3RD QUARTER 2007, VOLUME 9, NO. 3

IEEE COMMUNICATIONS SURVEYS
The Electronic Magazine of Original Peer-Reviewed Survey Articles

www.comsoc.org/pubs/surveys

SURVEY OF SECURE HANDOFF OPTIMIZATION SCHEMES FOR MULTIMEDIA SERVICES OVER ALL-IP WIRELESS HETEROGENEOUS NETWORKS

GIORGOS KAROPOULOS, GEORGIOS KAMBOURAKIS, AND STEFANOS GRITZALIS,
UNIVERSITY OF THE AEGEAN

ABSTRACT

In the very near future, we shall witness the coexistence of networks with heterogeneous link layer technologies. Such networks will naturally overlap each other and mobile users will need to frequently handoff among them for a number of reasons, including the quest for higher speeds and/or lower cost. Handoffs between such hybrid networks should be fast enough to support demanding applications, like multimedia content delivery, but also secure enough since different network providers are involved. This gets even more complicated considering that network providers may not simultaneously be multimedia service providers as it is the case today. In order to support security operations in a large scale the employment of an AAA protocol is mandated; however, this adds more delay to the handoff process. This article analyses and compares the prominent methods proposed so far that optimize the secure handoff process in terms of delay and are suitable for uninterruptible secure multimedia service delivery.

As the Internet has become an essential part of our everyday life, the progress of communication technologies will offer an "always connected" opportunity to everybody. To support this vision, the research community is suggesting a move towards an all-IP platform in order to take advantage of the high bandwidth of WLANs and the broad coverage of cellular networks and WMANs. The convergence of these heterogeneous wireless technologies will eventually lead to the 4th generation (4G) of mobile communication systems.

This environment, however, will not only provide the basis of new and better applications, but shall impose new technical problems as well. The trade-off between security and efficiency is one of the most challenging issues in wireless communications and this is not likely to change, at least in the near future. This trade-off is especially true in environments where the network provider is different from the service provider. In such cases, the end user must be authenticated to both providers in order to use a single service, and to many more if he plans to use more services or perform a handoff. That is, authenticate to the network provider and additionally perform a number of authentications as many as the service providers, or in case of a handoff, authenticate to the new network provider and re-authenticate to the services he already receives. For example, considering SIP registration in UMTS networks, the user must first authenticate in order to access the network and then authenticate (again) to access SIP services. These authentications are accomplished using AAA protocols, like RADIUS [1] or DIAMETER [2], which are in general costly, especially when the home network is many network hops away from the visiting network. This delay is even more crucial and must be carefully considered during handoffs.

Mobility management protocols like Mobile IP [3] and Cellular IP [4] do not consider AAA operations during handoff. In order to cope with long delays, a number of techniques have been proposed to optimize the handoff procedure. One way to achieve this is by minimizing the delays introduced by AAA interactions during the handoff phase. This survey looks into such schemes and compares them in terms of security, efficiency and scalability. Our work also provides a short description and a critical constructive view of each method

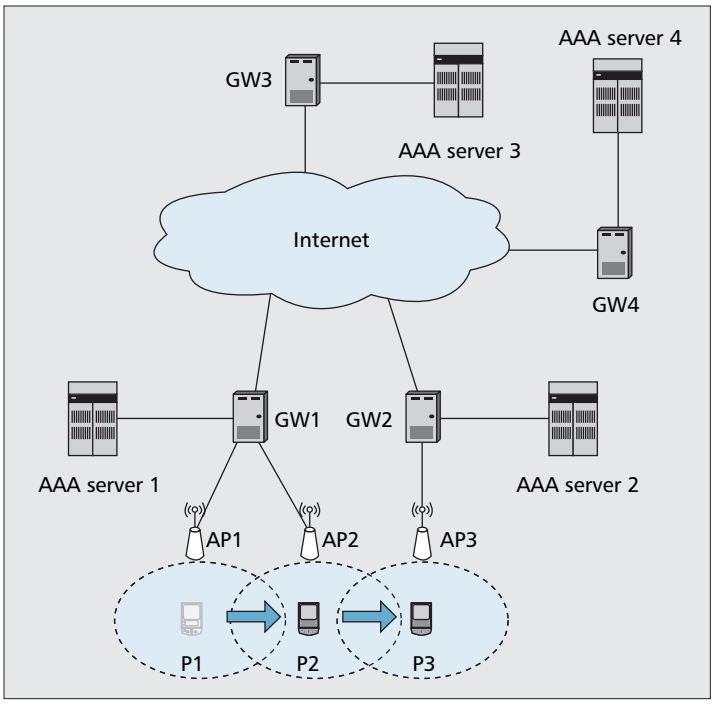

Figure 1. *General heterogeneous network architecture.*

presented.

To better analyze the problem, in the following we describe a typical scenario of using multimedia services over all-IP wireless heterogeneous networks. Under this context, a user terminal can roam between networks that utilize different access technologies like IEEE 802.11, 802.16 and UMTS. Each of these networks may belong to the same or to a different administrative domain. For example the user (terminal) is able to move from a WLAN to another WLAN, which belongs to the same operator (performing an intra-domain handoff) or from a WLAN to an UMTS network, which belongs to a different administrative domain (performing an inter-domain handoff). In general, inter-domain handoffs tend to be more expensive than intra-domain because of the network delays imposed by the distance, in terms of network hops, between the local and the home domain.

Figure 1 depicts the general architecture of a network composed of different technologies and administrative domains; for instance, access points AP1 and AP2, which reside in the same administrative domain, could use 802.11, and AP3 could be a cellular operator's access point. Each domain is represented by a gateway (GW) and an AAA server. This, of course, is a simplified representation and every gateway could act either like a true gateway to the Internet, or a multimedia server, or a directory for AAA servers lookup, etc.

Next, we consider a scenario where a terminal using a multimedia service from a server residing out of the local domain is executing an inter-domain secure handoff. Initially the terminal is at position P2, using a multimedia service from GW3 and its home domain is controlled by AAA Server 4. If the user moves to position P3, a handoff is going to occur. What should be assured during this handoff is the continuation of the multimedia service without severe quality degradation. The procedure that has to take place is as follows: the terminal first requests access to the network from GW2, which refers to its local AAA Server 2 which in turn refers to AAA Server 4 to authenticate it. After the terminal is granted access to the network it must access the multimedia service, so through GW2 the terminal requests the service from GW3, which refers to its local AAA Server 3, which in turn refers to AAA Server 4 to authenticate it.

The aforementioned example is the worst-case scenario where the home domain, the local domain and the multimedia server reside in three different places. In such cases the operations taking place during the handoff procedure result in long delays, especially when the involved servers are distant from each other. However, the previous worse case situation describes a non-optimized scheme, which does not consider the problems related to multimedia services during secure handoffs. This article is concentrated on schemes that try to solve or mitigate these problems.

The rest of the article is organized as follows. We describe the cardinal until now secure handoff optimization solutions together with some critical comments where this is necessary. We provide a comparison of the schemes in question, while the comparison criteria are analyzed and the schemes are thoroughly compared to each other. Finally, we offer concluding thoughts and future directions for this work.

Schemes Description

In the following we constructively describe all the major secure handoff optimization schemes proposed until now. This is necessary for the qualitative analysis provided later. Moreover, for the sake of completeness, we decided to only reference a number of other subordinate schemes which bare similarities with the ones presented hereupon.

OIRPMSA

In [5] the authors are examining the case of a secure handoff using Mobile IP [3] and SIP [6]. Their scheme namely "Optimized Integrated Registration Procedure of Mobile IP and SIP with AAA operations" (OIRPMSA) attempts to reduce the roundtrips needed between the mobile terminal and the home AAA server. Normally, 3 such roundtrips are needed:

1 Mobile IP registration (Fig. 2).
2 SIP register (Fig. 3, actions 1–6): This message gets a 401 (Unauthorized) response which, among others, includes challenge information.
3 SIP register (Fig. 3, actions 7–12): The terminal tries to authenticate using the previous challenge information and if the authentication is successful it gets a 200 (OK) response.

The suggestion of this work is to minimize the delay imposed by the second message by "converting" it to a local roundtrip between the mobile node (MN) and the Local AAA Server (AAAL). Their idea is illustrated in Fig. 4 and is as follows: when the MN sends the first message (Mobile IP registration) it states that a SIP registration is about to follow. The home AAA server's (AAAH) response includes some chal-

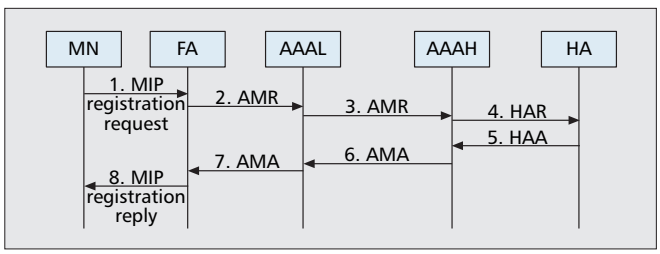

Figure 2. *Mobile IP registration with AAA operations [5].*

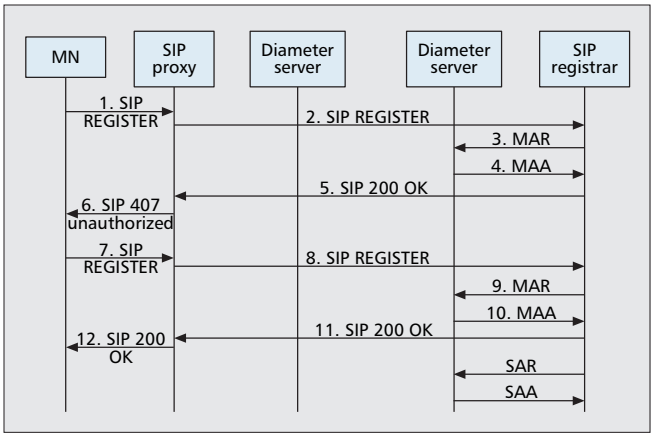

Figure 3. *SIP registration with AAA operations [5].*

lenge information which is stored in the local AAA server and will be used later. Then the mobile terminal sends a SIP register (action 9) towards the local AAA server which responds with a 401 (Unauthorized) response (action 10) that includes the previous challenge information. Finally, another SIP register message follows that goes all the way to the home AAA server (actions 11–16).

One shortcoming of this approach is that it is assumed that the network provider is the same as the service provider. Although this, in many cases, is true today, it is not the general case and of course it is not certain that it will still hold after a few years. Another weakness of this scheme is that the agents used by Mobile IP (FA-Foreign Agent, HA-Home Agent) should be co-located with the SIP proxy and SIP registrar respectively as shown in Fig. 4.

MPA

Another handoff optimization scheme, presented by Dutta et al. [7, 9], is MPA which stands for "Media — independent Pre — Authentication." MPA is a framework that can work over any link layer and can cooperate with any mobility management scheme. To support this claim some of the authors in another work [9] have combined MPA with IEEE 802.21 [10] as mobility management protocol. MPA framework assumes that the following elements exist in every network: Authentication Agent (AA), Configuration Agent (CA) and Access Router (AR). The basic steps taken by MPA are as follows:

1 Pre-authentication (Fig. 5, action 1): The mobile terminal finds the IP addresses of AA, CA and AR. It performs pre-authentication with the AA, creating security associations with AA, CA and AR.

2 Pre-configuration (Fig. 5, actions 4–5): When the mobile node is about to change its point of attachment, it performs pre-configuration using the CA to obtain new IP address and other configuration parameters (action 4). Using a tunnel management protocol, the mobile node sets up a tunnel with an access router from the candidate network (action 5).

3 Secure proactive handover (Fig. 5, actions 6–7): The terminal starts a binding update over the established tunnel by using both the old and the new IP addresses. This means that it has already executed a higher layer handoff before a link layer handoff.

4 Switching (Fig. 5, actions 8–9): The mobile node completes the binding update and executes the link layer handoff. After that, the mobile node starts communicating from the new point of attachment and deletes or disables the established tunnel.

In [7] a complete handoff solution is proposed which optimizes a number of parameters that add to handoff delay, like IP address assignment. In the testbed implemented by the authors, a non-MPA handoff took 4 seconds, whereas MPA assisted handoffs to different platforms took from 14 to 600 ms.

Shadow Registration

In [11] the Shadow Registration method is proposed in order to optimize secure handoffs. According to this scheme a security association is established between the mobile terminal and every neighboring AAA server before the former enters the domain the server controls. Using Fig. 6 as reference, when the mobile terminal resides in the central cell, a registration procedure is performed with all 6 neighboring cells. When this happens the necessary AAA operations are processed locally in this new domain without communicating with the terminal's home domain. Specifically, the authors examine two cases where Shadow Registration could be used; the Mobile IP case and the SIP case. In both cases, during the handoff, the requested AAA operations are processed locally and after the completion of the handoff a separate process is executed where security associations are sent to the new neighboring domains of the mobile terminal.

Based on the concept of Shadow Registration, Han et al. [12] have proposed Region-based Shadow Registration (RSR). RSR is trying to solve the problems of heavy traffic and waste of resources introduced by Shadow Registration. Instead of establishing a security association with every neighboring domain, RSR divides the terminal's current cell in regions (a, b and c in Fig. 6) and performs a Shadow Registration only when the terminal moves to a section with high probability to handoff. When the mobile node resides near the cell core, no Shadow Registration is performed. The outer zone of the cell is divided in three regions and each region is adjacent to two neighboring cells; when the mobile terminal moves to one of these three regions, a Shadow Registration is performed for the two neighboring cells. For example in Fig. 6, when the mobile terminal moves from the Core to region b, a Shadow Registration procedure is performed with cells 3 and 4. By this way, the two schemes have the same effect in reducing the handoff delay while RSR reduces the traffic between the domains.

Another similar approach is [13] which uses the Frequent

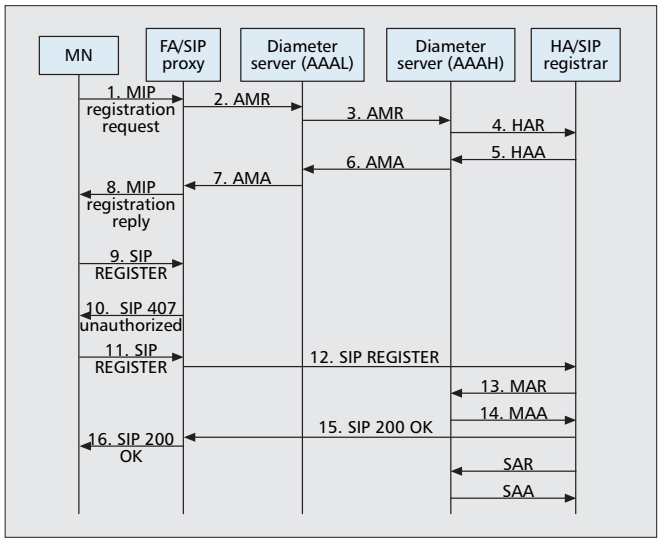

Figure 4. *OIRPMSA signaling [5].*

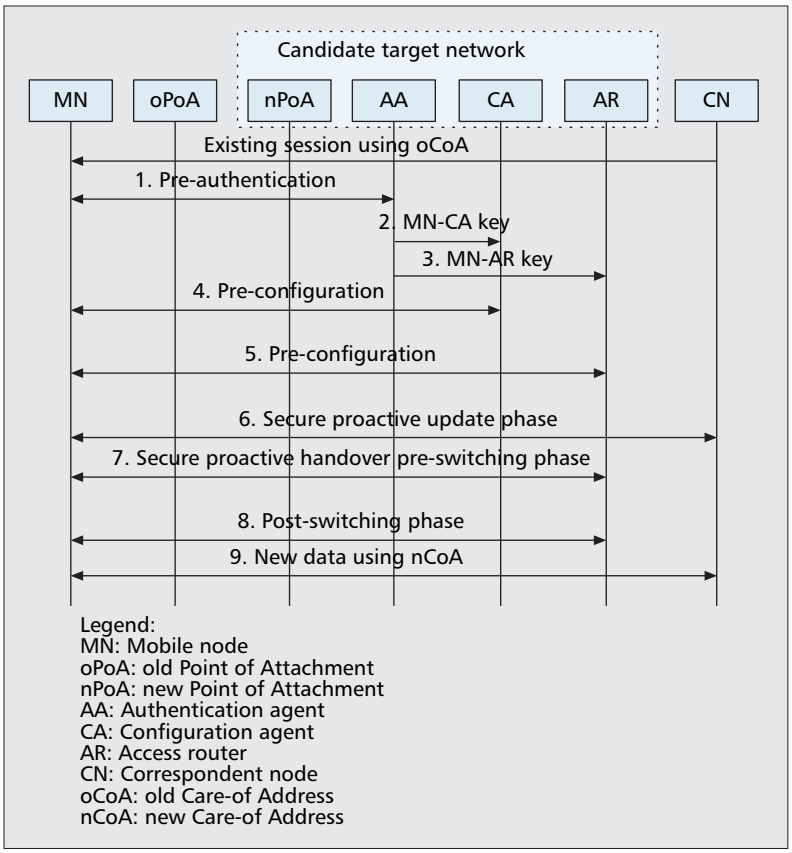

Figure 5. *MPA signaling flow [7].*

Handoff Region (FHR) concept. Considering this scheme, the network administrators collect information about the location of the access points and the movement of users and construct a weighted directed graph. With the help of FHR Selection algorithm, adjacent access points are grouped in FHRs and the mobile terminal is authenticated in advance towards the access points that belong to the same FHR.

A disadvantage of the above methods stems from the fact that in future heterogeneous networks the areas of coverage in most cases will overlap. In such an environment, when a mobile terminal roams in an area covered by a WLAN access point it is possible that this area is also covered by other WLAN, WMAN and/or UMTS access towers. Under these circumstances it is not obvious which the neighboring domains are, let alone that there can be many of them. This results to excessive signaling (especially in SR) and difficulties in determining which the neighboring cells (in RSR case) are. This maybe not seems to be a problem with FHR, but that would require from the administrators to collect new information every time a new network is deployed in the same area.

AAA Context Transfer

The solution proposed in [14] is product of the IST EVOLUTE project that tries to provide secure and seamless multimedia services over heterogeneous all-IP networks using the concept of context transfer. In RFC 3374 document [15] the context and context transfer terms are defined as:

- Context: The information on the current state of a service required to re-establish the service on a new subnet without having to perform the entire protocol exchange with the mobile host from scratch.
- Context transfer: The movement of context from one router or other network entity to another as a means of re-establishing specific services on a new subnet or collection of subnets.

EVOLUTE uses Mobile IP and SIP for inter-domain mobility management, while for intra-domain mobility uses protocols like Cellular IP and Hierarchical Mobile IP [16]. In order to provide secure access to multimedia services, the EAP-TLS [17] solution is used as the authentication protocol. Figure 7 depicts the signaling flow when the context transfer is not used. On the downside, Fig. 8 shows the same signaling flow when the context transfer is enabled. When the mobile terminal sends a request to handoff to a new gateway (NGW), this NGW gets the context from the previous gateway (PGW) whose IP is indicated in the terminal's request. The terminal is then authenticated to the NGW without contacting its home domain.

When the method of context transfer is employed, it is assumed that the new network can support the services offered from the previous one. However, in a heterogeneous environment, this might not always be the case and the mobile terminal might have to contact its home domain for renegotiation about the offered services.

Peer-to-Peer Security Context Transfer

The work of Braun and Kim [18] combines the concepts of security context and P2P networks to optimize authentication in heterogeneous wireless networks. According to this approach a security context contains authentication credentials in the form: {random number or nonce or challenge, expected response}. Such security contexts are created at the home domain by the home AAA Server and delivered to AAA Servers (or Brokers) that reside between the home domain and the local domain (and therefore are closer to the mobile terminal). The AAA Brokers take the authentication decision after a corresponding mobile terminal's request based on security contexts; for this reason they are referred as Security Context Controllers (SCCs) as well. SCCs are organized in a peer-to-peer manner and they are able to detect each other using mechanisms originated from P2P networks.

An example demonstrating peer-to-peer organization of SCCs is illustrated in Fig. 9. At first, the mobile terminal resides in the area covered by SCC 1 which has already acquired the security context from AAAH via SCCx. During this transfer, AAAH and SCCx have stored pointers to the current security context which resides in SCC 1. When SCC 1 gets the security context it broadcasts its acquisition to its neighbors. This way, when the user moves to the area covered by SCC 2, the new SCC acquires the security context from SCC 1 and informs AAAH. If this is not the case, say the user switches off the device in SCC 1 and moves to the area of SCC 3, SCC 3 is not aware of the stored security context in SCC 1 and has to request a new one from AAAH. This request is routed through SCCy and SCCx; when the request meets SCCx, SCCx returns a response that SCC 1 has a security context for the corresponding user. When SCC 3 gets this response it requests the security context from SCC 1 and informs AAAH that is now controlling the security context of the user.

Optimistic Access

In order to minimize the re-authentication delay, an alterna-

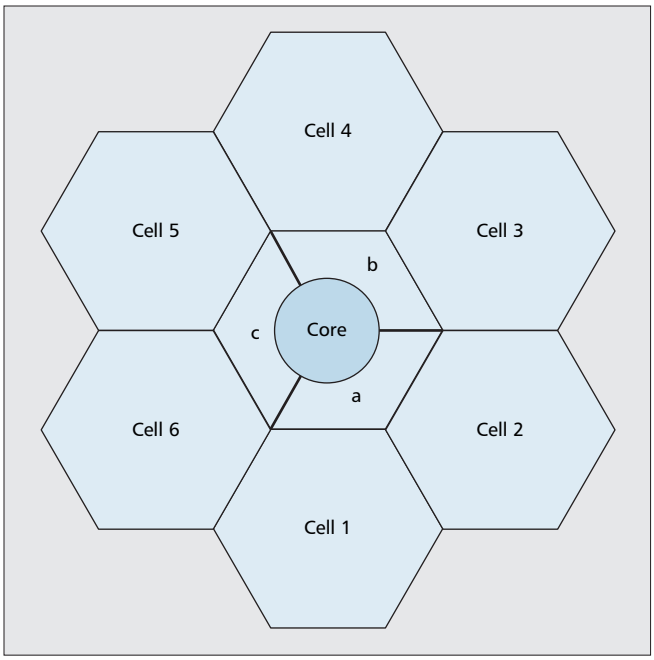

Figure 6. *Regional cell division [12].*

tive technique is proposed by Aura and Roe in [19]. According to this approach the mobile terminal, instead of executing a so-called strong authentication during the handoff process, it is granted optimistic access to the new network delaying the strong authentication which is held after the handoff is completed.

More specifically as shown in Fig. 10, when the mobile node (MN) handoffs to the new network a light (fast) authentication takes place. If this authentication is successful the MN is authorized a so-called optimistic access and can communicate through the new network. When the handoff process is complete, the MN must be involved in a new strong authentication to continue using the resources of the new network. After the end of this authentication the Optimistic Access scheme completes its purpose.

The target of the proposed work is to conclude to a protocol that allows optimistic access to well behaving mobile nodes while reducing the risk of possible misuse. A customer that has paid for some other service or is following some rules is considered well-behaving, while unknown users should perform a strong authentication during the handoff process. The protocol operates as following:

- The old access point sends to the mobile terminal a secret key K and a credential C over a secure channel.
- The new access point broadcasts challenges periodically and the mobile terminal retrieves one such challenge.
- The terminal computes and sends towards the new access point a keyed one-way function of the challenge and the secret key K. It also presents the credential C acquired from the old access point which contains some trust parameters about its previous good behavior. The new access point recovers K from the credential and decides to grant optimistic access or not by evaluating the trust parameters.

When the secure handoff procedure is completed, the strong authentication must take place in a short time. The exact period of optimistic access and authentication method are not covered by the above protocol and are matters of choice of the network administrator.

This protocol makes an obvious trade-off between security and performance. The vulnerability left is the small window of light authentication between the handoff and the strong authentication; still, in order for someone to misuse the resources of the network, the light authentication should be based on a weak protocol or no authentication at all. Another security issue of the above protocol is that all the access points of the network should share a secret key and this can be especially dangerous if a single access point leaks the key. Thus, key management issues concerning optimistic access must be carefully considered and further investigated.

Other Schemes

This section references secure handoff optimization schemes that could not be included in the conducted comparison. These methods are left out because they either exhibit many similarities with the mechanisms already described or they do not comprise a general solution supporting secure handoffs between heterogeneous networks for multimedia services delivery.

The first such scheme is Mobility-adjusted Authentication Protocol [20] (MAP) which utilizes symmetric cryptography in conjunction with the security context concept relying on special Security Context Nodes (SCNs). The work in [21] reviews fast authentication methods for 802.11 WLAN's for seamless mobility across administrative domains. The authors of [22] use the concept of AAA brokers while their novelty is a formula for finding the best spots within the network architecture to place these brokers. A method which is similar to the Shadow registration concept, and especially to the Frequent Handoff Region variation, is presented in [23]; the difference here is that this method does not require manual configuration and the system is auto configured instead. In [24] the authors are based on Hierarchical Mobile IPv6 (HMIPv6) which is an enhancement to Mobile IPv6 (MIPv6) that supports fast handoffs. Their proposal integrates the Diameter protocol to support authenticated access during roaming. Another approach is the Secure, QoS-enabled Mobility (SeQoMo) [25] architecture which is comprised of components that can be co-located with existing routers, access points, mobile nodes etc. to provide fast handoffs to HMIPv6 based networks. In [26] six approaches are proposed for session state re-establishment in intra-domain scenarios; these approaches are based on the combination of concepts like Fast Handover for Mobile IPv6 (FMIPv6), HMIPv6, AAA

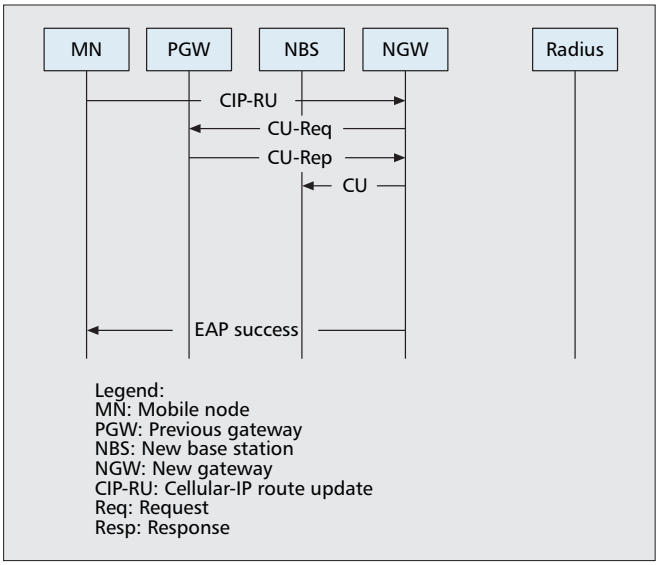

Figure 7. *EAP-TLS exchange without context transfer [14].*

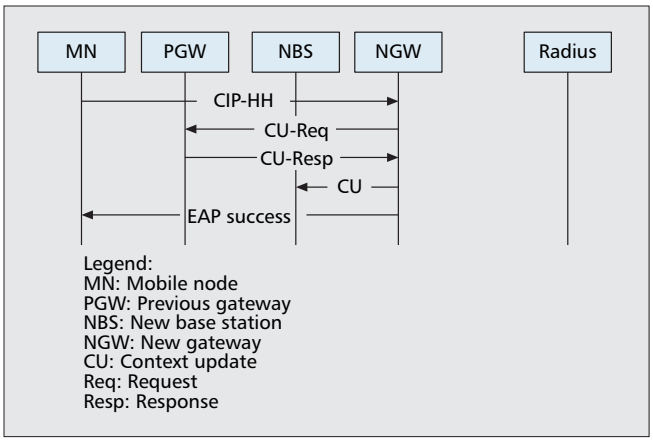

Figure 8. *EAP-TLS exchange with context transfer [14].*

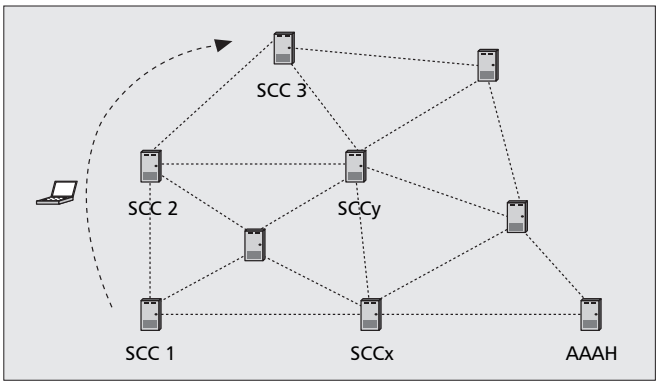

Figure 9. *P2P organization of SCCs [18].*

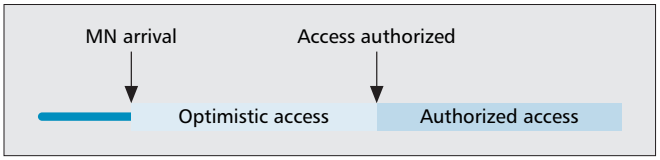

Figure 10. *Light and strong authentication in optimistic access scheme [19].*

and context transfer. Finally, the work presented in [27] shows how seamless handoffs can be realized in UMTS-WLAN integrated networks.

Discussion and Qualitative Analysis

Table 1 gives a comparison of the analyzed schemes based on selected criteria. In the rest of this section these criteria are explained and every scheme is compared to each other based on them. Using this approach, a clear view of the advantages and disadvantages of each scheme is provided. At the end of the section a discussion of the findings is furnished.

OSI Layer

This criterion shows in which OSI layer is the solution to the fast secure handoff problem implemented. In this article we only consider methods operating at either the network or application layer or both of them. When a protocol operates at the network layer, then it offers secure access to a different network even if the new network uses different link layer technology from the old one; this way the interconnection between heterogeneous networks is achieved. To put it in another way, it is possible to offer fast secure handoffs not only to multimedia services but to other applications as well. When a protocol operates at the application layer, then it is targeted at one application each time (in our case multimedia services offered by SIP) and this makes it possible to adapt better to the application's needs. Some schemes operate to both layers offering a complete solution to fast secure handoffs.

OIRPMSA combines authentication at network layer with authentication in application layer in order to provide optimized Mobile IP and SIP registration during handoff. *MPA* operates at the network layer and according to the authors it can be utilized in conjunction with any link layer and mobility management protocol. The *Shadow Registration* concept can be used to either layer, while nothing prevents it to operate to both layers in the case of multimedia services; moreover, two examples are provided, one for Mobile IP and the other for SIP registration. *AAA context transfer* operates either at network layer or at application layer or at both layers as well, and the testbed that the authors demonstrate uses Cellular IP and SIP protocols. Similarly, the *P2P context transfer* solution can be applied to any layer where fast re-authentication is required. The *Optimistic access* scheme is presented as an 802.11 technology solution. However the authors argue that it is also applicable to other technologies and it seems that it can be used to other OSI layers as well since it is rather a fast re-authentication method than a complete secure handoff scheme targeted specifically at one (specific) layer.

Security

In the security group some security related criteria are examined. The first one looks into whether each method uses *public or secret key* protocols to perform the necessary authentications. The next is *mutual authentication* which examines whether the authentication between the mobile terminal and the new access point is mutual or not. The *privacy* criterion checks if the actual identity of the mobile terminal is revealed to the new domain or not. The next criterion is about whether the new domain is able to claim the *non-repudiation* of the mobile terminal's actions. The last security related criterion is about whether it is assumed that there are *pre-established trust relationships* between the home and the visiting domain or not.

OIRPMSA, *MPA*, *Shadow Registration* and *P2P context transfer* do not dictate a special protocol neither the type of cryptography to be used, e.g. symmetric or asymmetric. By contrast, *AAA context transfer*, being based on specific technologies, uses the EAP-TLS protocol which is a public key protocol. *Optimistic access*, on the other hand, is based on shared key cryptography and keyed one way hash functions instead of public key signatures.

OIRPMSA, *Shadow Registration*, *P2P context transfer* and *Optimistic access* schemes do not support mutual authentication between the mobile terminal and the new access point. The *MPA* scheme, although it does not define an authentication protocol, it mandates that the chosen one should provide mutual authentication. As *AAA context transfer* utilizes EAP-TLS as the authentication protocol, it is straightforward that it can support mutual authentication.

The only protocol which supports privacy is *Optimistic access*. This protocol does not require any mobile terminal or user identity to be included into the exchanged messages. However, it does not specify what data are inserted into the credentials created by the access points and this is a possible breach of privacy.

The only scheme which offers non-repudiation services is *AAA context transfer*. This is based on the authentication protocol used which is EAP-TLS.

	OSI layer	Security					Efficiency			
		Public vs. secret key	Mutual authentication	Privacy	Non-repudiation	Assumes trust between domains	Roundtrips		Credential creation	Performance improvement (1%)[1]
Scheme identifier							During handoff	Total		
OIRPMSA	3, 7	not defined				full	3	3	on-the-fly	18.2%-33.3%
MPA	3	not defined	√√			no	1	6	on-the-fly	85%-99.65%[2]
Shadow Registration	3 or 7 or both	not defined				full	1	3	on-the-fly	—
AAA context transfer	3 or 7 or both	public key	√√	√	√	full	1	2	pre-computed	78.5%
P2P context transfer	any[5]	not defined				full	min 1, max 2	min 1, max 3	pre-computed	—
Optimistic access	any[7]	secret key	√	√		full	2	3	on-the-fly	—

Table 1. *Continued on next page…*

All the protocols except *MPA* assume that there exist pre-established trust relationships between the visiting and the home domain. The authors of the MPA scheme argue that their protocol works across different administrative domains based on trust relationships between the mobile terminal and each domain.

Efficiency

This group refers to criteria that examine the analysed schemes in terms of efficiency. The first two are about *roundtrips* occurring *during the handoff* process and the *total* number of handoffs required by the scheme. The next criterion shows whether the *credentials creation* is performed on-the-fly, e.g. when the credentials are requested from the authentication server (this does not imply that the request is done during the handoff process), or are being pre-computed before the actual request. Also, there is a *performance improvement* criterion which shows the percentage of performance improvement achieved by each scheme. Nevertheless, the methods discussed hereunder cannot be compared based on this criterion because every scheme concentrates on different specific network configuration which attempts to improve.

OIRPMSA performs 3 roundtrips during the handoff process, which is also equal to the total number of roundtrips performed by this scheme. *MPA* needs a total of 6 roundtrips, while only the last of them is executed during the handoff. When *Shadow Registration* is exploited, 3 roundtrips are performed in total, 1 of which is during handoff. The *AAA context transfer* has the minimum total number of roundtrips, requiring only 2, while 1 is entailed during the handoff process. In the case of *P2P context transfer*, the number of roundtrips during handoff is 1 when the previous SCC is known and 2 when is not. The total number of roundtrips is 1 and 3 respectively; the latter applies because the new SCC must inform the home AAA server that is the current SCC. *Optimistic access* needs 3 roundtrips in total, 2 of which during handoff, in order to complete its aim.

In *OIRPMSA*, *MPA*, *Shadow Registration* and *Optimistic access* schemes the creation of the credentials is done on-the-fly, whenever there is such a request from the AAA server. In *AAA context transfer* and *P2P context transfer* the credentials which are essential for the mobile terminal's authentication are pre-computed and can be communicated to foreign AAA servers before they are requested.

The authors of *OIRPMSA* provide a theoretical performance analysis on a system composed of Mobile IP, Diameter and Diameter SIP Application [20] protocols. This analysis showed an expected performance improvement between 18.2 percent and 33.3 percent. In the case of *MPA*, the results of a specific testbed are provided, which employs the following technologies: 802.11 as link layer technology, Protocol for Carrying Authentication for Network Access (PANA) protocol [29] for network access authentication, DHCP as the configuration protocol, SIP Mobility (SIP-M) [30] as the mobility management protocol, RTP/UDP [31] for carrying voice traffic and RAT (Robust Audio Tool) [32] as the media agent. This scheme does not only improve the delay imposed by

	Handoff types supported			Changes required	Standards used	Battery consumption	Scalability	4G ready
	Intra- or inter-domain	Pro-active/ reactive	Fast/smooth/ seamless					
Scheme identifier								
OIRPMSA	both	Reactive	not defined	Diameter Mobile IP and SIP applications, co-location of agents	Mobile IP, Diameter, Diameter SIP application	Depends on the implementation	low	√
MPA	both	Pro-active	seamless	Requires network elements: AA, CA, AR[3]	—	Depends on the implementation	medium	√
Shadow Registration	both	Pro-active	not defined	AAA protocol, SIP	Mobile IP, SIP	Depends on the implementation	low for SR and FHR, medium for RSR[4]	√
AAA context transfer	both	Re-active	seamless	Cellular IP	Hierarchical Mobile IP, Cellular IP, SIP	high	high	√
P2P context transfer	both	hybrid[6]	not defined	AAA protocol	—	Depends on the implementation	high	√
Optimistic access	both	Re-active	not defined	2nd layer protocol	—	low	high	

[1] The findings of this column cannot be used to compare the schemes because every scheme optimizes a differently configured network system.

[2] This scheme improves not only AAA related operations but other network parameters as well, like IP address assignment.

[3] AA: Authentication Agent, CA: Configuration Agent, AR: Access Router

[4] SR: Shadow Registration, RSR: Region-based Shadow Registration, FHR: Frequent Handoff Region

[5] It could be used to any layer where authentication is required.

[6] The first time the handoff is considered reactive but subsequent handoffs are considered proactive.

[7] The authors mainly consider 802.11 link layer technology but argue that the same ideas could be applied to other cases.

Table 1. *Secure handoff optimization schemes comparison.*

AAA operations but other delays as well, such as these imposed from the configuration protocol which tend to be higher. This results to an improvement in the order of 85 percent to 99.65 percent for MPA. For *Shadow Registration*, while a theoretical analysis is given, the performance improvement is heavily related to the distance between the mobile terminal and the home domain, and thus is difficult to be estimated. When the home domain is very close to the visiting domain the improvement is near zero and increases as the distance between the domains is increasing; so a representative numerical value could not be given. A testbed has been implemented in the case of *AAA context transfer* using Cellular IP, SIP and EAP-TLS protocols, resulting in a performance improvement of 78.5 percent in the case of multimedia service re-establishment, including both Cellular IP and SIP re-registrations, for an inter-domain handoff scenario. For *P2P context transfer* the analogy between performance improvement and domains distance applies as well; although a specific theoretical example shows an improvement of 35 percent this value cannot be used as a general improvement indicator. *Optimistic access* is designed for link layer handoff optimization, so no specific value can be given here.

Handoff Types Supported

This group of criteria refers to what types of handoffs each scheme is able to support. The first criterion examines whether *intra-domain or/and inter-domain* handoffs are afforded. Inter-domain handoffs tend to be most significant because this handoff type is the most expensive one in terms of delay. Next, it is examined whether the handoff each scheme supports is *proactive or reactive*; when a handoff is proactive, the

operation of the scheme starts before the handoff is actually needed and signaling is exchanged with the new access point before the mobile terminal connects to it; when a handoff is reactive, the scheme is initiated when the handoff is taking place and signaling with the new access point is done when the mobile terminal connects to it. In RFC 3753 [33], where mobility related terminology is listed, the following definitions are given for *fast/smooth/seamless* handover types:

- Fast handover. A handover that aims primarily to minimize handover latency, with no explicit interest in packet loss.
- Smooth handover. A handover that aims primarily to minimize packet loss, with no explicit concern for additional delays in packet forwarding.
- Seamless handover. A handover in which there is no change in service capability, security, or quality. In practice, some degradation in service is to be expected. The definition of a seamless handover in the practical case should be that other protocols, applications, or end users do not detect any change in service capability, security or quality, which would have a bearing on their (normal) operation. As a consequence, what would be a seamless handover for one less demanding application might not be equally seamless for another more demanding application.

From the above definitions it seems that the most appropriate type of handoff for secure multimedia delivery is the seamless one. Smooth handoffs are more appropriate for file transfers, while fast handovers could be suitable for multimedia delivery with no security restrictions.

The analysis showed that all methods are able to support both types when the distinction is made between intra-domain and inter-domain handoffs. While *Optimistic access* scheme does not explicitly deals with domains, these two types of handoff can be supported with careful selection of the security credentials and trust parameters.

OIRPMSA is a reactive scheme, while *MPA* and *Shadow Registration* are considered proactive because the new access point has received signalling prior to the handoff initiation phase. *AAA context transfer* and *Optimistic access* methods support reactive handoffs. Finally, *P2P context transfer* is using a mix of the two types and is considered a hybrid solution. The first time the user makes a handoff, this is a reactive one, while the subsequent are proactive; when the user cannot find a security context in the path between the mobile terminal and the home domain then this is also considered a reactive handoff.

MPA and *AAA context transfer* are designed with seamless handoffs in mind. The rest of the methods do not define the support of any special handoff type between fast/smooth/seamless types.

Changes Required

This section describes the changes required for the deployment of each scheme. It is stressed that the number, the nature and (most important) the cost of modifications to existing systems required by a scheme for its deployment plays a crucial role in its adoption and the transition to it from existing solutions.

OIRPMSA uses Diameter as AAA protocol and introduces the use of reserved flags of Diameter's Mobile IP and SIP extensions. Another modification required by this method is the co-location of Mobile IP's Foreign Agent with SIP Proxy into a FA/SIP Proxy and Mobile IP's Home Agent with SIP Registrar into a HA/SIP Registrar. MPA requires the introduction of three functional elements to each network:

- An Authentication Agent (AA) which is responsible for pre-authentication
- A Configuration Agent (CA) which is used for secure delivery of IP address and other parameters to the mobile terminal (first part of pre-configuration)
- An Access Router (AR) which executes the rest of the pre-configuration phase.

Shadow Registration necessitates the modification of existing messages of the AAA protocol in use; also, when SIP is used, a new SIP message introduced, namely the ANSWER message. AAA context transfer scheme requires modifications or adaptations to the Cellular IP protocol which can be summarized as follows: introduction of three new types of messages, modification of one existing message and a need for a context cache at each gateway. The changes mandated by *P2P context transfer* scheme relate with the AAA protocol; some AAA servers should also act as SCCs (Security Context Controllers) and these nodes should be capable of forming a secure Peer-to-Peer network. *Optimistic access* scheme in its current form requires the modification of the link layer protocol in use; if it is to be used for network or application layer re-authentication then the respective protocols should be altered.

Standards Used

This section summarizes the existing standards used by each scheme. The utilization of existing standards plays an important role in the commercial deployment of the proposed systems because it solves most problems causing incompatibilities of implementations between different vendors.

OIRPMSA operates based on Mobile IP, Diameter and an extension of it, namely Diameter SIP application. *MPA*, *P2P context transfer* and *Optimistic access* are more generic approaches and are not based on specific standards. The standards used by *Shadow Registration* method are Mobile IP and SIP. *AAA context transfer* uses Hierarchical Mobile IP, Cellular IP and SIP in its deployment.

Battery Consumption

This criterion is concerned with the level of power consumption which is very important in wireless networks where the mobile terminals work on batteries and therefore have limited power reserves. The criteria with which battery consumption is related are mainly the number of roundtrips and the type of cryptography used; asymmetric cryptography tends to be very expensive in terms of power consumption for mobile devices in contrast to symmetric cryptography.

OIRPMSA, *MPA*, *Shadow Registration* and *P2P context transfer* schemes do not clarify what type of cryptography will be used for the re-authentication of the mobile terminal, thus the cost in battery consumption is highly depended on the chosen implementation. *AAA context transfer* is considered a high demanding scheme because it uses EAP-TLS as authentication protocol, whereas *Optimistic access* with the use of only symmetric cryptography and one way hash functions is regarded a rather low consumption solution.

Scalability

The level of scalability shows how well is the scheme adapted when the number of networks, network elements and subscribers is increasing. It shows how dynamic is the system considering such changes and its possibility to be deployed in a large scale.

OIRPMSA is perceived to be a low scalability solution mainly because it requires the co-location of Mobile IP and

SIP servers. More specifically, it imposes the co-location of Mobile IP's Foreign Agent (FA) with a SIP Proxy into a single FA/SIP Proxy and the co-location of Mobile IP's Home Agent (HA) with a SIP Registrar into a single HA/SIP Registrar node. This could end up to a performance penalty when the number of serving mobile terminals is high. MPA is a scheme with moderate scalability; this arises from the resource demanding nature of this method. When a mobile terminal is about to handoff the procedures of pre-authentication and pre-configuration engage a considerable portion of resources, especially when the request rate is high, which may never be used if the handoff will not take place. In the *Shadow Registration* case a distinction is made between the three analyzed variations, namely Shadow Registration (SR), Region-based Shadow Registration (RSR) and Frequent Handoff Region (FHR). SR is considered a low scalability solution because of the excessive signaling during the registration phase, problem which is partially solved by RSR which in turn is considered a medium scalability method. Finally, FHR is a low scalability scheme because it needs the manual collection of information each time new access points are deployed. *AAA context transfer*, *P2P context transfer* and *Optimistic access* manage to keep the level of signaling rather low, resulting in high to moderate scalability. This must be proved thought considering either a real deployment or a wide scale simulation.

4G Ready

According to the 4G vision, future wireless heterogeneous networks will converge into an all-IP platform. This criterion designates whether the schemes in question are ready to support 4G networks.

From the analysis already provided earlier it is obvious that all schemes except *Optimistic access* are ready to support the forth generation of wireless networks. Optimistic access is not included in the 4G capable schemes list because it operates in the link layer; if, however, the ideas presented by its authors were adapted to higher layers then it could be considered a 4G capable solution.

Conclusions and Future Work

It is envisioned that future wireless networks will converge to an all-IP platform offering more bandwidth consuming services at higher speeds. In such an environment the security of multimedia services, being a demanding class of applications, without perceived degradation by the user is a very challenging issue. The realization of this objective includes the cooperation of mobility management schemes with AAA protocols for the secure and uninterrupted multimedia services provision.

In this article an overview of the current most representative secure handoff optimization schemes trying to achieve the aforementioned goals was given. Each scheme was briefly presented and some comments were provided where this was considered purposeful. Finally, a comparison of the schemes was conducted and the criteria of this comparison were further analyzed and explained.

The purpose of this work is to mark each scheme's advantages and disadvantages utilizing not only qualitative but quantitative criteria where this was possible. This way, it can be used as a basis for the evaluation of new proposed schemes and as a reference for the properties a secure handoff scheme should posses. As a statement of direction, we are currently working on expanding this work by proposing a new optimized secure handoff scheme, which exploits the advantages of the presented methods while at the same time minimizes the drawbacks pointed out throughout this article.

References

[1] C. Rigney *et al.*, "Remote Authentication Dial In User Service (RADIUS)," RFC 2865, June 2000.

[2] P. Calhoun *et al.*, "Diameter Base Protocol," RFC 3588, Sept. 2003.

[3] C. Perkins, "IP Mobility Support for IPv4," RFC 3344, Aug. 2002.

[4] A. Campbell *et al.*, "Cellular IP," IETF Internet Draft, Oct. 1999.

[5] P. Xu *et al.*, "Optimized Integrated Registration Procedure of Mobile IP and SIP with AAA Operations," *20th Int'l. Conf. Advanced Info. Networking and Applications*, vol. 1 (AINA'06), 2006, pp. 926–31.

[6] J. Rosenberg *et al.*, "SIP: Session Initiation Protocol," RFC 3261, June 2002.

[7] A. Dutta *et al.*, "MPA Assisted Optimized Proactive Handoff Scheme," *Proc. 2nd Annual Int'l. Conf. Mobile and Ubiquitous Systems: Networking and Services 2005 (MobiQuitous 2005)*, 17–21 July 2005, pp. 155–65.

[8] A. Dutta *et al.*, "A Framework of Media-Independent Pre-Authentication (MPA)," IETF Internet Draft, draft-ohba-mobopts-mpa-framework-03, work in progress, Oct. 2006.

[9] A. Dutta *et al.*, "Secured Seamless Convergence Across Heterogeneous Access Networks," White Paper, Columbia University, Apr. 2006.

[10] IEEE P802.21/D00.05: Draft IEEE Standard for LAN/MAN: Media Independent Handover Services, Jan. 2006.

[11] T. Kwon, M. Gerla, and S. Das, "Mobility Management for VoIP: Mobile IP vs. SIP," *IEEE Wireless Commun. Mag.*, vol. 9, no. 5, Oct. 2002, pp. 66–75.

[12] S.-B. Han *et al.*, "Efficient Mobility Management for Multimedia Service in Wireless IP Networks," *Proc. 4th Annual ACIS Int'l. Conf. Computer and Information Science (ICIS'05)*, 2005, pp. 447–52.

[13] S. Pack and Y. Choi, "Fast Inter-AP Handoff using Predictive-Authentication Scheme in a Public Wireless LAN," *Networks 2002 (Joint ICN 2002 and ICWLHN 2002)*, Aug. 2002.

[14] M. Georgiades *et al.*, "AAA Context Transfer for Seamless and Secure Multimedia Services over All-IP Infrastructures," *5th European Wireless Conf.*, Spain, Feb. 2004.

[15] J. Kempf, "Problem Description: Reasons For Performing Context Transfers Between Nodes in an IP Access Network," RFC 3374, Sept. 2002.

[16] H. Soliman *et al.*, "Hierarchical Mobile IPv6 mobility management (HMIPv6)," IETF Internet Draft, draft-ietf-mobileip-hmipv6-08.txt, June 2003.

[17] B. Aboba and D. Simon, "PPP EAP TLS Authentication Protocol," RFC 2716, Oct. 1999.

[18] T. Braun and K. Hahnsang, "Efficient Authentication and Authorization of Mobile Users Based on Peer-to-Peer Network Mechanisms," *Proc. 38th Annual Hawaii Int'l. Conf. System Sciences (HICSS '05)*, Jan. 2005.

[19] T. Aura and M. Roe, "Reducing Reauthentication Delay in Wireless Networks," *1st Int'l. Conf. Security and Privacy for Emerging Areas in Commun. Networks (SecureComm 2005)*, Sept. 2005.

[20] H. Kim, K. G Shin, and W. Dabbous, "Improving Cross-domain Authentication over Wireless Local Area Networks," *1st Int'l. Conf. Security and Privacy for Emerging Areas in Commun. Networks, (SecureComm 2005)*, Sept. 2005, pp. 127–38.

[21] M. S. Bargh *et al.*, "Fast Authentication Methods for Handovers Between IEEE 802.11 Wireless LANs," *Proc.2nd ACM Int'l. Wksp. Wireless Mobile Applications and Services on WLAN Hotspots (WMASH '04)*, ACM Press, 2004, pp. 51–60.

[22] H. Kim, W. Ameur, and H. Afifi, "Toward Efficient Mobile Authentication in Wireless Inter-domain," *Proc. Wksp. Applications and Services in Wireless Networks (ASWN), IEEE*, July 2003, pp. 47–56.

[23] A. Mishra, M. Shin, and W. A. Arbaugh, "Pro-active Key Distribution using Neighbor Graphs," *IEEE Wireless Commun.*

Mag., Feb. 2004.
[24] P. Engelstad, T. Haslestad, and R. Paint, "Authenticated Access for IPv6 Supported Mobility," *Proc. IEEE Symp.Computers and Commun. (ISCC'2003)*, Turkey, vol. 1, 2003, pp. 569–75.
[25] X. Fu *et al.*, "Secure, QoS-Enabled Mobility Support for IP-based Networks," *Proc. IP Based Cellular Network Conference (IPCN)*, Paris, France, Dec. 2003.
[26] T. Chen *et al.*, "A Performance Study of Session State Re-Establishment Schemes in IP-based Micro-Mobility Scenarios," *Proc. IEEE Computer Society's 12th Annual Int'l. Symp. Modeling, Analysis, and Simulation of Computer and Telecommun. Systems (MASCOTS 2004)*, Oct. 2004, pp. 159–66.
[27] H. Kwon *et al.*, "Consideration of UMTS-WLAN Seamless Handover," *Proc. 7th IEEE Int'l. Symp. Multimedia (ISM '05)*, 2005, pp. 649–56.
[28] M. Garcia-Martin *et al.*, "Diameter Session Initiation Protocol (SIP) Application."
[29] D. Forsberg *et al.*, "Protocol for Carrying Authentication for Network Access (PANA)," IETF Internet Draft, draft-ietf-pana-pana-12, work in progress, Aug. 2006.
[30] H. Schulzrinne and E. Wedlund, "Application-Layer Mobility using SIP," *SIGMOBILE Mob. Comp. Commun. Rev.*, vol. 4, no. 3, ACM Press, 2000, pp. 47–57.
[31] H. Schulzrinne *et al.*, "RTP: A Transport Protocol for Real-Time Applications," RFC 3550, July 2003.
[32] http://www-mice.cs.ucl.ac.uk/multimedia/software/rat/
[33] J. Manner and M. Kojo, "Mobility Related Terminology," RFC 3753, June 2004.

BIOGRAPHIES

GIORGOS KAROPOULOS (gkar@aegean.gr) is currently a Ph.D. candidate at the University of the Aegean, department of Information and Communication Systems Engineering. He holds a diploma in Information and Communication Systems Engineering and a M.Sc. in Information and Communication Systems Security both from the University of the Aegean. His current research focus is in mobile multimedia security in all-IP heterogeneous networks.

GEORGIOS KAMBOURAKIS (gkamb@aegean.gr) was born in Samos, Greece, in 1970. He received the Diploma in Applied Informatics from the Athens University of Economics and Business and the Ph.D. in Information and Communication Systems Engineering from the department of Information and Communications Systems Engineering of the University of Aegean (UoA). He also holds a M.Ed. degree from the Hellenic Open University. Currently he is a Lecturer in the Department of Information and Communication Systems Engineering of the UoA, Greece. His research interests are in the fields of Mobile and ad-hoc networks security, VoIP security, security protocols and PKI and he has more than 35 publications in the above areas.

STEFANOS GRITZALIS (sgritz@aegean.gr) holds a B.Sc. in Physics, an M.Sc. in Electronic Automation, and a Ph.D. in Informatics all from the University of Athens, Greece. Currently he is an Associate Professor, the Head of the Department of Information and Communication Systems Engineering, University of the Aegean, Greece and the Director of the Info-Sec-Lab. His published scientific work includes several books and more than 150 journal and international conference papers on Information and Communication Security topics.