication and generation of encryption key, and the same secret key, Ki, (see Figure 2). However, from the network side, the whole procedure is executed by the SGSN (instead of the base station) and employs a different random number (GPRS-RAND), and, thus, it produces a different signed response (GPRS-SRES) and encryption key (GPRS-Kc) than the GSM voice counterpart

To achieve authentication of a mobile user, the serving SGSN must possess security related information for the specific user. This information is obtained by requesting the HLR/AuC of the home network that the mobile user is subscribed. It

includes a set of authentication vectors, each of which includes a random challenge (GPRS-RAND), and the related signed response (GPRS-SRES) and encryption key (GPRS-Kc) for the specific subscriber. The authentication vectors are produced by the home HLR/AuC using the secret key Ki of the mobile subscriber.

During authentication the SGSN of the serving network sends the random challenge (GPRS-RAND) of a chosen authentication vector to the MS. The latter encrypts the GPRS-RAND by using the A3 hash algorithm, which is implemented in the SIM-card, and the secret key, Ki. The first 32 bits of the A3 output are used as a signed response (GPRS-SRES) to the challenge (GPRS-RAND) and are sent back to the network. The SGSN checks if the MS has the correct key, Ki, and, then, the mobile subscriber is recognized as an authorized user. Otherwise, the Serving Network rejects the subscriber's access to the system. The remaining 64 bits of the A3 output together with the secret key, Ki, are used as input to the A8 algorithm that produces the GPRS encryption key (GPRS-Kc).

User data and signaling protection over the GPRS radio access network is based on the GPRS ciphering algorithm (GPRS-A5) [24], which is also referred to as GPRS Encryption Algorithm (GEA) and is similar to the GSM A5. Currently, there are three versions of this algorithm: GEA1, GEA2 and GEA3 (that is actually A5/3), which are not publicly known, and, thus, it is difficult to perform attacks on them. The MS device (not the SIM-card) performs GEA using the encryption key (GPRS-Kc), since it is a strong algorithm that requires relatively high processing capabilities. From the network side, the serving SGSN performs the ciphering/deciphering functionality protecting signaling and user data over the Um, Abis, and Gb interfaces.

During authentication the MS indicates which version(s) of the GEA supports, and the network (SGSN) decides on a mutually acceptable version that will be used. If there is not a commonly accepted algorithm, the network (SGSN) may decide to release the connection. Both the MS and the SGSN must cooperate in order to initiate the ciphering over the radio access network. More specifically, the SGSN indicates whether ciphering should be used or not (which is also a possible option) in the *Authentication Request* message, and the MS starts ciphering after sending the *Authentication Response* message (see Figure 2).

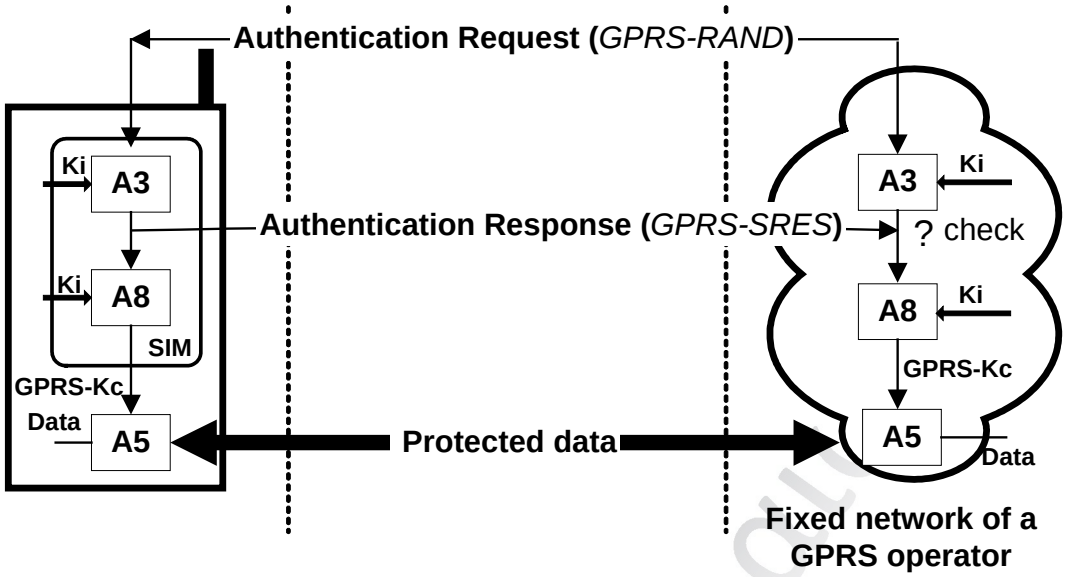


**Fig. 2: GPRS Authentication Process**

# 3. The Proposed Risk Analysis Approach

The proposed risk analysis approach for deliberate threats consists of three distinct levels of details (see figure 3). The highest is the conceptual foundation level, which defines and analyzes the terminology involved using UML class diagrams. The intermediate is the modeling tools level, which introduces certain tools that help in modeling conceptually the relations among different concepts. Finally, the bottom level is the mathematical foundation level, which includes all the different mathematical formulas and techniques used to estimate risk values for each threat.

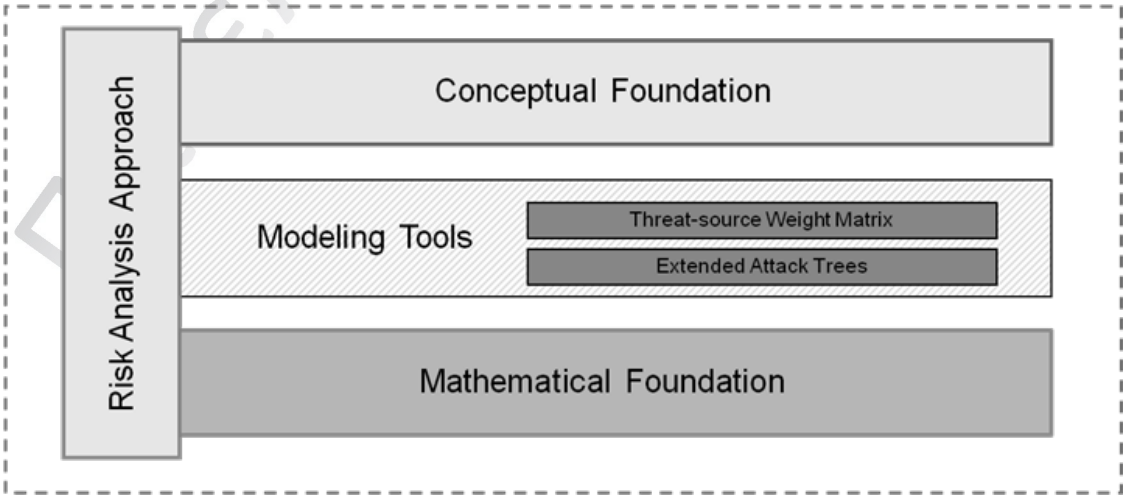


**Fig. 3: The proposed three-level risk analysis approach**

## 3.3 Conceptual Foundation

The conceptual foundation level achieves concept formalization using UML class diagrams [12]. The latter present, formally, how different concepts, involved in risk analysis, are related and which attributes of each concept participate in the risk estimation process. In order to create these diagrams, the concepts involved in the proposed risk analysis approach for deliberate threats should be identified and defined.

ISO 27005 classifies threats into three main categories: deliberate, accidental and environmental. Each of these categories is directly related to a set of concepts involved in a risk analysis process. An exception is the deliberate threats, which are related to an extra concept; the concept of "attack" (see Table 1). In the following, the concepts that are involved in the proposed risk analysis approach for deliberate threats, are defined according to the ISO 27000-series:

**Table 7: Concepts related with different threat categories**

| | Environmental | Accidental | Deliberate |
|---|---|---|---|
| Asset | ✓ | ✓ | ✓ |
| Risk | ✓ | ✓ | ✓ |
| Vulnerability | ✓ | ✓ | ✓ |
| Threat-source | ✓ | ✓ | ✓ |
| Attack | × | × | ✓ |

- Asset is "anything that has value to the organization" and which therefore requires protection.
- Threat is the potential cause of an unwanted event (i.e., an attack), which may result in harm of a system or organization.

- Vulnerability is a weakness of an asset or control (i.e., in ISO 27000-series, a control is a synonym of a countermeasure), which may be exploited by a threat. This general definition covers all threats categories. However, for deliberate threats, vulnerability is a weakness of an asset or control, which may be exploited by an attack to realize a threat.
- Risk is the combination of the probability of an event and its consequence.
- Attack is an attempt to destroy, expose, alter, disable, steal or gain unauthorized access to or make unauthorized use of an asset.
- Threat-source is anyone whose intention is to exploit an asset's vulnerability, launching an attack and thus, realizing a threat. Threat-source is a synonym of an attacker.

Figure 4, presents the UML class diagram of the concepts defined above using three different types of relations: association, aggregation and composition [12]. Association is represented with a simple line between two classes and denotes a simple relationship between two classes. Aggregation is represented with a transparent diamond shape and denotes a part-whole or part-of relationship between two classes. Finally, composition is represented with a solid diamond shape and denotes a strong life-death relationship between classes. Notation at the ends of each relation in the diagram is called multiplicity and indicates the number of objects that participate in the relation. For example, in figure 4 we can see that a threat may harm one or more (1…*) assets. On the other hand, an asset might be at risk by zero or more threats (0…*).

As illustrated in the UML diagram, deliberate threats are realized through information security incidents which involve the occurrence of one or more attacks. The latter exploit one or more of the asset's vulnerabilities to realize threats, and thus, harm the assets. The self-association of the attack class represents attacks series (i.e., a sequence of attacks) realizing one or more threats. Attack series may be either dependent events (i.e., occurrence of one affects occurrence of another) or independent events (i.e., occurrence of one does not affect occurrence of another).

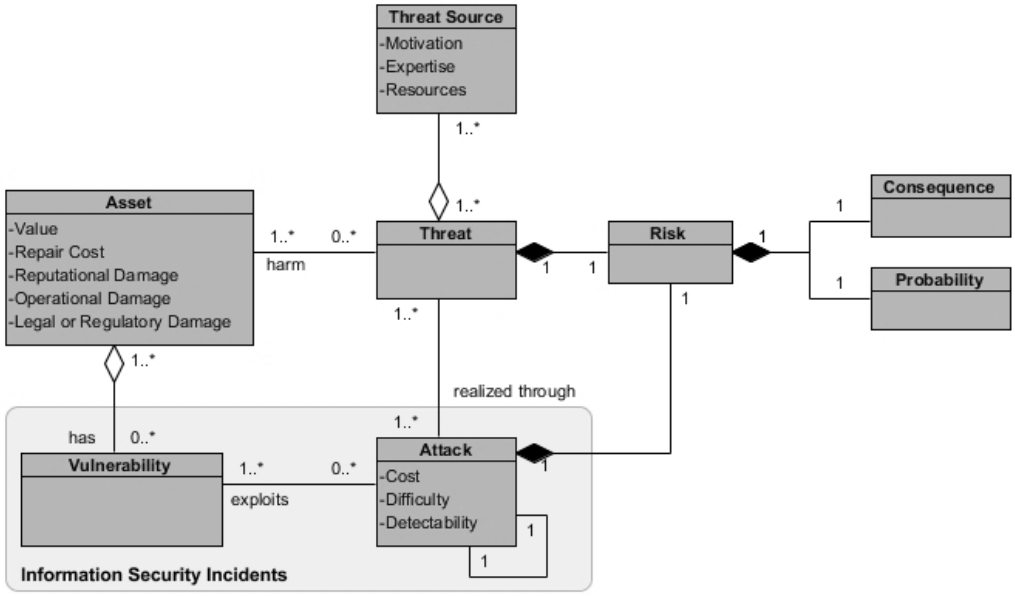


**Fig. 4: UML class diagram for the proposed risk analysis approach**

A threat may harm one or more assets and is related with one or more threat-sources. The aggregation relation between these two classes (i.e., threat and threat-source) denotes that if a threat is removed, then the same threat-sources may still exist for other threats. The risk analysis process estimates a risk value for each identified threat. This value is related to the probability of each threat to be realized and the corresponding consequences that occur. As mentioned previously, deliberate threats may be realized through one or more attacks, each of which has its own probability of occurrence and consequences, and, thus, its own risk value (Figure 4). The estimated risk value of a threat equals to the maximum risk value of all single attacks or series of attacks, which realize the threat. Consequently, risk values of the identified attacks and series of attacks should be estimated prior to estimating the risk value of the related threat. Since each threat is related with a unique risk value, if this threat is removed, then the corresponding risk value no longer exists. To represent this life-death dependency between the threat class and the risk class, as well as between the risk class and the attack class, the composition relation is employed.

Each of the concepts identified and modeled formally using the UML class diagrams, has certain attributes that should be considered during the risk estimation process. Some of these attributes are related to the probability of occurrence of a threat; while others are related to the consequences following the threat occurrence. However, not all of these attributes are always relevant with the threat under examination, and thus, each case should be studied separately. The attributes, which

are included in the modeled concepts of the proposed approach and consider both business and attackers' perspectives, are:

- Threat-source
  - Motivation: what motivates a particular threat-source (or attacker) to launch an attack.
  - Expertise: the level of knowledge of a particular threat-source (or attacker) related to an attack.
  - Resources: the resources (money, equipment) that a particular threat-source (or attacker) has in its possession.
- Attack
  - Cost: the cost of equipment needed to launch an attack.
  - Difficulty: the level of expertise needed for someone to launch an attack.
  - Detectability: the easiness or difficulty of an attack being detected.
- Asset
  - Value: the value of a specific asset (it may also be considered as the cost of replacement).
  - Repair Cost: the cost to repair an asset.
  - Reputational Damage: the damage in reputation occurs if an asset is compromised.
  - Operational Damage: the damage occurs in an organization's or system's operation due to the compromised asset.
  - Legal and Regulatory Damage: the fines and penalties that will be paid because of the compromised asset.

## 3.4 The Modeling Tools

The modeling tools used in the proposed risk analysis include: (i) the threat-source profile matrix and (ii) the extended attack trees. The chosen tools link the conceptual with the mathematical foundation level of the proposed approach, as explained bellow.

**The Threat-source Profile Matrix.** The threat-source profile matrix is a two-dimensional matrix, which contains the weights of the attributes involved in an attack probability estimation (i.e. cost, difficulty, detectability) for all different threat-sources. The UML class diagram of Section 3.3, shows that a threat is related with one or more threat-sources. Each threat-source has each own motivation, resources and expertise level and thus there is a different probability for each threat-source exercising a specific attack. For example, the high cost of an attack wouldn't be for a professional hacker as deterrent as it would be for a script kiddy. In other words, while the cost attribute does not carry too much weight in the attack probability

estimation for a professional hacker, it does for script kiddies which have limited resources. In order to reflect this diversity in probability values for different threat sources, profiles are created by assigning weight values for each attribute taking part in an attack probability estimation.

**The Extended Attack Trees.** Attack Trees [30] [31] represent a formal method of representing the varying of attacks that a system is exposed to, using a tree structure. The root node of the tree symbolize an identified threat, while the leaf nodes stand for the information security events (single attacks or attacks series) that realize the specific threat. The intermediate nodes of the tree can be either alternative subgoals, each one satisfying the parent goal (OR Nodes), or partial subgoals, whose composition satisfies the parent goal (AND Nodes). In a compound system there are several threats and consequently, attack trees, which form an attack forest. Attack trees can be illustrated both graphically and textually. However, graphical representation is not appropriate for composite systems, due to the enormous size that the tree could reach. In the proposed risk analysis approach, an extended version of the attack trees [9], which incorporates the CAND (Conditional AND) node is used. The classic attack trees cannot formally represent all the previously described information security incidents. Although the AND nodes of a tree can be used for depicting attacks series of independent events, they cannot be used for attacks series of dependent events, where the attacks occur under certain occurrence conditions. This is achieved by adding the CAND node (i.e., extended attacks trees). The CAND relation between nodes represent that the upper node is accomplished if all sub-nodes are attained under certain conditions.

## 3.5 Mathematical Foundation

As mentioned previously, the risk is the combination of the probability of an attack event and its consequences. In this approach, the risk value of an attack is derived by multiplying the probability of occurrence value of the attack with the estimated consequences, as shown in equation (1):

$$Risk(Attack) = P(Attack) \times C(Attack) \quad (1).$$

In order to estimate the attack probability value, utility curves from the multi-attribute utility theory are adopted that convert the attribute values into utilities. In the current approach, the utility curve chosen is the $\frac{1}{X}$. We chose this utility function because the probability attributes are in inverse proportion with the probability itself. Furthermore, it can accurately represent residual risk as the probability can never become equal to zero. Risk can only become equal to zero if the consequences are equal to zero or the corresponding vulnerability is removed. Each utility is then multiplied by the corresponding weight of the threat-source under examination and summed up to the probability value, as shown in equation (2):

$$P(Attack) = W_{cost} \times U(cost) + W_{diff} \times U(diff) + W_{dete} \times U(dete) \quad (2),$$

Where:

cost = Cost of an attack,

diff = Difficulty of an attack,

dete= Detectability of an attack,

Wcost = weight of the attack cost for a specific threat-source,

Wdiff = weight of the attack difficulty for a specific threat-source,

Wdete = weight of the attack detectability for a specific threat-source,

U(x) = utility function of the attributes.

The consequences of an attack equal to the sum of the related attribute values, as shown in equation (3). These attributes are: the asset value, the repair cost, the reputational damage, the operational damage and the legal damage, as defined in section 3.3.

$$C(Attack) = assetValue + repaCost + repuDam + OperDam + LegalDam \quad (3).$$

In order to estimate the risk of a specific threat, a risk aggregation over the constructed attack tree is required. Starting from the leafs and moving toward the root of the tree, the total risk value is aggregated according to the following:

In the OR nodes, the total risk value equals to the maximum risk value of its sub-nodes (SubN), as shown in equation (4).

$$Risk(N) = MAX(Risk(SubN_1), Risk(SubN_2), \ldots, Risk(SubN_i)) \quad (4).$$

In the AND and CAND nodes, the total risk value equals to the product of the joint probability of the sub-node events and the total consequences of the sub-node events (equation 5).

$$Risk(N) = P(SubN1 \cap SubN2 \cap \ldots \cap SubNi)\ x$$
$$C(SubN1 \cap SubN2 \cap \ldots \cap SubNi) \quad (5),$$

Where:

$$C(SubN_1 \cap SubN_2 \cap \ldots \cap SubN_i) = C(SubN_1) + C(SubN_2) + \cdots + C(SubN_i) \quad (6).$$

The joint probability of the sub-node events for independent attack series events equals to the product of the probabilities of each independent attack event (equation 7). On the other hand, the joint probability of the sub-node events for order-dependent attack series equals to the product of the probabilities of each attack event, in series, given the preceding events (equation 8).

$$P(SubN_1 \cap SubN_2 \cap \ldots \cap SubN_i) = P(SubN_1)x\ P(SubN_2)x \ldots xP(SubN_i) \quad (7).$$

$$P(SubN_1 \cap SubN_2 \cap \ldots \cap SubN_i) =$$
$$P(SubN_1)\ x\ (SubN_2|SubN_1)x \ldots x(SubN_i|SubN_{i-1} \ldots SubN_2SubN_1) \quad (8).$$

# 4. Risk Analysis Methodology: Step-by-step

The purpose of risk identification is twofold: (i) to determine what might happen causing potential loss, and (ii) to gain insight into how, where and why the loss occurs. To achieve this, both the business and the attacker's point-of-view should be

taken into consideration. The risk identification process consists of the following five (5) steps:

## 4.3 Risk Identification Process

**Step 1: Asset Identification.** In this step, anything that is important to the organization should be considered. This includes both primary assets (i.e., such as business processes/activities or information) and secondary (i.e., such as hardware, software, network, personnel, site and organization's structure). Assets' identification can be performed in various levels of details. However, the most appropriate is the one that provides sufficient information for the risk estimation process, which follows risk identification. However, since risk analysis is a recurrent procedure, the level of detail can be changed, accordingly, in further iterations of the risk analysis process.

**Step 2: Threat Identification.** In this step, anything that threatens assets and originates from deliberate threat-sources should be identified. These threats may arise either from inside or outside the organization. Threats should be identified as general as possible and elaborated further (i.e, going into a greater level of details), where appropriate. For every identified threat, possible threat-sources should be defined. Moreover, for each threat-source, a profile should be created giving weights to each attack attribute, using the threat-source profile matrix described in section 3.4.

**Step 3: Existing Controls Identification.** In this step, existing controls, if any, are identified in order to avoid unnecessary work in the next steps of the risk analysis process. According to ISO 27000, a control is the synonym of a countermeasure. Controls may reduce, minimize or even abolish the risk of a potential threat. Furthermore, in this step the efficiency of the existing controls should be verified. In many cases controls does not work as expected, creating new vulnerabilities, which should be treated either by replacing them or by implementing complementary controls.

**Step 4: Vulnerability Identification.** In this step, the vulnerabilities that may harm assets should be identified. Vulnerabilities may exist in an organization, processes and procedures, management routines, personnel, physical environment, information system configuration, hardware, software or even related external parties.

**Step 5: Information Security Incident Identification and Identification of Corresponding Consequences.** This step gets as input the identified assets, threats and vulnerabilities of the previous step, and identifies the entire set of information security incidents, related to the identified threats. Information security incidents fall into three main categories: single attacks, independent attacks series events (i.e., occurrence of one does not affect occurrence of another), and dependent attacks series events (i.e., occurrence of one affects occurrence of another). Furthermore, the consequences that will occur by a security incident should be identified in terms of asset value, repair cost, reputational damage, operational damage and legal damage.

## 4.4 Risk Estimation Process

This process estimates, quantitatively, the risk of each threat using the tools and the mathematical formulas, described in sections 3.4 and 3.5, respectively.

**Step 1: Assigning Values to Probability and Consequences Attributes.** In this step specific values are assigned to each attribute related to the probability of occurrence and consequences of the identified attacks. In case of order-dependent attacks series, the attack should be examined as part of a sequence of events.

**Step 2: Mapping Information Security Incidents with Threats.** This step involves the construction of an attack forest. For each threat, a separate attack tree is constructed, as described in section 3.4. Extra nodes that represent intermediate system states or sub-threats should be added where necessary.

**Step 3: Aggregating Risk using Attack Trees.** In this step, the risk is aggregated from the leafs to the root of a tree, using the formulas described in 3.5.

# 5. Implementing the Methodology on the GPRS Network

In this section, we provide a simple example where the proposed risk analysis approach is applied as a proof of concept in the GPRS Network.

## 5.3 Asset & Threat Identification

Assets in GPRS include subscriber identity, user data and network access. In current work, we focus on user data and more specifically user data transmission over the radio interface. The threat under examination is the interception of these data over the radio interface. For the purposes of the analysis we created threat theat-source profiles: Professionals, Hackers/Crackers and Script Kiddies. Professionals have a high level of expertise and plenty resources, however they are reluctant to launch attacks that can be detected and probably compromise their identity. Hackers/Crackers on the other hand while they have the same level of expertise they do not have the same amount of resources and thus the attacks they launch are limited to only those that are within their budgetary constraints. Script Kiddies only launch attacks that are fully documented and cheap. They are also unaware about the detectability of the attack their launching. According to the profiles we just described we created the threat-source profile matrix shown in table 8.

**Table 8: Threat-source Profile Matrix**

| | Professionals | Hackers/Crackers | Script Kiddies |
|---|---|---|---|
| Wcost | 0.1 | 0.4 | 0.6 |
| Wdetectability | 0.6 | 0.3 | 0.1 |
| Wdifficulty | 0.3 | 0.3 | 0.3 |

## 5.4 Existing Controls Identification

For the radio interface of GPRS it is assumed that no extra security control has been applied and only the GPRS standard security mechanisms for user data confidentiality are present, as described in 2.3.2.

## 5.5 Vulnerability Identification

There are plenty well-known vulnerabilities in GPRS that can be exploited by an attacker in order to realize the threat under examination. An important vulnerability of the GPRS authentication procedure is that there is no mutual authentication (one-way authentication). During authentication only MS is authenticated to the network while there is no sufficient mechanism that authenticates the network to the MS. In addition, the absence of a mechanism that ensures data integrity over the radio access network makes active attacks possible. Another basic weakness of the GPRS security architecture is that encryption of user and signaling data over the radio interface is optional and in some countries GPRS operators never switch on encryption in their networks. In these cases, data are conveyed in clear-text exposing them to potential attacks. If encryption is switched on, then during authentication the MS and SGSN indicate which type of encryption they support. However, no encryption or data integrity mechanism is employed in authentication messages exchange. This may lead to either the modification of the MS and the network capabilities regarding encryption, or the suppression of encryption over the radio interface.

## 5.6 Information Security Incident Identification and Identification of Corresponding Consequences

Exploiting the aforementioned vulnerabilities there are three security incidents that may occur: The class-mark attack (single event), the man-in-the-middle attack (dependent event series) and the GPRS-Kc retrieval by exploiting GSM network (dependent event series). GPRS user data interception incidents may have consequences related to reputational damage and legal damage due to fines. Since there is no asset destruction, we do not consider asset value or repair cost. Moreover, network operation is not affected by interception attacks and thus there is no operational damage.

**The Class-mark Attack (Incident 1).** By compromising the integrity of specific signaling/control data, an attacker may be able to suppress the encryption over the radio interface. The Class-Mark message that an MS sends to the network at the

beginning of a session to indicate its encryption capabilities, might be modified by an intermediate, so that the network is convinced that MS does not support encryption. The intermediate impostor may transmit a hoax class-mark message (E1) at the same time that the victim's MS transmits the original, but using a much stronger radio signal. Thus, the attacker's signal overrides the original message at the BTS**.**

**The Man-in-the-middle Attack (Incident 2).** When an attacker intervenes between the network and the MS providing the requested services (by the user), the so-called man-in-the-middle attack is launched. First the attacker masquerades as a BTS and impersonates the GPRS network to the victim MS (E2). In the subsequent authentication process, the attacker is authenticated to the network using its own subscription discarding the MS authentication data (E4). Then, the attacker requests to turn off the encryption (Cipher Mode Command) between the MS and the false base station (E3).

**Retrieving GPRS-Kc (Incident 3).** An attacker may eavesdrop on GPRS-RAND, sent by the network to MS under attack, for authentication purposes. Then, the attacker may impersonate the voice network (E5) (GSM) that initiates a radio session with MS, and starts the authentication procedure by "replaying" the eavesdropped GPRS-RAND (as a GSM-voice RAND) (E6). After the authentication completion, the attacker asks the MS to start encrypting with A5/2 (E8) or A5/1 (E9), which are the weaker versions of the A5 algorithm, sending the appropriate Cipher Mode Command (A7). After receiving a few milliseconds of encrypted voice traffic and performing one of the well-known attacks against the GSM's algorithms, the attacker is able to recover the corresponding encryption key, Kc. Since the retrieved Kc equals to the GPRS-Kc, the attacker is able to decrypt the GPRS traffic exchanged between the MS under attack and the legitimate network. Alternatively, the attacker may record the exchanged traffic, and carry out the impersonation attack to retrieve GPRS-Kc later on. Since many network operators rarely trigger reauthentications and use the same keys for a relatively long time, the attacker may use the retrieved key to intercept more than one session.

## 5.7 Assigning Values to Probability Attributes and Consequences Attributes

The values we assign to information security events' attributes that are related to probability (i.e., cost, detectability, difficulty) are in the scale from 1 to 4 and are given according to tables 9, 10 and 11. The values for attack detectability and difficulty are semi-quantitative in nature as they are based on qualitative characteristics. The *"quality"* in each value is carefully selected so that it cannot be interpreted in different ways and thus, avoids introducing subjectivity. On the other hand, the values for attack cost are quantitative based on euro values.

Attack difficulty values (see Table 9) are based on the assumption that a better documented attack is easier to be launched and thus, its occurrence probability is higher than a worse documented attack. The identified qualities include: reported attacks (i.e., attacks that are only announced with no further information), reported attacks with rough description (i.e., attacks that are not described in detail and only basic information about attack execution are given), reported attacks with detailed steps (i.e., attacks that are described step-by-step and in great detail) and reported attacks with detailed steps and available tools (i.e., attacks of the previous category that also have tools that facilitate the execution of the attack, these tools can be software or/and hardware).

**Table 9: Attack Difficulty Values**

| **Attack Difficulty** | |
|---|---|
| Attack reported, detailed steps & widely available tools | 1 |
| Attack reported, detailed steps | 2 |
| Attack reported, rough description | 3 |
| Attack reported | 4 |

Different *"qualities"* in the attack detectability table (Table 10) try to capture different detection cases. They are based on the assumption that attackers are reluctant to launch attacks that can be detected or can compromise their identity. The different

identified qualities include: attacks that cannot be detected, attacks that can be detected but not the involved attacker, attacks in which under certain conditions both the attack and the attacker can be detected, and attacks in which both the attack and attacker are detected.

**Table 10: Attack Detectability Values**

| **Attack Detectability** | |
|---|---|
| Cannot be detected | 1 |
| Attack is detected but not Attacker | 2 |
| Attack and Attacker can be detected under certain conditions | 3 |
| Both Attack & Attacker are detected | 4 |

The attack cost table (Table 11) includes four different cost groups in euro: 0-100, 100-500, 500-1000, 1000+. This table may require adjustments to reflect what is considered as expensive or cheap in each country. An index that can be used for such decisions can be the per capita income of each country.

**Table 11: Attack Cost Values**

| **Attack Cost** | |
|---|---|
| 0 – 100 euro | 1 |
| 100 – 500 euro | 2 |
| 500 – 1000 euro | 3 |
| Over 1000 euro | 4 |

In tables 12 and 13, we assign values according to the previously described tables (i.e., table 9, 10 and 11) to each attribute and for each event taking part in the information security incidents, identified in step 5 of the risk identification process. It is worth noting that in attacks series incidents, not all events necessarily contribute to the overall probability, as they may not add extra cost, difficulty or enhance the detectability of the attack.

Incident 1 (i.e. class-mark attack), requires special equipment and relatively high power by the attacker so that the forged message can override the original. The equipment cost is estimated between 500 - 1000 euro. Incident 1 can only be detected under the condition that providers log unencrypted connections as suspicious or the mobile device warns its user with a message or special symbol. Incident 1 has been roughly described and there are not any guides or available tools that facilitate the attack.

Incident 2 (i.e., man-in-the-middle attack) requires special equipment often referred as IMSI-catcher or GSM/GPRS interceptor. These devices are nothing more that notebooks equipped with special software and peripherals. The cost of such devices is above 1000 euro. As regards to detectability, just like incident 1, the incident can only be detected under the condition that providers log unencrypted connections as suspicious or the mobile device warns its user with a message or special symbol. However, in case that the incident is detected, the attacker might also be detected since the attack requires the use of a legitimate device in order o forward traffic to the network. The attack has been analyzed in details and there are available tools. However, these tools are only sold for lawful interception purposes and selling them to public is strongly prohibited in most countries.

Incident 3 requires similar and in many cases the same devices as those described in incident 2, depending on the features provided by the devices. The main difference between the two incidents is that there is no way for the network to detect incident 3. However, there might be an increased level of difficulty in case that A5/1 is used as encryption algorithm, instead of the weaker A5/2.

**Table 12: Incident 1 & 2 attributes value assignment**

| | **Incident 1** | **Incident 2** | | |
|---|---|---|---|---|
| | E1 | E2 | E3\|E2 | E4\|E2E3 |
| **Cost** | 3 | 4 | 1 | 1 |
| **Detectability** | 1 | 1 | 1 | 3 |
| **Difficulty** | 3 | 2 | 1 | 1 |

**Table 13: Incident 3 attributes value assignment**

| | Incident 3 | | | | |
|---|---|---|---|---|---|
| | E5 | E6\| E5 | E7\|E5E6 | E8\|E7E6E5 | E9\|E7E6E5 |
| **Cost** | 4 | 1 | 1 | 1 | 1 |
| **Detectability** | 1 | 1 | 1 | 1 | 1 |
| **Difficulty** | 2 | 1 | 1 | 1 | 3 |

Due to the fact that we do not have access to economic data related to reputational damage or legal damage for a GPRS system, we assume that the consequences of the threat under examination of each information security incident are constant and equal to C.

## 5.8 Mapping Information Security Incidents with Threats

Drawing the information security incidents to the single threat under examination, we provide the following extended attack tree:

**Threat: GPRS User Data Interception over the Radio Interface**
**OR** 1. To Suppress Encryption
1.1 To Manipulate Signaling
**OR** 1.1.1 To Modify Class-mark Message (E1)
1.1.2 To Launch Man-in-the-Middle Attack
**CAND** 1.1.2.1 To Masquerade as BTS (E2)
1.1.2.2 To Send Modified Cipher-mode command (E3)
1.1.2.3 To Reroute Intercepted data via a Legitimate device(E4)
2. To Retrieve Encryption Key (GPRS-Kc)
2.1 To exploit GSM network
**CAND** 2.1.1 To Masquerade as a BTS (E5)
2.1.2 To Eavesdrop and Replay GPRS-RAND as GSM-RAND (E6)
2.1.3 To Send Modified Cipher-mode Command (E7)
2.1.4 Employ Attack on Weak Encryption
**OR** 2.1.4.1 Attack on A5/2 (E8)
2.1.4.2 Attack on A5/1 (E9)

## 5.9 Aggregating Risk

Using equation (2) of Section 3.5, we calculated the probability of each information security event for all threat-source profiles. The results are presented in tables 14 & 15.

**Table 14: Probability of Occurrence Values for Different Threat Profiles for Incidents 1 & 2**

| | **Incident 1** | **Incident 2** | | |
|---|---|---|---|---|
| | P(E1) | P(E2) | P(E3\|E2) | P(E4\|E2E3) |
| **Professionals** | 0.73 | 0.775 | 1 | 0.6 |
| **Hackers/Crackers** | 0.53 | 0.55 | 1 | 0.8 |
| **Script Kiddies** | 0.4 | 0.4 | 1 | 0.93 |

**Table 15: Probability of Occurrence Values for Different Threat Profiles for Incident 3**

| | **Incident 3** | | | | |
|---|---|---|---|---|---|
| | P(E5) | P(E6\| E5) | P(E7\|E5E6) | P(E8\|E7E6E5) | P(E9\|E7E6E5) |
| **Professionals** | 0.775 | 1 | 1 | 1 | 0.8 |
| **Hackers/Crackers** | 0.55 | 1 | 1 | 1 | 0.8 |
| **Script Kiddies** | 0.4 | 1 | 1 | 1 | 0.8 |

Starting from the leafs of the tree and moving to the root we calculate the probability value in each node using equations (4) and (7) for each node respectively. The results are presented in the following tree.

**Threat: GPRS User Data Interception over the Radio Interface** (P:0.775 H/C :0.55 S:0.4)
**OR** 1. To Suppress Encryption (P:0.733 H/C :0.533 S:0.4)
1.1 To Manipulate Signaling (P:0.733 H/C :0.533 S:0.4)
**OR** 1.1.1 To Modify Class-mark Message (P:0.733 H/C :0.533 S:0.4)
1.1.2 To Launch Man-in-the-Middle Attack (P:0.465 H/C :0.44 S:0.373)
**CAND** 1.1.2.1 To Masquerade as BTS (P:0.775 H/C :0.55 S:0.4)
1.1.2.2 To Send Modified Cipher-mode command (P:1 H/C : 1 S: 1)
1.1.2.3 To Reroute data via a Legitimate device (P:0.6 H/C :0.8 S:0.93)
2. To Retrieve Encryption Key (GPRS-Kc) (P:0.775 H/C :0.55 S:0.4)
2.1 To exploit GSM network (P:0.775 H/C :0.55 S:0.4)
**CAND** 2.1.1 To Masquerade as a BTS (P:0.775 H/C :0.55 S:0.4)
2.1.2 To Eavesdrop and Replay GPRS-RAND as GSM-RAND (P:1 H/C : 1 S: 1)
2.1.3 To Send Modified Cipher-mode Command (P:1 H/C : 1 S: 1)
2.1.4 Employ Attack on Encryption (P:1 H/C : 1 S: 1)
**OR** 2.1.4.1 Attack on A5/2 (P:1 H/C : 1 S: 1)
2.1.4.2 Attack on A5/1 (P: 0.8 H/C : 0.8 S: 0.8)

# 6. Analyzing the results

As presented in table 16, there is a high risk for the threat under examination from professional hackers, but relatively low risk from hackers/crackers, (due to the high cost of the attacks) and script kiddies (because of both: the high cost of the attacks and the high level of expertise required). However, it should be noticed that the estimated risk is more likely to increase for all threat-sources in the future, since technology advances and becomes cheaper. Professionals have a clear preference in low detectability attacks. Hackers/Crackers and Script Kiddies have a balanced preference among the three studied attacks. Moreover, one may notice that GPRS user data interception over the radio interface is more likely to be realized through incidents 1 and 3, rather than incident 2. This happens because of the fact that incident 2 presents high cost and detectability. Incident 3, on the other hand, although costs more than the other incidents, it is the most likely to occur by all different threat-source profiles since it is well-documented with low detectability. Incident 1 is following closely incident 3 as regards to risk, except for script kiddies profile where the risk values are equal.

Table 16: Risk for each Incident and each threat profile

| | R(Incident 1) | R(Incident 2) | R(Incident 3) |
|---|---|---|---|
| **Professionals** | 0.73 x C | 0,465 x C | 0,755 x C |
| **Hackers/Crackers** | 0.53 x C | 0,44 x C | 0,55 X C |
| **Script Kiddies** | 0.4 x C | 0,37 X C | 0,4 X C |

By abolishing the use of the weak A5/2 and the use of non-encrypted sessions, the risk can be reduced significantly in a cost-effective way, as it is presented in table 17. Incident 1 disappears completely, as it involves a non-encrypted session between the network and the victim's MS, which is no longer possible. In addition, the risk of incident 3 is reduced significantly, as the weak A5/2 is no longer exploitable and the only way to launch the attack is by exploiting the stronger A5/1. On the other hand, the risk of incident 2 remains unchangeable, although it involves a non-encrypted session. This is because the non-encrypted session takes place between a fake BTS and the victim's MS, while the network is not participating in encryption algorithm negotiation.

Table 17: Risk Improvement by implementing specific security controls

| | R(Incident 1) | R(Incident 2) | R(Incident 3) |
|---|---|---|---|
| **Professionals** | - | 0,465 x C | 0,62 x C |
| **Hackers/Crackers** | - | 0,44 x C | 0,44 X C |
| **Script Kiddies** | - | 0,37 X C | 0,32 X C |

# 7. Conclusions & Suggestions for Future Research

In current work we proposed an ISO 27005-compatible, quantitative risk analysis method focused on deliberate threats, and implemented it on GPRS. More specifically, the conceptual and mathematical foundations of the approach, as well as

the tools that facilitate the process of risk estimation were elaborated. The entities and attributes that take part in the risk analysis were defined and represented, graphically, using UML class diagrams. The tool of threat-source profile matrix is introduced in order to get insights into who and how is more likely to attack to the system under examination. The specific steps of the proposed approach are defined and analyzed, in details, and all the necessary mathematical functions are explained. Furthermore, we examined the risk related to the GPRS user data interception over the radio interface. We identified potential threat-sources and created their profiles, and we set semi-quantitative values for probability attributes, which are clearly defined and minimize subjectivity.

The results show that user data interception over the radio interface is more likely to occur by professional threat-sources and that by implementing simple and cost-effective security controls we can reduce the risk significantly for all threat-source profiles. However, the potential risk still remains high, especially if we consider the consequences that will follow such an event (i.e., great reputational damage and enormous fines). The results, although indicate the need to further minimize the risk by implemented security controls or by moving on to 3G networks, should not be considered completed since they do not include consequence attribute values as we didn't have access to such economic data. However, it can be used as an out-of-the-box solution for GSM/GPRS providers, which have access to such data as we provide the attributes that should be considered. In future work, we suggest to further develop the proposed method to include automatic security controls suggestion compatible with those defined in ISO 27001 [32]. Furthermore, as risk is not static, it is recommended to enhance the method so that it can provide projection of a threat's risk in a future time.

# References


1. Computer Emergency Response Team (CERT), Carnegie Mellon University, Cert Statistics (Historical), http://www.cert.org/stats/

2. International Organization for Standarization, ISO/IEC 27000, "Information technology - Security techniques - Information security management systems - Overview and vocabulary", 2009

3. CRAMM User Guide, Version 5.0 & 5.1, http://www.cramm.com/, (2005).

4. The CORAS method, http://coras.sourceforge.net/

5. OCTAVE Information Security Risk Evaluation, http://www.cert.org/octave/

6. Landoll J. D.: The Security Risk Assessment Handbook – A complete guide for performing security Risk Assessments. Auerbach Publications, pp. 424-425, (2006).
7. Microsoft solutions for security and Compliance & Microsoft Security Center of Excellence: The Security Risk Management Guide. Microsoft Corporation, pp. 20-21, (2006).
8. Xenakis C., Apostolopoulou D., Panou A., Stavrakakis I..: A Qualitative Risk Analysis for the GPRS Technology. In Proc. IEEE/IFIP International Conference on Embedded and Ubiquitous Computing (EUC-08), Shanghai, China, pp. 61-68 (2008).
9. Zaobin, G., Jiufei T., Ping W., Vijay V.: A Novel Security Risk Evaluation for Information Systems, Proceedings of the 2007 Japan-China Joint Workshop on Frontier of Computer Science and Technology, pp 67-73, (2007).
10. Benini, M., Sicari, S.: Assessing the risk to intercept VoIP calls, Journal of Computer Networks, vol 52, issue 12, pp. 2432-2446, (2008).

11. International Organization for Standardization (ISO), ISO/IEC 27005: Information technology – Security techniques – Information security risk management, (2008).

12. Object Management Group (OMG), Unified Modeling Language Specifications, http://www.omg.org/technology/documents/modeling_spec_catalog.htm#UML

13. 3GPP TS 23.060 (V9.0.0): GPRS Service Description, Stage 2, 2009

14. Moore, P. A., Ellison, J.R., Linger, C.R., "Attack Modeling for Information Security and Survivability", Carnegie Mellon University, Technical Note, 2001.

15. Jaquith A.: Security Metrics – Replacing Fear, Uncertainty, and Doubt, Addison Wesley, (2007).

16. Payne S.: A guide to security metrics, Technical Report, SANS Institute, (2006)

17. McLaughlin P. M., "..the very game…- A tutorial on mathematical modeling", http://www.causascientia.org/math_stat/Tutorial.pdf

18. Shneier, B.: Attack Trees: Modeling security threats. Dr. Dobb's Journal, http://www.schneier.com/paper-attacktrees-ddj-ft.html, (1999).

19. Shneier, B.: Secrets & Lies: Digital Security in a Networked World. John Wiley & Sons, (2000).

20. Pratyusa Manadhata, Jeannette M. Wing, An Attack Surface Metric, CMU-CS-05-155, Carnegie Mellon University, http://reports-archive.adm.cs.cmu.edu/anon/2005/CMU-CS-05-155.pdf , (July 2005).

21. Pratyusa K. Manadhata, Kymie M. C. Tan, Roy A. Maxion, Jeannette M. Wing, An Approach to Measuring a System's Attack Surface, CMU-CS-07-146, Carnegie Mellon University, http://reports-archive.adm.cs.cmu.edu/anon/2007/CMU-CS-07-146.pdf , (August 2007).

22. Mell P., Scarfone K., Romanosky S., "A Complete Guide to the Common Vulnerability Scoring System Version 2.0", http://www.first.org/cvss/cvss-guide.pdf , (June 2007).

23. 3GPP TS 23.060 (V9.0.0): GPRS Service Description, Stage 2, (2009).

24. Xenakis C.: Security Measures and Weaknesses of the GPRS Security Architecture. International Journal of Network Security, Vol.6, No.2, pp:158–169, (2008).

25. Xenakis C.: Malicious actions against the GPRS technology. Computer Virology, Springer, Vol. 2, No. 2, pp. 121-133, (2006).

26. Xenakis C., Merakos L.: Vulnerabilities and Possible Attacks against the GPRS Backbone Network. In Proc. International Workshop on Critical Information Infrastructures Security, (CRITIS'06), LNCS 4347, Springer, pp. 262 – 272, (2006).

27. Home of GSM Association,"Market Data Summary (Q2 2009) – Connections by bearer technology",http://www.gsmworld.com/newsroom/market-data/market_data_summary.htm

28. 3GPP TS 09.60 (V7.10.0): GPRS Tunneling Protocol (GTP) across the Gn and Gp Interface, 2002

29. 3GPP TS 02.17 (v8.0.0): Subscriber Identity Modules (SIM) Functional Characteristics, 1999

30. Shneier, B.: Attack Trees: Modeling security threats. Dr. Dobb's Journal, http://www.schneier.com/paper-attacktrees-ddj-ft.html, (1999)

31. Shneier, B.: Secrets & Lies: Digital Security in a Networked World. John Wiley & Sons, (2000)

32. International Organization for Standarization, ISO/IEC 27001, Information Technology – Security Techniques – Information Security Management systems – Requirements, (2005)