Wireless World Research Forum (WWRF)

Title: Enhancing end-users privacy in 3G networks

Christos Xenakis and Lazaros Merakos

Communication Networks Laboratory
Department of Informatics & Telecommunications
University of Athens, 15784 Athens, Greece.
Tel. +30 210 7275418, +30 210 7275323, Fax. +30 210 7275601
{xenakis,merakos}@di.uoa.gr

Subject Area: WG3 – Cooperative & AdHoc Networks

1. Objectives of the required research

Mobile/wireless Internet is becoming available with the advent of third generation (3G) mobile communication systems. Along with the variety of new perspectives, mobile Internet also raises new concerns on security issues. The radio transmission is by nature more susceptible to eavesdropping and fraud in use than wireline transmission. The user mobility and the universal network access certainly provoke security treats. The introduction of IP-based transport technology to the core of 3G networks brings along new vulnerabilities and potential threats. Mobile network operators do not deploy their own private networks, but they rather rely on the existing Internet infrastructure for the establishment of intra-network, and inter-network communications. Furthermore, the complex network topologies and the heterogeneity of the involved technologies increase the dependability challenge.

Security is a critical factor to realize the opportunities presented by the ubiquity of mobile devices and networks. It is therefore important to address information security concerns for wireless applications, especially for those applications that have anything to do with mobile e-business, e-government, e-finance and e-health. Moreover, seamless access to private networks by mobile workforce is expected to be a fashionable and valuable service. Telemetry and other un-tethered equipment, traveling sale forces, field maintenance crews, telecommuters, and other mobile professionals are driving demands for anywhere - anytime secure access to corporate intranets, databases, and e-mail servers.

Without privacy and security, personal users would perceive mobile communications as harmful and would not use the respective systems. However, user data in the UMTS backbone network are conveyed in clear-text exposing them to various threats. Moreover, inter-network communication is based on the public Internet, which enables IP spoofing to any malicious third party who gets access to the network. Thus, next generation mobile subscribers require generic, flexible, client initiated security mechanisms, which can provide advanced security services to user data traffic according to the particular end-users needs, and will be available anywhere – anytime.

Security principles such as confidentiality, integrity and authentication can be guaranteed by the deployment of Virtual Private Network (VPN) technology [2]. VPN authenticates and authorizes user access to corporate resources, establishes a secure tunnel, and encapsulates and protects data conveyance over a network. It extends dedicated connections between remote branches, or remote access to mobile users over a shared infrastructure. The advantages of using the transport facilities of a public network, combined with advances

in the field of network security, make VPN services extremely attractive compared to traditional private line services. However, mechanisms such as VPN and IPsec were originally conceived to address network security issues for fixed-point networks. Wired environment solutions can often be extended for applications to wireless environments, but they might need some changes or a complete rebuild. It is critical to ensure that security services provided in wireline network should be available in wireless environment too.

In this article, the incorporation of dynamic, client-initiated, IPsec-based VPN solutions over the UMTS network infrastructure is proposed. Specifically, three alternative schemes for secure VPN deployment over the UMTS network are outlined. The UMTS infrastructure provides the mobile users access to the public Internet, and allows them to employ IPsec tunnel technique to traverse firewalls, access private networks, and convey sensitive data securely. This type of access is referred to as voluntary tunneling, since it enables a mobile node to establish a secure communication channel back to a private network. The proposed schemes differ on the position where the IPsec functionality is placed within the UMTS network architecture (mobile node, access network, and UMTS network border), and whether data in transit are ever in clear-text or available to be tapped by outsiders. The different security models are: a) the end-to-end, b) the network-wide, and c) the border-based.

2. State of the art in the area

2.1. 3G-security architecture

To satisfy the security requirements and counteract against security threats, 3G-systems have incorporated a specific security architecture. This architecture is built on the security principles of 2G-systems with improvements and enhancements in certain points in order to provide advanced security services. Its main objective is to ensure that all information generated by or relating to a user, as well as the resources and services provided by the serving network (SN) and the home environment (HE), are adequately protected against misuse or misappropriation. Fig. 1 gives an overview of the complete 3G-security architecture, illustrating five major security classes: (I) network access security, (II) network domain security, (III) user domain security, (IV) application domain security and (V) visibility and configurability of security [6, 7]:

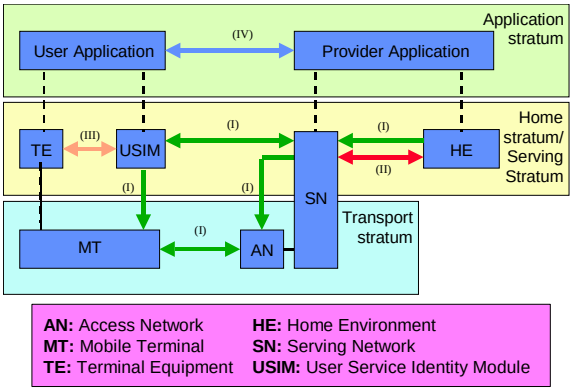

Fig. 1 Security architecture

Network access security is a key component in the 3G-security architecture. This class deals with the set of security mechanisms that provide users with secure access to 3G services, as well as protect against attacks on the radio interface. Such mechanisms include: i) user identity confidentiality, ii) authentication and key agreement iii) data confidentiality and iv) integrity protection of signaling messages. Network access security takes place independently in each service domain.

Network domain security features ensure that signaling exchanges within the UMTS core, as well as in the whole wireline network are protected. Various signaling protocols and interfaces are protected by standard procedures based on existing cryptographic techniques. Specifically, the IP–based protocols are protected at network level by means of IPsec [1], while the realization of protection for the SS7-based protocols and the access network interfaces are accomplished at the application layer.

User domain security ensures secure access to the mobile station (MS). It is based on a physical device called UMTS Integrated Circuit Card (UICC), which can be easily inserted and removed from terminal equipment, containing security applications such as the User Service Identity Module (USIM). The USIM access is restricted to an authorized user, or to a number of authorized users. To accomplish this feature, the user and the USIM must share a secret (e.g., a PIN). The user gets access to the USIM only if he proves knowledge of the secret.

Application domain security deals with secure messaging between the MS and the SN or the SP over the network with the level of security chosen by the network operator or the application provider. A remote application should authenticate a user before allowing him to utilize the application services, and it could also provide for application-level data confidentiality. Application-level security mechanisms are needed because the lower layers' functionality may not guarantee end-to-end security provision. Lack of end-to-end security could be envisioned when, for instance, the remote party is accessible through the Internet.

Although the security measures provided by the SN should be transparent to the end user, for certain events and according to the user's concern visibility of the security operation as well as the supported security features should be provided. This may include: a) indication of access network encryption; b) indication of network wide encryption; and c) indication of the level of security (e.g., when a user moves from 3G to 2G).

Finally, configurability enables the mobile user and the HE to configure whether a service provision should depend on the activation of certain security features. A service can only be used when all the relevant to it security features are in operation. The configurability features that are suggested include: a) enabling/disabling user-USIM authentication for certain services; b) accepting/rejecting incoming non-ciphered calls; c) setting up or not setting-up non-ciphered calls; and d) accepting/rejecting the use of certain ciphering algorithms.

2.2 Complementary network security features

Besides the security features that are included in the 3G-security architecture, the mobile network operators can apply traditional security technologies used in terrestrial networking to safeguard the UMTS core network, as well as the inter-network communications, such firewalls, and pre-established VPNs [2].

Firewalls can be characterized as a technology providing a set of mechanisms to enforce a security policy on data from and to a corporate network. They are established at the borders of core network allowing traffic originated from specific foreign IP addresses. Thus, firewalls protect the UMTS backbone from unauthorized penetration. Furthermore, application firewalls prevent direct access through the use of proxies for services, which analyze application commands, perform authentication, and keeps logs.

However, firewalls were originally conceived to address security issues for fixed networks, and thus, are not seamlessly applicable in mobile scenarios. They attempt to protect the clear-text transmitted data in the UMTS backbone from external attacks, but they are inadequate against attacks that originate from other mobile network malicious subscribers, as well as from network operator personnel or any other third party who gets access to the UMTS core network. Mobility may imply roaming between networks and operators, possibly changing the source address, which because of the static configuration of firewalls, may potentially

lead to discontinuity of service connectivity for the mobile user. Moreover, firewalls security value is limited because they allow direct connection to ports and cannot distinguish services.

Since firewalls do not provide privacy and confidentiality, pre-established VPNs have to complement them to protect data in transit. The current type of VPN fails to provide the necessary flexibility, to establish reliable secure connections for typical mobile users. VPN services for UMTS subscribers can be established in a static manner between the border gateway of the UMTS core network and a remote corporate security gateway. This makes the realization of VPN services feasible only between the security gateway of a large organization and a mobile operator, when a considerable amount of traffic requires protection. Thus, if the static VPN parameters or the VPN topology has to be changed, then the network administrators in both ends must reconfigure it. Furthermore, the aforementioned security scheme can provide VPN service neither to individual mobile users, who may require on demand VPN establishment, nor to enterprise users that may roam internationally.

2.3 Application-layer security

Application-layer security builds security features into individual applications, and they operate independently of any network security measures. Many applications have special security requirements that simply cannot be met by network security services. Security at this level is by far the easiest to deploy, as long as all users are running a homogeneous application on a standard platform. While these methods are effective for solving specific security problems, such solutions are by their nature limited to their specific niches. For instance, Transport Layer Security (TLS) [2] works fine for simple client-service cases, but, in case that the service contains a considerable number of cross-references to other servers, for each one a separate key-exchange operations is needed overloading unnecessarily the client.

Wireless Application Protocol (WAP) is a suite of standards for delivery and presentation of Internet services on mobile phones and wireless terminals. A lightweight layered protocol architecture is introduced in the wireless network segment, as a wireless equivalent to the Internet protocol stack, taking into account the limited bandwidth of mobile networks, as well as the restricted processing capabilities of mobile phones. To connect the wireless domain to the Internet, a WAP gateway is needed to translate the protocols used in WAP to the protocols used on the public Internet. The WAP architecture has been standardized in two releases (ver. 1.2.1 and ver. 2.0) [3].

To secure data transmission in WAP architecture (ver. 1.2.1) the Wireless Transport Layer Security (WTLS) protocol [3], which is based upon the TLS protocol, is employed. WTLS has been optimized for use over narrow-band communication channels providing also datagram support. It ensures data integrity, privacy, authentication and denial-of-service protection. For Web applications that employ standard Internet security techniques with TLS, the WAP gateway automatically and transparently manages wireless security and conveys protected data between the WTLS and TLS security channels (see Fig. 2).

However, this scheme does not guarante data privacy, since it can not provide end-to-end security. Although encryption is used, the WAP gateway constitutes a security hole since, inside the gateway, the data is transmitted in its original un-encrypted form. WTLS is only used between the mobile device and the gateway, while TLS can be used between the gateway and the web server. From a security point of view, this means that the gateway should be considered as an entity-in-the-middle. This security weakness means that data exchanged may be available to people with privileged access to the WAP gateway, and, thus, the privacy of the data depends the gateway’s internal security policy.

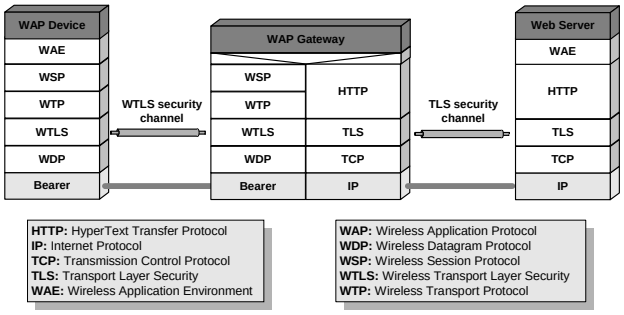

Fig. 2: WAP 1.2.1 architecture

WAP 2.0 proceeds to the re-design of the WAP architecture introducing the existing Internet protocol stack, including the Transmission Control Protocol (TCP), into the WAP environment for the entire wired and wireless part. The new architecture allows a range of different gateways, which enables conversion between the two protocol stacks anywhere from the top to the bottom of the stack. A TCP-level gateway allows for two version of TCP, one for the wired and another for the wireless network, on top of which a secure TLS channel can be established all the way from the mobile device to the server (see Fig. 3). The availability of wireless profile of the TLS protocol, which includes cipher suites, certificate formats, signing algorithms and the use of session resume, enables end-to-end security support at the transport level allowing interoperability for secure transactions.

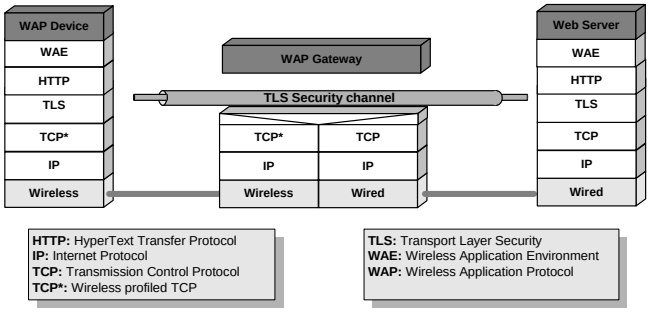

Fig. 3: WAP 2.0 architecture

WAP 2.0 does address the gap in security caused by protocol translation at the WAP gateway. However, the mobile phone would have to use a much higher latency incurring and more bandwidth consuming IP protocol stack. Although TLS can be used to secure the communication of any application, it must be integrated into the to application, and, thus to a large extend it is used for web-based applications. Interaction with the end-user is needed, for example, to check with whom a secure session has been established, or to explicitly request the client to authenticate to the server. TLS is generally a resource consuming protocol for deployment on mobile devices with limited processing capabilities and low bandwidth high-latency wireless network. Moreover, this overhead may be increased by complex key-exchange operations in case that the service contains cross-references to other serves.

3. Possible approach

Summarizing the security analysis, it can be perceived that the existing UMTS security architecture provides

sophisticated security services, and addresses many security concerns that have been listed in the context of next generation mobile networks. However, data communications over the UMTS may experience security threats since, there is a lack of a general-purpose mechanism that can provide advanced security services to user data traffic according to the particular end-users needs, inside and outside the UMTS core network. Despite the ciphering at the air interface, the IP data traffic goes unencrypted all the way from the radio access network to the remote site, and vice-versa. Given that the UMTS core network is based on IP and it is connected to the public Internet, the UMTS backbone may be considered as vulnerable and easily accessible network segment. Firewall technology cannot adequately ensure data transfer within the UMTS core network. Pre-established, static VPN deployment cannot be applied in any mobile scenario protecting all the type of potential services. Moreover, WAP architecture and application layer security are mainly related to web-based applications, which can integrate TLS or WTLS functionality.

The advanced protection of data traffic according to end-users' needs is expecting to empower the security services being provided in 3G networks. Complementary to the 3G-security architecture the incorporation of dynamic VPN technology in the 3G-security framework will further increase the supported level of protection by providing general-purpose security services. A mobile user decides when and where to establish a VPN across a public network managing the initialization process directly. This is well suited to mobile users, particularly those who roam globally and require access to their home networks, as well as off-the-street users needing a VPN for a short time, perhaps for a distributed game. Enterprises find this highly attractive, since their users may roam internationally avoiding the expense of international phone calls back to connect securely to their home networks. Moreover, VPN technology guarantees interworking with existing and forthcoming IP terrestrial network infrastructure.

The most prominent technique for deploying VPN across IP networks guaranteeing interworking with any type of carried services is the network layer security, and the IPsec standard [3]. IPsec facilitates the authentication of the communicating entities, as well as the transparent encryption and integrity protection of the transmitted packets in both IPv4 and IPv6 networks. It is especially useful for implementing VPNs, and remote access to private networks. Concerning VPN deployment there are two general approaches. The first is based on customer premises equipment (CPE) approach, where the VPN capabilities are integrated into CPE devices. The second scheme pertains to network-assisted, where the security functionality and the VPN operation are outsourced to the network operator, or a service provider

3.1 End-end-deployment scheme

Following the CPE approach, the end-to-end [8] security model is realized (see Fig. 4). The communicating endpoints establish by themselves a sesurity association (SA), which extends over the entire multi-nature communication path. Sensitive data are privatized as they leave the originator's site, and remain protected while they are conveyed over the radio interface, the UMTS backbone network, and the public Internet eliminating the possibilities to be intercepted or to be altered by anyone.

For VPN establishment the Internet Key Exchange (IKE) [4] protocol is employed. Both IKE and IPsec have to operate in a mobile UMTS environment, where dynamic IP adressing and Network Address Translation (NAT) [5] may be used. NAT maps an isolated address realm with private unregistered addresses to an external realm with globally unique registered addresses. The conjunction of NAT with IPsec arises many incompatibilities, listed in [10], since the later either hides private addresses through encryption and thus let them escape translation, or it experiences integrity violations as a consequence of the NAT manipulating on protected IP addresses.

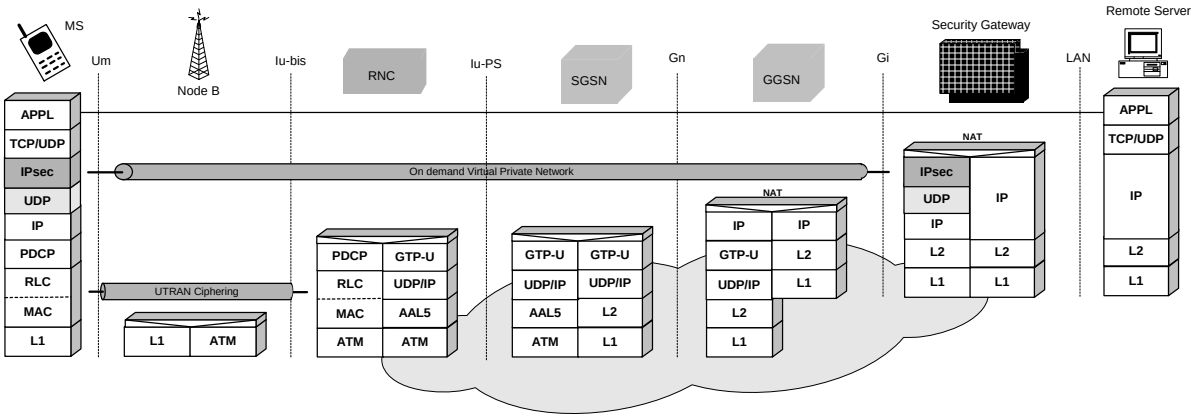

Figure 4: End-to-end VPN deployment scheme over UMTS

This model has minimal impact on the existing network infrastructure, and provides the best security services to the end users. The mobile subscriber may freely move within the UMTS coverage area maintaining network connectivity and VPN service provision. However, it requires from end-users to incorporate the appropriate security software (IPsec) to negotiate and apply security. This imposes computational costs on the lightweight end-user devices, and duplicates encryption (packet encapsulation) over the expensive radio interface. Moreover, the end-user must be aware of when encryption is required.

3.2 Network-assisted deployment scheme

An alternative approach to the end-to-end VPN scheme pertains to a network-assisted scheme. This scheme integrates VPN functionality into the network infrastructure eliminating the need for end-user involvement. The network operator offers responsive, reliable and flexible VPN services, where the administrative and the computational overheads for the end-user are minimized.

For the deployment of a network-assisted security scheme, the MS must be enhanced with a lightweight security client (SecC) module, which is used to request for VPN services and express the end-user preferences. Moreover, a fixed UMTS node should incorporate a security server (SecS) that establishes, controls, and manages VPNs between itself and remote security gateways (SGs) at corporate LANs on behalf of the mobile users (see Fig. 5). SecS comprises an IPsec implementation modified to adapt to the client-initiated VPN scheme, and the security service provision in a mobile UMTS environment. When a mobile user wants to establish a secure remote connection towards a SG, it uses the SecC to request for an IPsec SA from the corporate SecS

By relying on a sequence of concatenated protection mechanisms (Access network ciphering and VPN deployment), it is possible to provide secure data transfer without requiring an extra tunnel overhead on the radio link, or the implementation of computationally intense encryption algorithms in the MS. The VPN functionality is also based on IPsec. For VPN initialization and key agreement procedures an Internet Key Exchange (IKE) protocol proxy [9, 10] scheme has been proposed, which enables the mobile user to initiate a VPN, while shifting complex key negotiation to the network infrastructure.

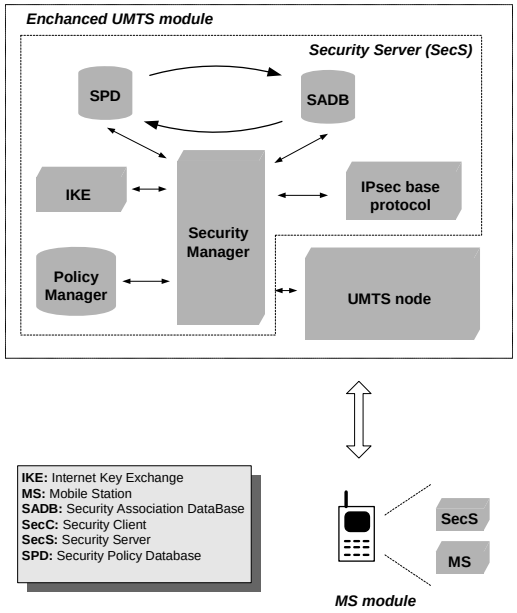

Figure 5: Security Client (SecC) and Security Server (SecS) modules

The deployed VPNs provide maximal security services to end users and is compatible with the legal interception option. The required enhancements for security service provision can be integrated in the existing network infrastructure, and therefore, the security scheme can be used as an add-on feature of the UMTS.

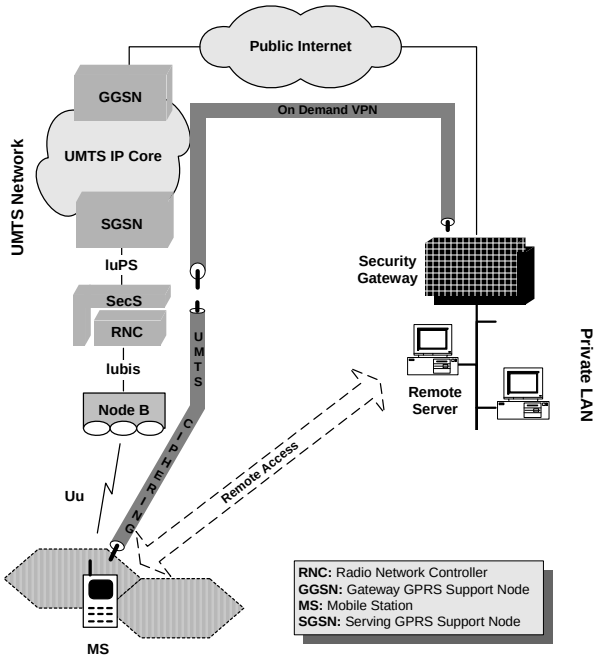

Figure 6.a: Network-wide VPN deployment scheme

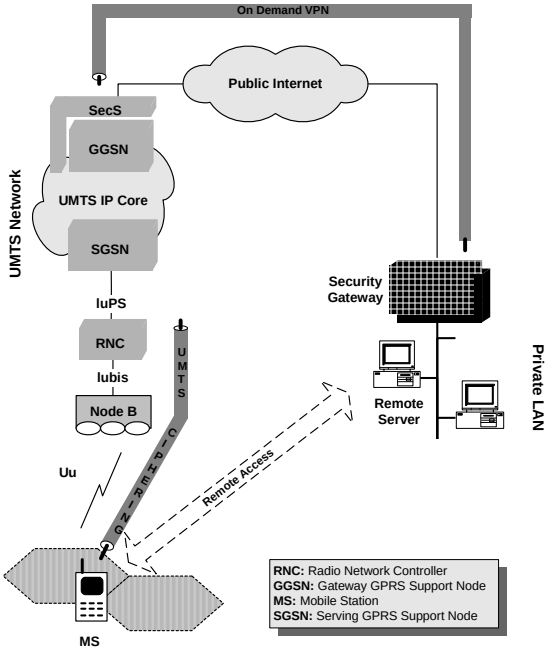

Figure 6.b: Border-based VPN deployment

scheme

There is significant interest in such solutions both by customers, seeking to reduce support costs, and by network operators, seeking new revenue sources. By placing security functionality in the UMTS access network or in the UMTS border, the network-wide or the border-based VPN scheme are deployed respectively [9, 10] (see Fig. 6.a, 6.b). In the network-wide scheme the deployed VPN is extended over the UMTS backbone and the pubic Internet, and thus, the NAT employment in the GGSN node requires special consideration. Moreover, the VPN management infuences the UMTS network management and vice-versa. On the other hand, in the border-based scheme, the VPN expands only on the public Internet segment operating transparently to the NAT presence and the MS movement as long as the MS is under the same GGSN. Otherwise, the current SA is dropped and a new should be established.

3.3 Evaluation and comparison

The proposed security schemes provides dynamic, secure, remote access of mobile users to corporate resources over UMTS network based on IPsec. Each one supports a different level of security, and follows a different model of deployment. Therefore, the particular scheme that will be selected in a potential scenario depends on the risk analysis performed, the required security services, the operator's security policy, as well as the applied network topology. In the following table 1, a comparison of the proposed security models, based on the deployment and operation features, is presented in a tabular form.

Evaluation Parameters	End-to-End scheme	Network-wide scheme	Border-based scheme
Dynamic networking	√	√	√
End to end security	√	√	
Security in one hop	√		
Strong encryption algorithm		√	√
Authentication in end-users' hands	√		
Minimum enhancements in MS		√	√
Less MS cost		√	√
Applicable to any MS type		√	√
Minimum enhancements in the mobile network infrastructure	√		
Become an add-on feature of UMTS		√	√
Less computational overhead in MS		√	√
No encapsulation (IPsec, UDP) over UTRAN.		√	√
No encapsulation (IPsec, UDP) over the UMTS backbone			√
None points of bottleneck in the UMTS network architecture	√		
No interrelation between the VPN management and the network management	√		√
Less control plane overhead	√		√
System reliability corresponds to network reliability	√		
Scalable model		√	√
VPN outsourcing model		√	√
No third party trust	√		
Less skilled end-user required		√	√
SA transparency to end-user		√	√
Solid VPN management		√	√
No SLA required	√		
No enhancements required to support mobility	√		√
MS movement transparency in VPN operation	√		
VPN continuity when user roams	√	√	
Compatible with legal interception option		√	√
Compatible with NAT appliance			√

Table 1 : Comparison table of the proposed security models

4. List of References

[1] S. Kent and R. Atkinson, "Security Architecture for the Internet Protocol," RFC 2401, Nov. 1998.
[2] B. Gleeson, A. Lin, J. Heinanen, G. Armitage, and A. Malis, "A Framework for IP Based Virtual Private Networks," RFC 2764, Feb. 2000
[3] Wireless Application Forum (WAP), WAP specifications, URL: http://www.wapforum.org/what/technical.htm.
[4] D. Harkins and D. Carrel, "The Internet Key Exchange (IKE)," RFC 2409, Nov. 1998.
[5] L. Phifer, "The Trouble with NAT," *Cisco The Internet Protocol Journal*, vol. 3, no. 4, Dec. 2000, pp 2-13.
[6] 3GPP TS 33.102 (v3.12.0) 3G Security Security Architecture", release '99, June 2002.
[7] C. Xenakis and L. Merakos, "Security Architecture Standardization and Services in UMTS"," Proc. Mobile Venue 2002, Athens, Greece, May 2002, pp. 585-592, available from: http://www.telecom.ntua.gr/mobilevenue02/presentations/RNM_contribution_xenakis.pdf.
[8] C. Xenakis, E. Gazis and L. Merakos, "Secure VPN Deployment in GPRS Mobile Network," Proc. European Wireless 2002, Florence Italy, Feb. 2002, pp. 293-300.
[9] C. Xenakis and L. Merakos, "Dynamic Network-based Secure VPN Deployment in GPRS," Proc. PIMRC 2002, Lisboa, Portugal, Sept. 2002, pp. 1260-1266.
[10] C. Xenakis and L. Merakos, "On Demand Network-wide VPN Deployment in GPRS," IEEE Network, Vol. 16, No. 6, Nov/Dec. 2002, pp. 28-37.