

Security for Mobile Operators in Practice

Christos K. Dimitriadis

Department of Informatics, University of Piraeus, 18 Eptanisou, 15236, Nea Penteli, Athens, Greece

(Email: cricodc@unipi.gr)

(Received May 31, 2007; revised and accepted Oct. 23, 2007)

Abstract

Mobile operators are organizations that have to manage a great amount of critical information, including legal sensitive subscriber data. A number of assessment projects on the infrastructures of major mobile operators, revealed a number of vulnerabilities that if exploited may lead to important business impact. The scope of this paper is to publish these vulnerabilities towards the enhancement of security and privacy of mobile operators. A threat model was created, according to which countermeasures tailored to the specific environment of a mobile operator are proposed.

Keywords: Mobile operators, security, vulnerabilities

1 Introduction

During the recent years, mobile telecommunication networks were transformed from infrastructures that provided voice and very limited data services to infrastructures that provide a wide range of multimedia data services [16]. These new demands in service provision, required enhanced versions of supporting services such as charging and billing, roaming, interworking and addressing, along with the necessary security services for protecting confidentiality, integrity and availability of all types of information, including user traffic, signalling and control data.

The outcome of this transformation was an upgrade of the existing closed Signalling System 7 (SS7) based networks to Internet Protocol (IP) based systems that combined a number of old and new technologies and applications under the pressure of timely service delivery start-up. This situation created a number of security vulnerabilities reported by several research studies in the field [6, 7, 9, 15]. This paper describes the consolidated results of security assessments in a number of major mobile operators, sharing experiences and highlighting existing security issues in the specific environment of mobile operators. The scope of this paper is to publish vulnerabilities and countermeasures towards the enhancement of security and privacy of mobile operators, according to the fundamental principle that security in practice is based on the knowledge we have regarding the vulnerabilities applied

in a specific environment [12, 13]. The problem is defined by presenting vulnerabilities, threats and the corresponding business impact, identified through the security assessment projects. The paper is organized in five main sections, not including the introduction and conclusions. Section 2, presents the importance of security and privacy for mobile operators. Section 3, describes the areas of the mobile operators' infrastructure, upon which the security assessments were applied and provides basic knowledge about the systems involved. Section 4 presents a threat model, as an outcome of processing the results from the assessments. Section 5 proposes countermeasures tailored to the specific environment of mobile operators.

2 Security and Privacy for Mobile Operators

Mobile Operators are organizations with very strict and nonnegotiable security and privacy requirements [10]. These requirements derive from the need for protecting the company's shareholders' interests and are very essential in order to assure and maintain business continuity, profitability and growth, respected organization image and competitive edge and legal compliance. Unfortunately, the vast amount of critical information handled by mobile operators, in combination with the complexity of their infrastructures is making the addressing of these requirements more and more difficult to achieve. At the same time the European Union Directive on privacy and electronic communications [5] considers the privacy of communication and of subscriber personal data, as a fundamental element of the society, while the continuously increasing demand for the world to be able to effectively face terrorism and organized crime forced governments to permit the lawful interception of communications [5, 14].

Security is defined by three fundamental elements, as far as information is concerned: confidentiality, integrity and availability [8]. Information types, include user related technical data and corporate data. User-related technical data include user traffic, charging data, billing data, location data, addressing data, identity data, security management data, access control management data and service profile data [1]. Corporate data include cus-

tomers' personal and financial data, contract data, employee information, organizational and corporate financial, sales and marketing data and roaming data. Except from the need to balance privacy and lawful interception in practice and from the complexity of the operator's business functions and services, including a vast amount of information types, the path towards security and privacy is further complicated by the nature of the environment of the mobile operator, which includes the following parameters:

- The large number of employees that have access to sensitive data and most of them belong to the lower payroll streams.
- The large number of partners' and suppliers' staff that gain at least temporary access to sensitive data.
- That electronic monitoring equipment and tools (such as cameras, tape recorders etc) are neither legally acceptable nor welcomed in any working environment.
- Some of the information accessed by employees may have a very high value (under various circumstances).
- Various outsiders are trying to gain access to sensitive information, either by technical means or by offering very significant amount of money to employees to get it for them.
- Official public authorities may often put a lot of pressure to employees to gain easier and faster access to data (than following the legally accepted and established procedures)
- No detailed employee screening policy is usually followed during the hiring process

A security incident may cause important business impact to the mobile operator. Breach of confidentiality can result in severe embarrassment, financial loss, and even litigation for a mobile operator. Further types of serious disclosure involve secret commercial information, plans and strategic directions, customer personal data and subscribers call data or information disclosed to legal representatives. A breach of integrity can also be particularly devastating. Especially in the case of billing data the impact may be direct monetary loss. Legal problems may also be a result of the realization of this type of threat, since for example subscriber call data may be related to the investigation of a legal case, or subject to lawful interception. As far as availability is concerned, the loss of records or data can be particularly disruptive. Especially for mobile operators, where most applications and services are time critical, a breach of availability may cause several impacts, including direct monetary loss, increase in costs through additional working hours and a great deal of embarrassment where information is unexpectedly not available. Legal problems may also be a result when not being able to keep up with service level agreements.

3 Domains of Assessment

The rationale behind the security assessments, is based on the fact, that the improvement of security in Third Generation (3G) mobile systems mainly focuses on the air-interface, by implementing mobile terminal to Universal Mobile Telecommunications System (UMTS) Terrestrial Radio Access Network (UTRAN) mutual authentication, as well as solving a number of existing problems caused by vulnerabilities of the underlying cryptographic technology of Second Generation (2G) networks. Special attention is also given to user privacy in UTRAN, by the deployment of an identity management scheme, which protects the confidentiality of the user identity and position, as well as of service delivery to the user [2].

The target of assessment is to build a threat model, regarding elements of the service provision path, other than the UTRAN, including the core mobile network, as well as the corporate network of the mobile operator as depicted in Figure 1.

The 3G core network consists of the Circuit Switched (CS) domain, the Packet Switched (PS) domain and the IP Multimedia Subsystem (IMS) [3]. The CS domain serves traffic switching and signaling for voice mobile connections, linking the UTRAN with other voice networks such as the Public Switched Telephone Network (PSTN). The PS domain, serves traffic switching and signaling for data connections, linking the UTRAN with other Packet Domain Networks (PDNs), including the Internet. For that purpose, the PS domain consists of the Serving General Packet Radio Service (GPRS) Support Node (SGSN) and Gateway GPRS Support Node (GGSN). The IMS is a complementary subsystem, providing multimedia services over the PS domain. SGSN and GGSN connect to roaming partners - through the Gp interface - as well as with critical systems's including the Home Location Register (HLR), the Mobile Switching Center (MSC)/Visitor Location Register (VLR), the Charging Gateway (CG) and the Radio Network Controllers (RNC). The Gi interface is the medium for communication of GGSNs with external PDNs. Communication between SGSNs and GGSNs is realized through the GPRS Tunnelling Protocol (GTP), over the Gn interface [4].

The corporate network indicatively hosts the billing system, the Enterprise Resource Planning system, the fraud detection system, the Data warehouse, systems for subscriber management and customer relations, as well as other systems supporting the business functions of the mobile operator. Most of these systems handle critical data, including call details, personal subscriber data and sensitive corporate data. The architecture of the corporate network depends on the specific implementation of each mobile operator. This also applies for the connection of these systems with the core mobile network, which is necessary in order to implement business processes. For example, the fraud detection system may receive information from many systems of the core network, as well as systems from the corporate network and especially the

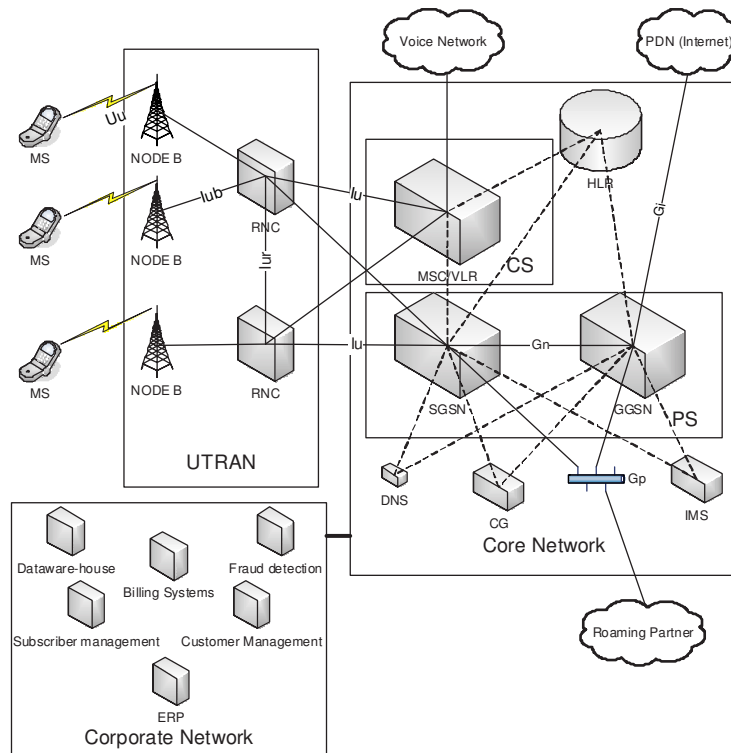


Figure 1: Basic elements of the mobile infrastructure

billing system, in order to combine and compare information depending on its specific functions.

4 Threat Model

The security assessments of the corporate and core networks of a number of major mobile operators, lead to the creation of a threat model. The target of the threat model is not only to identify new vulnerabilities regarding mobile operators, but also describe some known vulnerabilities of information systems that become very important and difficult to address due to the specific nature of the environment of a mobile operator.

The threat model describes threats than can be realized by the exploitation of vulnerabilities, through attacks. The graphical description of the threat model is based on a combination of attack trees [12]. An attack tree has a root node, leaf nodes and child nodes. The root node represents the target of the attack, the leaf nodes, the means for reaching the target and the child nodes, represents the events that consist the attack. All leaf and child nodes have an OR logic. Figure 2 presents a combination of attack trees.

The combined attack trees have three root nodes, one for each final targets that correspond to the breach of the system's security cornerstones: confidentiality, integrity and availability. An intruder may use as a mean

for reaching the target, one of the leaf nodes, which are the compromise or disruption of a critical system or information. Vulnerability Groups (VG), reflect the main mobile-operator-specific findings of the security assessments, while 19 possible attacks (A) can be deployed for exploiting the vulnerabilities.

- **VG1. Inadequate Asset Registry:** The infrastructure of mobile operators, involves a great number of assets, while the technical personnel usually has a long catalogue of updates to implement in a daily basis. Asset registries are kept by the individual departments and are not updated on time, due to the increased workload. The update of the registries does not follow a globally specified procedure and thus there is no regular way of auditing the accuracy of their contents. This vulnerability may permit the installation of unauthorized software or non-permitted configuration of a device altering its behavior (A1). This attack may lead to unauthorized access or disruption of a critical system, including call or data interception, for example by attacking the MSC or a data gateway, realizing all possible threats.
- **VG2. Flaws in Identity Management:** The multitude of systems that process or store critical and legal sensitive data in combination with the fast infrastructure upgrades is the main reason for the existence of this group of vulnerabilities. Exploiting this vul-

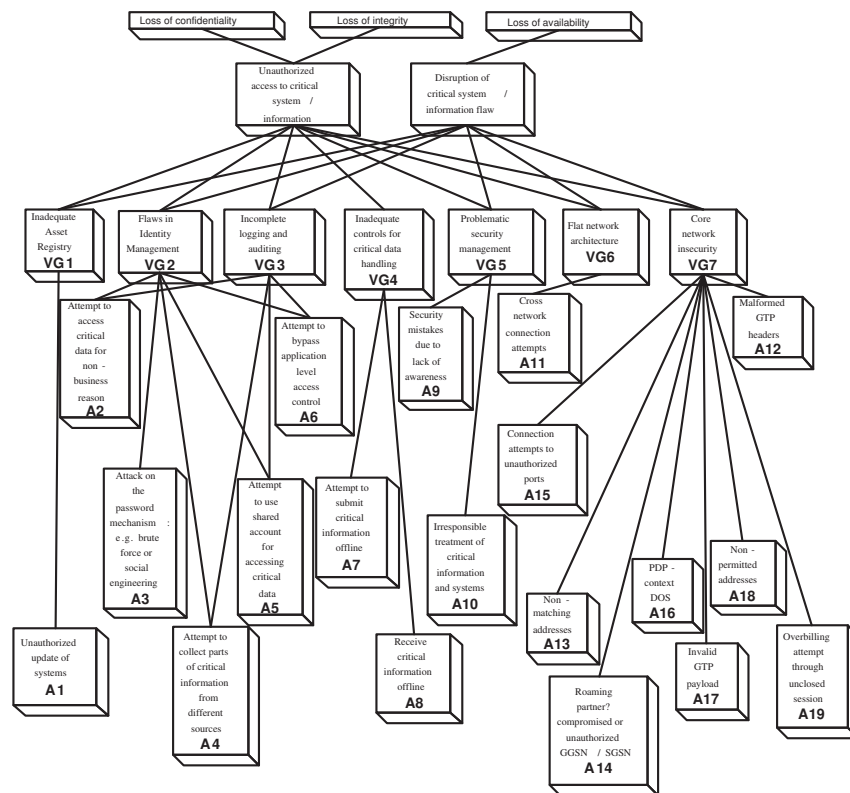


Figure 2: Combined attack trees

nerability group, may lead to unauthorized access or disruption of a critical system, realizing all possible threats. The following have been identified:

- There is no detailed description of the roles of all of the users so as to be able to define the exact access rights in specific information types that each user needs in order to complete a work-task. The absence of the detailed description does not permit any attempt in refining and re-defining the access rights of each user and thus there is no way of optimising the access process by eliminating surplus access rights. There is no central management of the users and no central mechanism for auditing the access rights of the users and imposing central access policies. User management is implemented for each specific system and/or application. Access rights cancellation for users that either transfer from one department to the other or do not any more belong to the workforce of the mobile operator is accomplished with a time lag that poses serious security concerns. For some systems the business need of some departments does not justify access to critical data. This fact may permit access to critical data, such as call details or personal data, for personal reasons, without a specific business need (A2).
- The access control mechanism used in all systems is password based. This is a weak access mechanism especially if the criticality of the data of the mobile operator is taken into consideration. Attacking the password-based mechanism (A3), most likely by deploying social engineering, combining this vulnerability with the lack of security awareness of the personnel, may permit access to critical data and systems, causing loss of confidentiality and integrity, as well as loss of availability, in the case of attacking the configuration of a critical system towards denial of service.
- The same types of critical information are being accessed in different systems, by the same users. Although access to these systems may be necessary for gaining complementary parts of information, this operation makes information access very complex and might influence an investigation procedure in the case of an incident. A possible attack (A4) regards the collection of parts of critical data from different sources (e.g. personal data linked to call details), towards a possibly undetected security breach.
- Shared accounts are activated, with the excuse that a large number of technical personnel has to commonly and quickly access raw data records in systems. Using a shared account for accessing critical data (A5), not only breaches

system security, but is also an extra motive that makes forensic investigations very difficult to implement and define responsibility.

- Direct queries to databases containing critical data can be executed by business users. This fact permits bypassing the application layer authentication and authorization mechanisms (A6). This facility has been added for specific business users, who had to quickly verify information for customers, when the corresponding facility was not provided by the application. Although the application has been upgraded, the capability for direct queries has not been disabled.
- VG3. Incomplete logging and auditing: This group of vulnerabilities, provides an extra motive to the attacker, who can deploy attacks A2, A4, A5 and A6, without being detected.
 - Lack of adequate logging capabilities of legacy systems. Logging capabilities vary from system to system even when these systems contain critical information of the same level. There exist systems for which the activated logging capabilities are inadequate for the criticality and impact of the data they contain. For example, view logging is of primary importance for call details and this facility is not provided by a number of legacy systems.
 - Huge number of system logs, not correlated and not managed. Logs of systems with critical information are not audited in a scheduled and comprehensive way. There are no means to correlate logs from various sources (applications, operating system, Database, network elements) in order to be able to create an audit trail for the user actions.
- VG4. Inadequate controls for critical data handling. When accessing critical information, some facilities like copy/paste and printing are available, permitting the transfer of critical data off-line. Some business processes, permit the submission of unencrypted e-mails with sensitive marketing and sales data as attachments.
 - Critical data are being distributed by email without the use of mechanisms for ensuring their confidentiality and integrity. This may permit the deployment of simple attacks that breach critical data confidentiality (A7).
 - Mailing lists are updated with a time delay and thus sometimes users that should have been removed from them still receive critical data through them (A8).
- VG5. Problematic security management.
 - The security awareness of the personnel is low and thus the voluntary participation of the people in the process of following the specified security guidelines and procedures is not achieved. Hardcopies of critical data do not always bear appropriate labeling. Hardcopies of critical data are not always handled according to their criticality. This situation may lead to mistakes that cause security breaches (A9).
 - Information ownership is not clear due to the complexity of the mobile operators' environment. Call Detail Records for example, exist in many forms, from the time of their generation by the mobile network elements (for example SGSNs) to their storage and processing by the corporate billing system, making the segregation of responsibilities on information types and as a result, information ownership difficult to define. Since responsibility is vague, this situation may become a motive for causing security breaches from the inside (A10).
- VG6. Flat network architecture: These vulnerabilities include flat mobile core networks and management networks, shared operations and management networks also connecting to corporate networks, insecure billing and lawful interception connectivity. Uncontrollable communication may permit cross network attacks (A11) causing access to critical systems of belonging to other domains than those that the user should have access to.
- VG7. Core network insecurity: Core network security is usually not a priority, causing inadequate firewall and intrusion detection architectures. This fact may lead to critical system compromise or disruption, realizing all possible threats. Attacks may be deployed by the submission of GTP packets with malformed GTP headers (A12) - for example incorrect value of GTP or element length or non-permitted message types - may lead to GGSN/SGSN compromise or disruption. This attack exploits badly configured on non-existent GTP firewalls. When the GGSN/SGSN is compromised, it may become a vaulting house for compromising or disrupting other critical systems, such as the HLR, the MSC, the CG or the billing gateways, by exploiting the flat core network architecture of the mobile operator and the lack of security zones. Moreover, the core network of the mobile operator becomes a logical extension of the core network of the roaming partner with limited control, exposing both to serious threats, impacting security from and to external (roaming partners) nodes. In any case, the compromise of a critical system, including the GGSN/SGSN may lead to the loss of confidentiality or integrity of critical and legal sensitive information, while disruption leads to loss of availability (Denial of Service). Other attacks

that exploit badly configured on non-existent GTP firewalls, with the same results are based on the following techniques:

- The submission of non-matching addresses between the GTP payload (encapsulated packet) and the assigned address as defined in the Packet Data Protocol (PDP) context handshake (A13).
- Connection attempts from unauthorized or compromised SGSNs or GGSNs of a roaming partners, exploiting the poorly controlled communication path that connects the two networks (A14).
- Attempts to connect to the target systems in administration ports and generally any attempt to connect to a port other than the standard GTP ports (A15).
- Attempts to initiate a great number of PDP Contexts that may result the disruption of a GGSN/SGSN, leading to loss of availability of these systems (A16).
- GTP packets with invalid GTP payload (including GTP in GTP encapsulation and encapsulation of non-IP based protocols in GTP) may result packet spoofing and not permitted sessions (A17).
- GTP packets with payload that contains non-permitted source and/or destination addresses may result packet spoofing and not permitted sessions (A18).
- Over-billing attacks which are based on the existence of unclosed sessions (A19). A specific IP address requests data from a server and releases the IP address, which is re-assigned to a new user. The new user keeps receiving the data, causing the over-billing of the new user. This attack exploits badly configured on non-existent GTP firewalls and leads to loss of integrity of billing information, through the creation of false billing data.

5 Enhancing Security and Privacy

While some of the identified vulnerabilities seem that can be addressed by the implementation of well-known security controls, it is important to explain specific solutions that address the particularity of each vulnerability in the environment of the mobile operator. Regarding VG1 and VG5, the first step is to combine the implementation of an automated asset registry mechanism, with strict procedures defined by the corporate security policy, as well as security awareness. The personnel of the mobile operator should be informed regarding the importance of keeping an updated registry and to follow a well-defined change management procedure, since the opposite could

permit attacks (for example eavesdropping when altering the MSC) that have major business impact, including legal implications. The assignment of information ownership is also a very important step that troubles mobile operators due to the reasons explained in the previous section. The most appropriate information owners are those individuals who best understand the information's value and threats, and have the authority to balance risk against cost [13]. Based on that principle, we propose the following dual-level information ownership schema for mobile operators, consisting of two information owners who also interact with system owners that host the information. This schema requires interdepartmental cooperation:

A first level information owner should be a business user, holding a position in the upper levels of the hierarchy, whose responsibilities are most relevant and most dependent to this specific information. This information owner should have the supervision of the specific information handling and should provide guidance to the second level information owner. The first level information owner should be responsible for any decision taken regarding the information. This information owner should also interact with system owners that host the information in the various types that it is transformed to, after being processed by the various corporate systems.

The second level information owner, should be a business user at lower levels of the hierarchy, whose business duties are most relevant with the information. This information owner should be responsible for the every-day handling of information and should notify the first level information owner in case a decision has to be made, as far as the information is concerned.

Regarding VG2, a unified user management system should be employed in order to perform user-provisioning related tasks from a central point in the network. This user provisioning and management system should enable easy generation of a user's profile according to the department they belong. Based on a predetermined set of business rules, the system should allow users' profiles to be set up automatically and the users to be granted the minimum access privileges they need to perform the respective business functions. The users will thus be registered to every application, database, middleware and operating system instance to which they need to have access. Furthermore, the system should allow for the self-organization of each user account (reset passwords, change personal information). The system should be extensible in order to support different types of access control on different systems according to the sensitivity of the data or a documented policy. When the user's status in the organization changes the system should allow for full revocation of the granted access rights or the assignment of a new profile. Additionally, the system should be able to provide tailored reporting facilities (e.g. per department, per system etc) depending on management needs. In order to implement strong access control characteristics on the new system, a clean separation of user duties

and roles should be implemented. It is desirable that wherever possible, a mechanism for providing on-demand authorization to view selected sensitive data should be developed. In this mechanism a separate user should be designated to dynamically assign privileges to users that need to perform a business function on the selected sensitive data following the request-approve-resolve principle.

Regarding VG3, the ability to monitor and audit user activity is fundamental for security. Specific provisions need to be applied in order for the security officers to be able to identify abnormal user behavior and to track a security incident regarding leak of sensitive information to specific user accounts in an undeniable manner. Therefore, applications should be extended or updated to provide the generation of detailed audit logs that record user activity in both system and application level, while logs should be collected by a centralized mechanism for further analysis and correlation.

Regarding VG4, the technical measures for protecting access and viewing of critical and sensitive data, provisions must be made for the security of the data during transfer between the departments of the organization. Procedures that identify explicitly which user within the organization is authorized to transmit and receive sensitive data from other departments and under which circumstances or business functions this is necessary, should be drafted. Furthermore, the procedures should foresee the manipulation of the sensitive data after their transmission with respect to storage and validity period in order to ensure that only authorized personnel will be able to view the data and that data can be securely discarded if their business purpose or validity is terminated. In addition, complementary technical mechanisms must be applied in order to ensure that the security attributes of authentication, confidentiality, integrity and non-repudiation are maintained during transmission and receipt of sensitive electronic messages. Copy/paste should be disabled were not necessary and printing must be implemented in a controlled environment. All e-mail communications that contain sensitive data should be digitally signed by the sender, in order to ensure the authenticity of the source and the integrity of the data in transit.

VG6 and VG7, concern network security controls. Filtering and detecting gateways should separate the various networks into zones, as well as the various critical systems within its network. Figure 3, presents an indicative architecture for separating the mobile network into zones.

There are 7 indicative zones, separating different types of critical systems into zones. Control Points (CP), act as filtering and detecting gateways, which separate the zones and implement security policies depending on the zones they are related with. Control points related to the core network, should be configured in order to detect and block all attacks described in the previous section, including attacks against the GTP protocol, denial of service attacks and unauthorized connection attempts. Over billing attacks are addressed by configuring the CP, in order to detect session termination and update the IP

address pool. The corporate and management network, may access each of the critical systems by connecting to an interface of the corresponding control point. For example, a department that needs to automatically modify a subscriber's profile in the HLR, should use an application which sends commands to the HLR through interface A of CP4.

6 Conclusions

Mobile operator's security should be addressed as a whole, involving all elements in the service provision path. The multitude of systems that handle critical data in combination with the time-sensitive requirements for service provision upgrade, lead to the creation of immature security architectures that hinder vulnerabilities. The exploitation of these vulnerabilities may lead to important business impact, including legal implications. Addressing these vulnerabilities requires a deep understanding of the problem, taking into account the particular nature of the mobile operator's environment. This understanding is a prerequisite for the interpretation and customization of security controls defined by international security standards and best practices. The threat model and countermeasures proposed in this paper were based on security assessments on major mobile operators and aimed to contribute towards that direction, increasing security and privacy.

References

- [1] 3rd Generation Partnership Project, TS 21.133, *3G Security; Security Threats and Requirements*, 2004.
- [2] 3rd Generation Partnership Project, TS 33.102, *3G Security; Security Architecture*, 2004.
- [3] 3rd Generation Partnership Project, TS 23.002, *Network Architecture*, 2004.
- [4] 3rd Generation Partnership Project, TS 29.060, *GPRS Tunnelling Protocol (GTP) Across the Gn and Gp Interface*, 2005.
- [5] DIRECTIVE 2002/58/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications).
- [6] W. Donald, and L. Scott, "Wireless security threat taxonomy," *IEEE Workshop on Information Assurance*, United States Military Academy, West Point, NY June 2003.
- [7] N. E. Fishway, M. Nofal, and A .Tadros, "An improvement on secure communication in PCS," *Performance, Computing, and Communications Conference*, pp. 175-182, 2003.
- [8] ISO/IEC, 17799, *IT - Code of Practice for Information Security Management*, 2005.

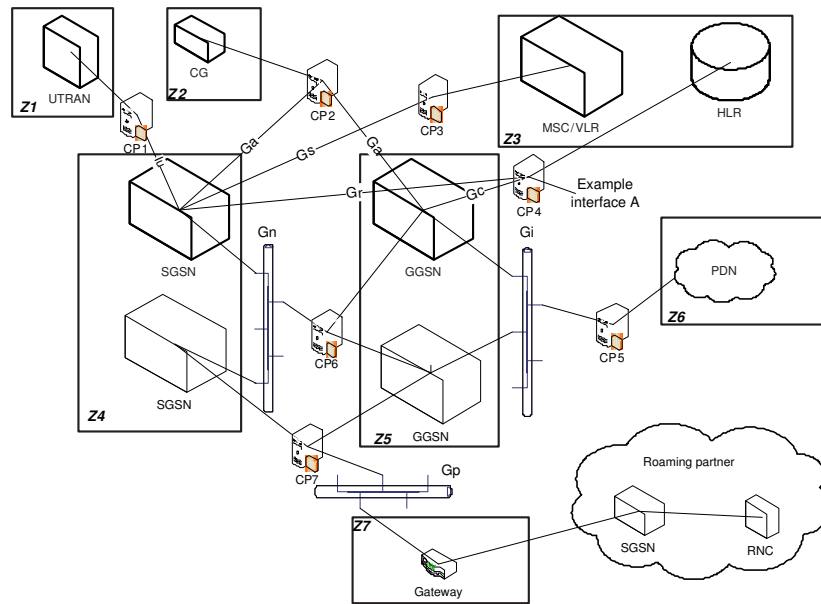


Figure 3: Indicative segregation of mobile network into zones

- [9] K. Kameswari, L. Peng, S. Yan, and F. L. Thomas, "A taxonomy of cyber attacks on 3G networks," *ISI 2005*, LNCS 3495, pp. 631-633, 2005.
- [10] V. Neimi, and K. Nyberg, *UMTS Security*, John Wiley and Sons, 2003.
- [11] J. Ramachandran, *Designing Security Architecture Solutions*, John Wiley and Sons, 2002.
- [12] B. Schneier, "Attack trees," *Dr. Dobbs' Journal*, vol. 24, no. 12, pp. 21-29, 1999.
- [13] *The Honeynet Project Know Your Enemy. 2nd ed Learning About Security Threats*, Addison-Wesley, 2004.
- [14] Uniting and Strengthening America by Providing Appropriate Tools Required to Intercept and Obstruct Terrorism (USA PATRIOT ACT) Act of 2001.
- [15] O. Whitehouse, and G. Murphy, *Attacks and Counter Measures in 2.5G and 3G Cellular IP Networks*, @stake press, 2004.
- [16] D. Wisely, P. Eardley, and L. Burness, *IP for 3G-Networking Technologies for Mobile Communications*, John Wiley and Sons, 2002.
- Christos Dimitriadis** is specialized in prevention, detection and response Information Technology (IT) security mechanisms for 2G and 3G mobile systems. Dimitriadis has 34 publications in the field of IT security and privacy and he is a founding member of the international Mobile-Government Study Group - MGSG. He has been invited by several organizations to provide lectures and security training, including the ITU, NIST and the TAIEX office of the EC. Dimitriadis received a PhD from the University of Piraeus, a diploma of Electrical and Computer Engineering from the University of Patras and holds the CISM and CISA certificates from ISACA.