Support of subscribers’ certificates in a hybrid WLAN-3G environment

Georgios Kambourakis *, Angelos Rouskas, Stefanos Gritzalis, Dimitrios Geneiatakis

Department of Information and Communication Systems Engineering, University of the Aegean, Samos 83200, Greece

Received 11 September 2004; received in revised form 25 April 2005; accepted 8 August 2005

Available online 22 September 2005

Responsible Editor: W. Wang

Abstract

Third Generation Partnership Project (3GPP) has recently provided a cellular-WLAN interworking architecture as an add-on to 3GPP system specifications. This architecture can offer IP-based services, compatible with those obtainable by 3G packet switched domain, to a 3G subscriber who is connected via a WLAN. Following this approach, in this paper we propose extensions to current 3GPP specifications, implementing and experimenting with a hybrid WLAN-3G network architecture capable of supporting subscriber’s certificates. We focus on attribute certificates, which are of major importance for user authorization and, due to their temporary nature, entail minimum concern regarding revocation issues. We emphasise on the necessary public key infrastructure incorporation which requires minimum changes in 3G core network elements and signalling and provide a list of the potential threats, which can be identified in a presumable deployment. Apart from the description and requirements of the proposed WLAN-3G architecture, particular emphasis is placed on the experimental evaluation of the performance of two alternative test-bed scenarios, which shows that digital certificates technology is not only feasible to implement in present and future heterogeneous mobile networks, but also can deliver flexible and scalable services to subscribers, without compromising security.

© 2005 Elsevier B.V. All rights reserved.

Keywords: Security; Certificates; Attribute certificates; PKI; Mobile and wireless networks; Heterogeneous wireless environments; Evaluation

1. Introduction

In the very near future, mobile users will want to access specific time-limited services, like buying something from an on-line store, settle down some stock transactions with a bank, or download a file from a protected site. This can be accomplished by using temporary or attribute short-lived certificates. Attribute Authorities (AA), or Certification Authorities (CA), bind the characteristics of an entity (called attributes) to that entity by signing the appropriate Attribute Certificate (AC) [1].
Attributes can specify group membership, role, security clearance, or other authorization information associated with the AC holder. Therefore, ACs are particularly well suited to control access to system resources and implement role-based authorization and access controls, accordingly [1,2]. They can also effectively implement and support popular authorization mechanisms such as Role-Based Access Control (RBAC) [3].

ACs are theoretically similar to Privilege Access Certificates (PACs), as used in SESAME and Windows 2000 operating system. The use of ACs has been included into both the ANSI X9.57 standard and the X.509 standards and recommendations of both ITU-T [4] and ISO/IEC, as a better alternative to X.509 public key certificates (PKC), for carrying authorization information. AC-based authorization is also an extension to the IETF Transport Layer Security Protocol (TLS). The basic structure of an AC is shown in Fig. 1. One of the advantages of these temporary certificates having a short life is that they do not usually need to be revoked and will therefore not need to be included in any Certificate Revocation List (CRL). If they are issued in respect of a pre-paid subscription service, they certainly not require any revocation at all. Finally, this mechanism can support non-repudiation services.

Another application area for ACs is mobile code technology, used by applications in wired and wireless computer networks in the last few years. Making code mobile means that programs or code segments are exchanged between computer networks and systems and the heterogeneity of platforms is hidden by a common language in which the program code is actually written [5]. A solution to protect the execution environment (e.g. mobile device), against potentially malicious mobile code is to authenticate the mobile code before it is actually executed. This approach is known as Shrink-Wrap. So, although it is not possible for someone to decide if a portion of mobile code contains malicious code, he can at least authenticate it. This can be very useful to a software developer who digitally signs the mobile code and distributes it together with the attribute certificate that is needed to verify the signature.

For example, let us consider a mobile Palm user who connects via GPRS in a mobile-portal and seeks for games. He wants to be sure that the gaming-code he decides to download is at least authentic. On the other hand, a developer who programs an application for specific phones wants to sign his code and put it along with the matching certificate in a mobile-portal. Therefore, he also needs to obtain an AC. In such an environment, we assume that there are many Attribute Authorities, which can issue that kind of certificates, certainly in collaboration with the service-offering parties. For example, if an organization already runs a directory service for public key certificates and related status information, this service can also be used to distribute ACs.

When implementing such scenarios with ACs, we need to examine interworking alternatives between the mobile core network and the necessary public key infrastructure (PKI) [6,7]. Taking into account recent 3GPP specifications about WLAN-3G interworking [8,9,11–13] in a Beyond-3G (B3G) vision, we propose a hybrid WLAN-3G architecture to support ACs issuing. 3GPP does not assume any specific type of WLAN system, but for the purpose of this paper we presume that the WLAN is of the

<table>
<thead>
<tr>
<th>Version</th>
<th>Holder</th>
<th>Issuer</th>
<th>Signature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serial Number</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attribute Certificate Validity Period</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attributes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Issuer Unique ID</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extensions</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The attribute field can specify group membership, role, security clearance, or other authorization information (octets) associated with the AC holder.

Fig. 1. Attribute certificate structure.
IEEE 802.11 type. The proposed architecture, which extends undergoing work by 3GPP, enables a Wi-Fi user, who is also a subscriber to a 3G mobile network operator, to move across WLAN segments administrated by different WLAN operators and to acquire on-demand ACs. Consequently, the user needs to know only his home 3G network operator, who is responsible of establish and maintain Roaming Agreements (RAs) with various intermediate visited 3GPP and ending WLAN operators. Additionally, we experiment with on-the-fly certificate generation, testing the performance of two prototype implementations. The measurements show that ACs issuing is attainable in terms of service time, while simultaneously can deliver flexible and scalable solutions to both future mobile operators and users. Finally, we list all possible impending threats suggesting, where applicable potential countermeasures.

The rest of the paper is organized as follows. In Section 2, we present and discuss a feasible interworking architecture between 3G core network and a PKI capable of providing certificates, under the assumption that the user may roam between different visited and probably heterogeneous network domains. Section 3 gives an overview of our experimental test bed and procedures, as well as a security risk analysis, while Section 4 presents the derived performance measurement results. The last section concludes the paper and points to future work.

2. 3G network architecture with PKI

Currently, 2.5G and 3G systems lack such a large-scale infrastructure, as PKI, to authorize and consequently charge mobile users for new services, as well as to provide digital signatures, certificates and non-repudiation services. However, in the years to come it is very likely that mobile operators will incorporate PKI technology or become associated to Trusted Third-PKI Parties (TTPs), also known as Certification Service Providers (CSP).

Successful wireless PKI implementations and solutions from companies like Sonera Smarttrust, Lucent Technologies and Entrust, strengthens the assertion that PKI has become an acknowledged and promising component of standards. Projects like ASPeCT [14] and USECA [15], 3GPP discussion papers for Universal Mobile Telecommunication System (UMTS) [16,17], as well as other papers [18,19], foresees that evolution. The eNorge 2005 strategy calls for a shared PKI for Norway [20], while advanced standards such MexE, WAP and i-mode from NTT DoCoMo have moved forward to introduce public key methods. Furthermore, WAP specifications [21] mention the use of role-certificates to be included in later versions.

More importantly, very recent 3GPP specifications documents for UMTS Release 6 [22,23] explore the possibility of deploying PKI and supporting subscriber’s certificates by mobile operators. However, 3GPP approach enables certificate issuing only from the Home 3GPP network. In addition, the whole mechanism is based on symmetric key derivation, after the user has been authenticated against a bootstrapping server, rather on long-term private keys. Finally, the issued certificates can be used to obtain certain services only by the home operator-controlled Network Application Function (NAF).

2.1. The proposed architecture

Fig. 2 depicts the proposed architecture that integrates PKI elements with 3G core network and conforms with 3GPP WLAN-3G interworking architecture specified in [8]. Note, that this scheme clearly extends and complements undergoing work by 3GPP and other forums. Consequently, all CA/AA network elements are assumed to be part of Network Domain Security (NDS), meaning that the protocols, needed for secure communications between them, are already in place. For example, as specified in [24,25], Mobile Application Part Protocol (MAP) or IPsec can be used to protect inter or intra communications.

The architecture introduces a new gateway element, which acts as a Certificate Provisioning Gateway (CGW) for the user. As a result, new IP interfaces and protocol messages have to be specified, for instance, between this CGW and the corresponding CA/AA. Of course, other alternative solutions can be proposed. For example, the direct connection of CA/AA with Gateway GRPS Support Node (GGSN) seems to be an attractive and natural choice. However, such a solution can only support certificate issuing from 3GPP environments. Additionally, it will require standardization of new messages between the User Equipment (UE) and GGSN, and GGSN and SGSN, as well as in active Packet Data Protocol (PDP) contexts. For instance, the user has to activate a secondary PDP context, if he wants to request a certificate from the visited network assuming that PDP
context activation is allowed by the visited network. Nevertheless, CGW can still be implemented, at the 3G-network side only, as a software module embedded in the GGSN. This solution will also reduce the corresponding costs for 3GPP network operators.

Obviously, before being capable of acquiring certificates, the user has to be successfully authenticated by the network. Current 3GPP specifications for UMTS Release 6 [8,9,11,12] describe an 3G-WLAN interworking architecture, where the home network is responsible for access control, while 3GPP Authentication, Authorization and Accounting (AAA) proxy relays access control signalling to the home 3GPP AAA server (see Fig. 2). Universal Subscriber Identity Module (USIM) based authentication mechanism can be based on the existing UMTS Authentication and Key Agreement (AKA) method. As this method should be independent of the underlying WLAN standard and should be supported by a standard authentication mechanism, 3GPP seems to choose the EAP-AKA protocol described in [11,26]. EAP is a general protocol for PPP authentication, which can support multiple authentication mechanisms. Consequently, EAP-AKA provides a way to exchange AKA authentication messages encapsulated within the EAP protocol. Other interesting works, which propose alternative authentication methods combining the AAA framework and the UMTS security features, can be found in [27,28].

2.2. Description and requirements

Before the user can acquire certificates, he has to be authenticated. Fig. 3 depicts a scenario where a 3GPP user is roaming in an area covered by a number of Wi-Fi access hot-spots. The user’s terminal performing active or passive scanning procedures defined in [29] can obtain a list of all available WLAN Service Set Identifiers (SSIDs). Once that list is made up, the UE will connect to the most preferred WLAN. For instance, the UE shall use a SSID that has a direct connection to home 3GPP network operator. If this is not possible, then the UE attempts to select a SSID that has a connection to one of the 3GPP operators in the preferred 3GPP operators list, stored in Universal Integrated Circuit Chip (UICC) card.

Network selection and authentication procedure are based on user’s Network Access Identifier (NAI) [31], which is of the form International Mobile Subscriber Identity (IMSI)@realm or Packet Temporary Mobile Subscriber Identity (P_TMSI)@realm. P_TMSI also known as pseudo-

Fig. 2. Generic architecture to support authentication and certificate issuing in 3G-WLAN environments (compatible with [8]).
nym, is generated as some form of encrypted IMSI in order to support user identity privacy (anonymity) in WLAN access [11]. Note, that NAI is used for authentication purposes in the IP multimedia subsystem (IM), introduced in UMTS Release 5 [32]. The access request is forwarded to the AAA proxy that translates the AAA request into the equivalent 3G AAA protocol request. The Access Point (AP) communicates with the AAA server that provides EAP server functionality using an AAA protocol, such as RADIUS [33] or Diameter [34]. During the authentication phase, EAP messages are encapsulated using a mechanism such EAP over LAN (EAPoL) between the UE and the AP and re-encapsulated in RADIUS messages from the AP to the home 3GPP service network. It is not our intention to provide a detailed description of the EAP-AKA mechanism [9,10]. However, the overall attach and authentication procedures are depicted in the first two parts of Fig. 4 for reasons of clarification and convenience to the reader.

The Home Subscriber Server (HSS) shall also check if there is a 3GPP AAA server already registered to serve this subscriber. In this case, it shall provide the previously registered AAA server address to the current AAA server with the previously registered AAA server address, causing authentication signalling routed to this server. At some point, the AAA server will respond to the UE indicating the status of the AKA authentication. If the procedure is successful and the subscriber is authorized to utilize WLAN or VPLMN assets, the AAA server will return the other essential certificate-related subscriber’s profile information, retrieved from HSS to the corresponding CGW.

Note, that in case the user is not directly connected to his home network, the home AAA server knows where to back-forward these profile data, because the EAP-Response Identity packet, previously received, contains, source WLAN and/or VPLMN IDs among other things (see fifth message transfer in Fig. 4). Another available option for 3GPP networks only, is to download this information to the corresponding SGSN, which then forwards it to the local CGW. Additionally, in case of roaming, e.g. in another VPLMN or WLAN domain, the authentication procedure is repeated, thus, there is no need to transfer any users’ profile data between CGW’s that belong to different providers.

As a final point, it is quite rational for each WLAN or 3GPP provider to install only one CGW in its domain, as the traffic load is not expected to be heavy. In order to increase availability to and perform better traffic load balancing, it is also possible to have another backup CGW, which is updated e.g. using mirroring techniques from the master CGW. For this reason, all the subscribers’ profile related-data are always downloaded to the same CGW in each distinct WLAN or VPLMN realm. As a result, there is no need to transfer users’ profile data between CGW’s that belong to the same provider.

This profile information will allow the CGW to decide if certificate issuing for this subscriber is permitted, which types of certificates can be acquired, the RA identifier applied, etc. After the user has been authenticated and has obtained an IP (actually, the UE can obtain two IPs an inner and an outer as specified in [8]), the UE as the case may be, has to find the appropriate home or visited network’s CGW. The UE can discover the CGW via one of the following methods:

(a) The address information shall be published. The UE shall store all the parameters as part of the initial establishment of IP connectivity.
(b) The address information shall be pushed automatically to the UE.
(c) The location information shall be discovered dynamically, based on DHCP, after IP connectivity has been established. The DHCP server shall inform the UE with the domain...
name of the local CGW and the address of a Domain Name Server (DNS) that is capable of resolving the Fully Qualified Name (FQDN) of the CGW (see the last part of Fig. 4).

(d) If the user has a direct connection to the home 3GPP network during PDP context, activate or update procedures [30].

Subsequently, certificate issuing is possible for the user at any time, assuming that this is in accordance with the credentials retrieved from HSS in the previous step. A certificate can be requested either from the CGW in the home network or from the CGW, if supported, in the visited network. The visited network can be a 3G or a WLAN network. Details on the certificate issuing procedure and the exchange of protocol messages, e.g. between CGW and CA/AA, are provided later in Section 3.2.

It is worth noting, that the UE has to support the EAP-AKA authentication mechanism, while
the subscriber’s profiles have to be updated to encompass new certificate-related fields. This will enable the home 3GPP operator to control the issuing of certificates, including the subscribers who are allowed to obtain them and the types of issued certificates. As noted earlier, trust issues between the home and visited network operators are subject to specific RAs or service agreements established beforehand. Besides that, PKI services can be offered by the 3G or WLAN operator itself or by another third party in the form of a CSP. In this case, supplementary service agreements between mobile service providers and associated CSPs have to be considered. More on trust issues between 3GPP and WLAN providers can be found in [11].

Standardization of messages between CGW and UE and CGW and AAA server (for transporting certificate-related fields) are required. Such a solution also assumes that the user has a dual mode mobile station supporting both WLAN and UMTS, or the WLAN device can be occasionally linked with a UE, which supports USIM capabilities (Bluetooth, USB or IrDA). Furthermore, this approach supposes the existence of at least one long-term private + public key pair per subscriber. Private key can be stored safely in his tamper-resistant UICC card and can be accessed, e.g. by an application (USIM), providing a password or a separate PIN known only to the user. How USIM or UICC functionality and Secure Signature Creation Devices (SSCD), which hold the private(s) keys of a subscriber, can be combined and how these keys can be obtained on demand, are described in [35,36].

Concluding, this architecture provides an access independent IP-based approach, which is relatively easy to deploy in current Packet Switched (PS) domain subsystems as in GPRS and hybrid 3G-WLAN environments, because it is in accordance with current 3GPP specifications and requires minimal changes to existing 3GPP core network elements and protocols. Moreover, this solution has minimum affection on WLAN standards and equipment. For example, it requires no modifications on legacy APs and no modifications of SSID broadcasting.

3. Experimental framework for certificate issuing

3.1. Test-bed setup

To test the feasibility of the aforementioned proposed architecture presented in Fig. 2, we used as a case study the delivery of ACs over IEEE 802.11b and GPRS networks. We constructed two experimental network architectures, which are illustrated in Figs. 5 and 6. The difference between these two topologies is the type of the network the user is connected to. In Fig. 5 the visited network is a WLAN, while in Fig. 6 the user is connected via GPRS to his home network. As already noted in the previous section, test-beds do not consider EAP-AKA authentication procedures, because certificate provisioning is possible after the user has been authenticated and obtained an IP. In other words, the authentication procedure is totally complementary to AC acquisition and consequently has no direct impact on the measurements provided below. Some interesting implementation and evaluation details on this issue are reported in [37].

The mobile device is a Compaq iPAQ H3970 Pocket PC (PPC) that uses Windows PPC 2002 operating system. The client uses a Nokia D211 dual GPRS class 7/WLAN IEEE 802.11b PCMCIA card inserted in iPAQ’s expansion pack plus module. The PPC incorporates a 400 MHz Intel X-Scale PXA250 CPU and has 64 MB of RAM and 48 MB of flash ROM available. It also utilizes a user-accessible section of ROM that can hold approximately 22 MB of data, applications, and other files.

The two (home and visited sub-network) CGW machines use Pentium 4 at 1 GHz processors and 128 MB of RAM. At the other end, the two CA/AA servers have Pentium 4 1.4 GHz processors with 192 MB of RAM. All the aforementioned devices run the Windows XP professional operating system.

All intra-network communications between CA/AAs and CGWs in our scenarios are protected by IPsec [41] Authentication Header (AH) protocol in transport mode. Inter-network communications in Fig. 5, are also protected by IPsec Encapsulating Security Payload (ESP) protocol in tunnel mode. IPsec uses Internet Key Exchange (IKE) protocol [42] for peer authentication. IKE can be configured to use pre-shared secrets or public key-based
authentication with certificates. In our implementations we used IKE main mode with digital signatures and certificates. Note that 3GPP also chooses IPsec and IKE with pre-shared secrets to secure network communications between UMTS IP inter- or intra-network entities [25]. As IPsec is operating only in the wired link, the proposed scheme does not suffer from well-known problems caused by this protocol when used in wireless links, like packet fragmentation and performance slowdown.

We wrote the applications in Microsoft’s Embedded C++ version 4.0 and employed the well-known open-source Apache-style license OpenSSL toolkit in version 0.9.7b (http://www.openssl.org) [43] to make them public key enabled and create the necessary certificates for the experiments. The necessary RAM space for the client, the CGW and the CA/AA applications to run, are 100, 98 and 96.1 Kb, respectively. The GPRS coding scheme was CS1 (9.05 Kb/s) and the time slots for GPRS were varying from 3 to 4, thus having wireless network speeds in the range from 27 to 36 Kb/s. Network speeds for 3G will be 144 Kb/s up to 348 Kb/s for wide and up to 2 Mb/s for low coverage and mobility, which will substantially reduce transfer times.

CA/AA server and CGW processes, which are multi-threaded, open TCP listening sockets, and wait for transactions. When they receive a message, they dispatch a thread to process and respond to the request. We also used a multi-threaded process to load the AA with virtual requests for AC certificate issuing. This process is running on another laptop machine that incorporates a Celeron 1.2 GHz processor with 256 MB RAM and is wired to the home or visited (sub)network with a speed up to 10 Mbps. The inter-arrival times between successive AC issuing requests, generated by that process, are exponential. A more realistic test-bed would require the inclusion of background internet traffic, or extra traffic from other sources. However, the purpose of our test-beds is to measure the additional overhead induced by our proposed elements and message exchanges, and thus we restricted our study on these only. Hereunder, we describe in detail the procedure of obtaining a temporary certificate from the CA/AA.

Fig. 5. Scenario A: Topology and test-bed.

Fig. 6. Scenario B: Topology and test-bed.
It is worth noting, that a PKI can also be implemented using an opensource product like OpenCA (www.openca.org) or a commercial product like those from RSA. Moreover, to increase security and support standard X.509 certificates and effective revocation mechanisms, providers can even install and use (distributed) mediated PKIs [44].

3.2. Details on the procedure for attribute certificate issuing

The aim of this section is to provide further implementation details, over request and AC creation and handling phases followed during our experiments. Depending on the nature of the request, the user or an automaton (process, daemon, service) on behalf of the user, constructs a certified request locally filling up values in the following fields. Here we supply some demo values. Of course, depending on the implementation, the request can also include (or exclude) some other fields e.g. the time that the request was created, identifiers of the hash and signature algorithms used, serial number of request, etc. Request’s creation time and serial number can be considered as mandatory in order to deal with replay-attacks discussed later in Section 3.3.

\[
\begin{align*}
COUNTRYNAME &= \text{"US"} \\
STATEORPROVINCENAME &= \text{"VA"} \\
LOCALITYNAME &= \text{"FAIRFAX"} \\
ORGANIZATIONNAME &= \text{"ZORG.ORG"} \\
ORGANIZATIONALUNITNAME &= \text{"SERVER DIVISION"} \\
COMMONNAME &= \text{"NAI(IMSI/P-TMSI@realm)"} \\
SUBJECTALTNAME &= \text{"DNS:195.251.161.167"} \\
TYPEOFREQUEST &= \text{"0.000" (bit pattern)}
\end{align*}
\]

The sixth field appoints the P-TMSI [11] assigned by AAA server during authentication or IMSI if temporary identity is not available, stored in the non-volatile memory of the UE. The last field in the form of a bit-pattern determines the type of the AC that the user wants to be issued (last three bits) and the issuing network’s (home or visited) CA/AA (first bit). For instance, a value of “1.101” for the last field designates: “Visited Network” and “Request type = 101”.

When the request-certificate creation phase is completed, the application incorporates the user’s public key, hashes the entire block using the MD5 function and signs the hash with the user’s RSA 1024 bits private key. 16_bytes_Hash = MD5 (Request + Public_Key) and Digital_Signature = (Hash)User’s_Private_Key. Note that the user’s long-term private key is stored in his UICC card and can be generated and associated with him during registration time or later on-demand as noted earlier in Section 2.2.

The certified request block (Request||User’s_Public_key||Digital_Signature), which is about 740 bytes, is then transmitted to CGW. According to Figs. 5 and 6, as soon as the visited network’s CGW receives the request, it validates its signature (generates the request hash, decrypts the signature using the public key in the certificate, compares the two hashes), thus certifying that it has not been altered in any way. If the request is valid, CGW checks the first bit of the last field to decide where to route it. The recipient can be the CA/AA in which the CGW is connected to or the CGW in the user’s home network. Naturally, in case the visited network is also the user’s home network, the request can only be sent to the home network’s CA/AA. Before transmission, the CGW can attach to the request possible certificate-specific and other requisite parameters from the user’s profile needed by the CA/AA to issue the certificate. Alternatively, request validation can be performed at the ending CA/AA instead of the serving CGW. However, if the request is not valid it will needlessly be forwarded to the corresponding CA/AA, which will discard it generating the analogous error code.

Upon reception, CA/AA shall issue and sign with its private key an AC based on the user’s values found in the request. The CA/AA will store the certificate, which is about 1 Kb, locally or in a corresponding certificate repository (LDAP directory or other repository). To end up, the certificate is delivered back to the user according to one of the following routes:

(a) CA/AA (home or visited network) → CGW (home or visited network) → UE.
(b) CA/AA (home network) → CGW (home network) → CGW (visited network) → UE.

All the above communications are organized into protocol message exchanges depicted in Figs. 4, 7 and 8.

When the AC is received by the UE, the UICC has to verify that it is valid and authentic. The procedure requires from the corresponding process,
daemon or service, to check the received AC against its integrity. Moreover, it will ensure the UICC that the AC has been published by a CA/AA it can trust. That is, generate the AC’s hash, the decrypt AC’s signature and compare the two hashes. It is important that the procedure use the public key from UICC’s internal store to decrypt the AC’s signature, rather than the public key that might be provided (with the AC) by the CA/AA. Given that changes on those lists are rare, a CA/AA list with their matching public keys can be stored securely in the UICC. This list can be periodically updated.

After validating the AC, the user is ready to use it according to the push model [1]. Another option is to pass to the application server the corresponding link where the certificate lies, instead of the actual certificate (pull model). Closing, issued certificates records held by CA/AA, can effectively provide, if needed, non-repudiation services. For example, assume that an individual has an account at an online brokerage. The individual buys shares of some stock using an application on his mobile phone, and then deliberately denies the transaction. It is difficult for the online broker to prove that the individual did indeed buy those shares. However, if the application used an AC when the trade is requested, the online broker can easily verify the transaction.

3.3. Identification of most serious attacks

This section focuses on the most serious threats and attacks undermining AC request and delivering procedures. Potential threats can be identified either in the confidentiality and integrity of the exchanged data (requests, ACs, subscriber’s profile fields) or in the signalling integrity between the communicating network entities. Another general but related set of attacks are also possible, like Denial of Service (DoS) attacks and unauthorized use of the subscriber’s private key. Other sorts of attacks like radio jamming or disassociation of legitimate users from APs e.g. when layer 2 control signalling is not integrity protected, are outside the scope of this paper. Although most of the attacks described here are performed by an attacker in the WLAN domain they may have serious consequences on the 3GPP network too. For instance, the easiest way for someone to get access to a given hot-spot is to simply become a registered subscriber. Attacks can also be launched remotely over the Internet.

3.3.1. Confidentiality, integrity and anti-replay protection

Normally, after performing the EAP-AKA authentication procedure, ciphering of data and integrity protection of signalling between the UE
and the AP (using link layer WEP, TKIP, CTR, application layer TLS, etc) or the Radio Network Controller (RNC) (using EAP-AKA derived 128-bit symmetric keys) are enabled. So, further protection, as discussed promptly, can be viewed as an additional option (possibly enabled by each provider itself).

Supposing that ciphering between the UE and the AP is weak, an attack can be exploited over the wireless link, as the integrity and confidentiality of data and signalling in the wired links is protected by IPsec, as noted earlier in Section 3.1. For example, a man-in-the-middle type of attack over the wireless link or simply eavesdropping on the traffic around an AP would enable the intruder to intercept the request or the AC. The same applies for rogue AP network attacks, where the attacker utilizes an AP masqueraded as legitimate in a given hot-spot.

However, even if a perpetrator manages to do so, e.g. by-passing WEP protection, he will not be capable to forge the request or the AC and use it for self-provisioning, as he does not possess the user’s or the CA/AA’s private key and thus he cannot reproduce the right signature. As a result, the undertaken risk is not so high even if the request and the AC are transmitted in clear-text, as they are almost useless for anyone who intercepts them.

To further increase confidentiality, e.g. in case the AC includes sensitive data like the "holder" field, the AC can be delivered back to the UE in encrypted form. Consequently, the CA/AA will have to encipher the issued certificate using the subscriber’s public key found in the request, before forwarding it back to the UE. This will not only enable the UE to verify the integrity of the AC, by decrypting it using the subscriber’s private key, but also will ensure the CA/AA that the AC can be utilized only by the legitimate user.

Encryption can be also applied to effectively protect the request. Following this syllogism, UE encrypts the request after signing it using the CA/AA’s public key. The only field that must not be encoded is the TYPEOFREQUEST, as it is needed by the intermediate CGWs to route the request. However, the integrity of the whole request, including the aforementioned field, is guaranteed (see Section 3.2). CA/AA will decrypt the request using its private key. This scheme also supports identity protection, as the aggressor cannot discover the identity of the user who asks for an AC. Naturally, when the transmitted request is encrypted, the intermediate CGWs cannot perform any validation. As a result, this task is left to the corresponding CA/AA. Some experimental results on request/AC encryption issues are provided later in Section 4.3.

As already mentioned in Section 3.2, to cope with replay attacks, for instance in case an eavesdropper records the communication between a UE and a specific CGW, the request must include time and/or request sequential number ID fields. Having this requirement fulfilled, it is very difficult for the potential aggressor to replay the request to the CGW-victim at a later time. This is true as the request is signed with user’s private key leaving almost no possibility to the attacker to forge it. In case of encryption, these fields must remain not encoded too.

3.3.2. DoS, DDoS and Service spoofing attacks

For a given transaction, another option is to return to the user a certificate pointer message encoded with the user’s public key, which contains the necessary information (e.g. a pointer, a username and a password) for the user to retrieve the certificate, instead of the actual AC. In this case, the certificate retrieval and forwarding are safer to be done by the user. If the client delivers a certificate URL, rather than the certificate itself, to an application server, he implicitly requests from the server to do the work (retrieve the certificate). The danger is obvious: a DoS attack is possible when a client deliberately passes an invalid certificate URL.

Of course, a DoS attack, e.g. by modification of signalling messages, requests, certificates, or even ACKs, is possible at any time. Nevertheless, this situation is analogous to radio jamming and is very hard to deal with. Another possible threat known as Distributed Denial of Service (DDoS) [45] aiming e.g. at CGW or AAA servers availability, can be performed from the Internet using “bots”. The bot (robot), could for example listen for connections on CGW’s port waiting a specific command from the attacker. Upon reception, it starts flooding a given IP address with packets.

Another identifiable attack is service spoofing. Considering this type of threat, the attacker impersonates one or several services in the local network, e.g. a DNS server, a DHCP server or in our case a CGW. This kind of attack can be performed e.g. by employing a rogue AP. Depending on the signal strength, some WLAN UEs may connect to the rogue AP. As a result, the aggressor can modify the user’s data or redirect the traffic to another network.
Analogous to the aforementioned threat is a *man-in-the-middle* type of attack, where the aggressor can cause the MAC and IP addresses of his WLAN UE to be bound to the identification credentials of a legitimate user. Depending on WLAN technology, and the default level of security chosen, the MAC and IP addresses of some UEs may be sent unencrypted, provoking the attacker to record them. This can be also done e.g. by analyzing and spoofing Address Resolution Protocol (ARP) packets. Hence, the eligible user is denied access, while the perpetrator gains access to services and resources normally authorized to the legitimate user. In a volume-based charging model, an attacker could inject packets choosing concrete source or destination MAC and IP addresses, targeting to the inflation of user’s invoice. For this reason, 3GPP is suggesting that WLAN operators should not use IP address based accounting.

Firewalls, Intrusion Detection Systems (IDS) like Snort (www.snort.org) and traffic analyzers like Ethereal (www.ethereal.com) can be utilized to counter fight DoS, DDoS and service spoofing attacks. Another solution to increase *availability* is to install e.g. a secondary-backup CGW, which can serve users requests when the primary CGW goes down after a DoS or DDoS attack. However, a detailed analysis, description and signature identification of all possible attacks and their corresponding countermeasures are left out for future work.

### 3.3.3. UE-oriented attacks

Open platform terminals may be infected by viruses, Trojan or other malicious software. Hence, in case the Mobile Equipment (ME) is not secure, the attacker may install a program that shows the user to acquire an AC for a “purchase of 10 € stocks” but ask the USIM to sign a different request like “purchase of 1000 € stocks”. Also if that program manages to find the PIN or password to access the subscriber’s private key, it can command the USIM to generate signatures even in ignorance of the user. In this case the legitimate subscriber would have to pay for services he did not actually solicit.

Trojans may perform analogous activities such as monitor the user’s keyboard for private data and forward the information to the attacker’s machine. Generally, the subscriber’s private key is better protected when it is kept in the UICC, rather on the ME. An interesting work on topics surrounding mobile appliances security can be found in [46].

### 3.3.4. Privacy considerations

In case the pseudonym of the subscriber (P_TMSI) is not available, the subscriber’s IMSI has to be included in the request. This consist a breach in user’s privacy, as the intruder may detect who is obtaining an AC and probably in which location the subscriber is roaming (location privacy). An obvious solution here is to have the request and the AC encrypted as discussed earlier in this section. Finally, as stated in the introduction, ACs have limited time frame validity, meaning that revocation is not necessary. The holder is bound to use the AC expeditiously before it expires.

### 4. Measurement results

We experimented with various values for the arrival rate of AC requests, which determines the virtual load offered to the CA/AA. We varied this parameter from 20 to 60 requests per minute and the effect on the server performance was negligible. Measurements were gathered from a set of 2000 transactions between the CA/AA server and the client. Our experiments were conducted in different days and hours during a week period and 50% of the measurements were logged during peak hours. We tracked and measured the following times (Table 1):

#### 4.1. Scenario A: The visited network is a WLAN

Scenario A is depicted in Fig. 5. It consists of two sub-networks with an average ping time of about 80 ms between them. The requests are delivered to the visited network’s (WLAN) CGW and then are forwarded to the CGW in user’s home network. The ACs are issued by the user’s home network CA/AA. The average values and standard deviations of the time durations measured are presented in Table 2 and the probability density function (PDF) of client’s total time (CROT) is shown at the left side of Fig. 9.

#### 4.2. Scenario B: The visited network is a GPRS

Scenario B is depicted in Fig. 6. It consists of two sub-networks with an average ping time of about 1230 ms between them. The client is connected to its home network via GPRS and the requests are delivered to the home network’s CGW. The ACs are issued by the user’s home network CA/AA. The average values and standard deviations of the time durations measured are presented in Table 2 and the probability density function (PDF) of client’s total time (CROT) is shown at the left side of Fig. 9.
Table 1
Description of service times measured

<table>
<thead>
<tr>
<th>Time</th>
<th>Description</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Client side</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRCT</td>
<td>Client request creation time</td>
<td>Time for the user’s device to create the request. This is done by an automaton (daemon) installed in user’s device</td>
</tr>
<tr>
<td>CRTT</td>
<td>Client request transmission time</td>
<td>Elapsed time from request transmission, until the client receives an ACK from the CGW in visited/home network. (The request has been successfully delivered to CGW and it is valid.)</td>
</tr>
<tr>
<td>CRRT</td>
<td>Client request return time</td>
<td>Elapsed time from CGW-ACK, until the AC has been received</td>
</tr>
<tr>
<td>CROT</td>
<td>Client request overall time</td>
<td>Elapsed time from request transmission until the AC has been received and an ACK has been sent back to CGW. This time contains CRTT and CRRT</td>
</tr>
<tr>
<td>CGW side</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GWVT</td>
<td>CGW request verify time</td>
<td>Time for the CGW to verify the request (recalculate hash, validate signature, decide where to route the request)</td>
</tr>
<tr>
<td>GWFT</td>
<td>CGW request forward time</td>
<td>Elapsed time from request forwarding, until the CGW receives an ACK from the CA/AA. (The request has been successfully delivered to the CA/AA.)</td>
</tr>
<tr>
<td>GWOT</td>
<td>CGW overall time</td>
<td>Elapsed time from request forwarding, until the attribute certificate has been delivered back to CGW</td>
</tr>
<tr>
<td>CA/AA side</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AACT</td>
<td>Attribute authority certificate creation time</td>
<td>Time for CA/AA to create the AC, according to the request</td>
</tr>
<tr>
<td>AART</td>
<td>Attribute authority certificate return time</td>
<td>Elapsed time from AC forwarding, until the AA receives an ACK from the CGW. (The AC has been successfully delivered to the CGW.)</td>
</tr>
</tbody>
</table>

Table 2
Average service times in milliseconds for scenario A

<table>
<thead>
<tr>
<th>Time</th>
<th>CRCT</th>
<th>CRTT</th>
<th>CRRT</th>
<th>CROT</th>
<th>GWVT</th>
<th>GWFT</th>
<th>GWOT</th>
<th>AACT</th>
<th>AART</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>1074.3</td>
<td>2955.8</td>
<td>144.1</td>
<td>3144.9</td>
<td>10.8</td>
<td>0.6</td>
<td>73.8</td>
<td>62.1</td>
<td>0.2</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>31.1</td>
<td>92.6</td>
<td>171.8</td>
<td>196.6</td>
<td>55.0</td>
<td>5.7</td>
<td>26.2</td>
<td>11.3</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Fig. 9. CROT probability density functions for scenarios A and B.

CA/AA. The average values and standard deviations of the time durations measured are presented in Table 3 and the probability density function of client’s total time is shown at the right side of Fig. 9.

4.3. Comments on the measurement results

As we see, the average client’s total time for one transaction, is about 3.2 and 4.4 s when the user is roaming to a WLAN network or he is connected
via GPRS, respectively. These measurements are also reflected in the PDFs of Fig. 9. One can say that these times are not differing greatly. Nevertheless, the WLAN case includes the extra network delay for CGW-to-CGW roundtrip communications.

GPRS values are highly concentrated around 3.8, while WLAN values near 3.2. Naturally, these times are expected to grow depending on the visited and home network’s CGW distance expressed in ping times. The more the numbers of domains the request has to travel, the more CROT is expected to be. The increased standard deviation times for the client, which are spotted in Table 3, are explained considering the unreliability of the GPRS connection. On the other hand, the corresponding WLAN standard deviation times are typical for this type of connection.

Another important issue for scenario B is the extra network delay derived from the fact that the CGW and CA/AA did not reside inside the mobile provider’s core network. The request as well as the AC, has to traverse all the way back to the local network where the CGW and the CA/AA are located. Considering the average ping time between those domains we can presume the extra time needed for this task. Of course, this time is expected to differ depending on the distance between the SGSN, CGW and CA/AA inside the provider’s core network. Naturally, a more realistic test-bed would require the CGW and CA/AA installment inside the 3GPP provider’s network, which, in fact, is very difficult to arrange.

Finally, we gathered some measurements using RSA 1024 bit public keys to encrypt and decrypt requests and ACs, to deal with extra protection procedures raised in Section 3.3. The results (Minimum–Maximum values in milliseconds) from 10 runs performed on three devices with a key-length 1024 bits (PKCS#1 v.1.5 and padding 117 bytes) are summarized in Table 4. We notice that this extra protection comes at a quite reasonable cost, but is totally optional for the provider to enable.

5. Conclusions and future work

As users rush to adopt IP technology and want wireless access to IP networks, they also become aware of the need for security features and protection of their privacy. The constantly increasing population of users expects from mobile operators to provide features that will provide reliable authentication, authorization and accounting mechanisms and offer availability and quality comparable to that of the wired services. Thus, more flexible, dynamic and scalable mechanisms are necessary in order to support on-demand services and all-IP end-to-end solutions in a many-to-many trust model integrated with the Internet environment.

Based on the assumption that the necessary PKI to support ACs is about to be incorporated in the mobile network in the near future, in this paper we proposed and analyzed a viable hybrid 3G-WLAN network architecture. The anticipated scheme extends current 3GPP specifications, being capable of providing digital certificates to the 3G subscribers independently of the underlying access network. Focusing on attribute certificates, we experimented with on-the-fly certificate generation, testing the performance of two prototype implementations based on the discussed architecture. Furthermore, we listed all possible impending threats suggesting,
where applicable, potential countermeasures. Results showed that ACs issuing is attainable in terms of service time, while simultaneously can deliver flexible and scalable solutions to both future mobile operators and users.

Topics to be further investigated include mobility issues when the request has to cross multiple visited domains and possible cross certification procedures between CA/AAs, which belong to different visited WLAN or 3GPP domains that have roaming or service agreements with the home 3GPP network operator. Another important topic in common with RAs is trust issues between the CA/AAs operated by or collaborating with 3GPP or WLAN providers and end service providers. The possibility of the home or visited CA/AAs to ask directly from the HSS the necessary parameters from the subscriber profile, instead of downloading it to the serving CGW can be an interesting issue as well. This scenario can be particularly convenient when an AC is requested by the home 3GPP network. Finally, we are also planning to deploy OpenCA to further experiment and evaluate the proposed architecture in a wider scale.

Acknowledgements

We would like to thank the editor and the reviewers for their valuable comments that helped us to improve the overall quality and presentation of our work.

References

[23] 3GPP Technical Specification, Bootstrapping of application security using AKA and support for subscriber certificates; system description, TS ab.cde v.0.3.0, September 2003.


involved in several national and EU funded R&D projects in the areas of Information and Communication Systems Security. His published scientific work includes several books on Information and Communication Technologies topics, and more than one hundred journal and national and international conference papers. The focus of these publications is on Information and Communication Systems Security. He has served on program and organizing committees of national and international conferences on Informatics and is an editorial advisory board member and reviewer for several scientific journals. He was a Member of the Board (Secretary General, Treasurer) of the Greek Computer Society. He is a member of the ACM and the IEEE.

Dimitrios Geneiatakis was born in Athens, Greece, in 1981. He received the Diploma in information and communication systems in 2003, and the M.Sc. in security of information and communication systems in 2005, both from the department of Information and Communications Systems Engineering of the University of Aegean, Greece. His current research interests are in the areas of Security mechanisms in Internet telephony, Smart Cards and Network Security. He is a member of the Technical Chamber of Greece.