Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Hasan, Burkhard Stiller

Non-repudiation of Service Consumption

TIK-Report
Nr. 209, December 2004

Institut für Technische Informatik und Kommunikationsnetze
Computer Engineering and Networks Laboratory

Hasan, Burkhard Stiller
Non-repudiation of Service Consumption
December 2004
Version 1
TIK-Report Nr. 209

Computer Engineering and Networks Laboratory,
Swiss Federal Institute of Technology (ETH) Zurich

Institut für Technische Informatik und Kommunikationsnetze,
Eidgenössische Technische Hochschule Zürich

Gloriastrasse 35, ETH-Zentrum, CH-8092 Zürich, Switzerland

Non-repudiation of Service Consumption

Hasan[1], Burkhard Stiller[2,1]

TIK-Report No: 209

[1] Computer Engineering and Networks Laboratory TIK, ETH Zürich, Switzerland
[2] Computer Science Department IFI, University of Zürich, Switzerland
E-Mail: [hasan|stiller]@tik.ee.ethz.ch, stiller@ifi.unizh.ch

Abstract

Today's Internet technology is able to support different Quality-of-Service (QoS) classes to meet different application and user requirements. Combined with the support of user mobility, service providers can offer differentiated services not only to their own customers, but also to roaming users. This service offer is accompanied normally by more complex pricing schemes which require a complex accounting of the real service consumption. It is important that the commercial provisioning of Internet services needs to meet security requirements of service providers as well as service users. Besides the access control to services, the dedicated service consumption must be provable to justify billing and to protect users and providers against other malicious parties.

This paper develops an architecture termed NorCIS (Non-repudiation of the Consumption of Internet Services) and its detailed protocol interactions which allow for the generation and transfer of non-repudiation evidences of service consumptions in a mobile Internet Protocol (IP)-based environment. The respective evidence structure is proposed, which supports a variety of accounting schemes, and which includes information to be used for protection against various attacks. In addition, NorCIS proposes the use of virtual identifiers within evidences to support the privacy of users' identities.

Keywords: *Evidence, Mobile Internet, Non-repudiation, Privacy, Security, Service Consumption*

1 Introduction

In the near future many differentiated IP services will be made available to meet different application and user Quality-of-Service (QoS) requirements. Those services will be offered by numerous co-operating service providers to allow for a seamless access by mobile and roaming users. A service provider is a provider who offers connectivity, application, content, or any combination of them to the end users. The terms provider and service provider are used interchangeably.

Figure 1 depicts the scenario considered for a mobile environment where users are able to consume IP-based services not only from the home, but also from foreign domains. User U has a contract with his Home (network) Provider HP. This contract allows user U via his mobile terminal to reach network access from the home domain operated by HP. To allow for user U also gaining network access from different foreign domains, roaming agreements are established between HP and Foreign (network) Providers (FPs), who operate those domains independently. The deployment of Authentication, Authorization, and Accounting (AAA) [2] and a Mobile IP [9], [13] infrastructure is assumed to enable user mobility.

In such a mobile environment accounting data are important to justify charges for the consumption of services, particularly during the time the user is visiting a foreign domain. Users are expected to rely on accounting information that a provider collects. However, a user can deny having consumed a service, if there is no proof of this usage. In practice, this dispute can be solved by having written terms and conditions, "forcing" the user to accept the correctness of all providers' accounting information, mainly if the user cannot prove the opposite. This solution places the user in a disadvantage compared to the provider. Therefore, a fair solution for users and providers has to be achieved by applying non-repudiation (NR) mechanisms to generate evidences during the service consumption. This paper proposes that the provider is obliged to prove the correctness of his bill, if there is a dispute. Therefore, the approach termed NorCIS, Non-repudiation of the Consumption of Internet Services, has been developed.

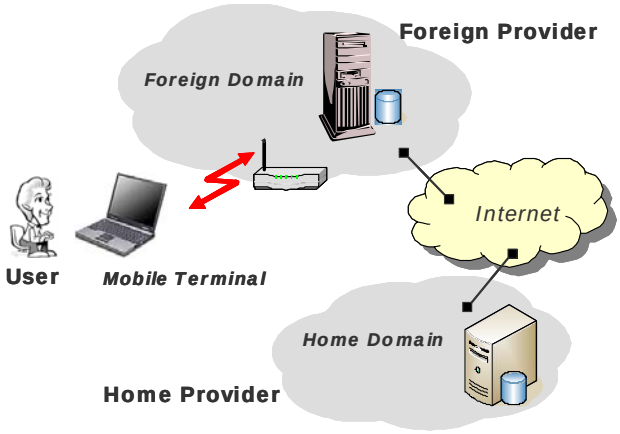

Figure 1: Roaming User Scenario.

The definition of repudiation is provided by the Internet Engineering Task Force (IETF) in [14] as a "denial by a system entity that was involved in an association (especially an association that transfers information) of having participated in the relationship." And a non-repudiation service is defined as "a security service that provides protection against false denial of involvement in a communication." The IETF makes a further remark that "A non-repudiation service does not and cannot prevent an entity from repudiating a communication. Instead, the service provides irrefutable evidence that can be stored and later presented to a third party to resolve disputes that arise if and when a communication is repudiated by one of the entities involved." Obviously, these IETF definitions are related or restricted to communication events. The approach presented in this paper, however, uses the term non-repudiation in a broader sense.

A special non-repudiation service is proposed, *i.e.*, a non-repudiation of service consumption. The objective of non-

repudiation aimed at is to protect against users' false denial of having consumed specified services. Hence, statements on service consumption need to be signed by the user, and securely stored by the provider. Note that the objective is not to protect against providers' false denial of having delivered specified services.

The provisioning of non-repudiation services is divided into different phases: (1) evidence generation, (2) evidence transfer, (3) evidence verification and storage, and (4) dispute resolution [15]. In all of these phases, Trusted Third Parties (TTP) can assist participants in various ways as identified by different roles a TTP can play, such as a Certification Authority, a Notary, a Delivery Authority, a Time-stamping Authority, and an Adjudicator [15]. The approach presented here deals with phases (1) to (3) and avoids the involvement of a TTP; at least it reduces a TTP involvement to an offline mode, whenever possible.

Services considered include multi-media and QoS-enabled services, which are offered using a variety of complex pricing schemes. Not only the duration of service consumption needs to be accounted for, but also, *e.g.*, the time of the day, traffic volume, and its QoS class need to be recorded. Moreover, a provider normally applies a different tariff to his own customers than for visiting mobile users. Additionally, a user may not want his real identity to be known to FPs, yet FPs must be able to account for service consumptions. This requires a complex description of service consumption and related evidence structure.

Due to these considerations the following list of important issues have to be solved for a fully decentralized and distributed system of co-operating and competing providers:

- Which instance needs to keep evidences of a service consumption: foreign provider, home provider, or both?
- Which instance should generate the statement on service consumption?
- What must a non-repudiation evidence comprise of in order to be able to prove the consumption of differentiated services with different accounting schemes, as well as to protect against various attacks? Potential attacks are *e.g.*, users' denial of having consumed a particular service from a particular provider at a particular time for the particular amount (duration, volume, etc.), and provider generated fake evidences to prove fictitious service consumption.
- How to deal with the change of the Care-of Address (CoA) in mobile communications between non-repudiation entities in the mobile terminal and the network?

The remainder of the paper is organized as follows. Section 2 compares different related work with respect to expected properties in the non-repudiation of service consumption. While Section 3 presents NorCIS describing all roles and functions of participating parties in the non-repudiation process, Section 4 develops the implementation architecture and outlines key considerations regarding its support of mobility and security, as well as its performance and scalability. Possible optimizations are discussed in Section 5. Finally, concluding remarks are offered in Section 6.

2 Related Work

The International Organization for Standardization (ISO) non-repudiation model deals with events of creating, sending, receiving, and recognizing the content of a message [8]. Several non-repudiation services are defined with each service related to the specific event or a meaningful combination of those events. *E.g.*, the non-repudiation of origin (NRO) is a non-repudiation service which is intended to protect against an originator's false denial of having created the content of a message and of having sent the message. NorCIS, however, focuses on specific messages containing statements on service consumption and on a non-repudiation protocol capable of inter-domain interactions within a mobile environment to transfer those evidences generated.

Current non-repudiation protocols reduce the involvement of TTPs to deal with keys only rather than with the content of transferred messages. Research has been performed in achieving specific requirements on the property of a non-repudiation protocol, *e.g.*, fairness. Since its introduction in non-repudiation protocols, the definition of fairness has evolved into different flavors: weak, strong, true, and probabilistic fairness [10], [12]. A non-repudiation protocol provides strong fairness if and only if at the end of a protocol execution either A received the non-repudiation of receipt evidence for the message M, and B received the corresponding message M as well as the non-repudiation of origin evidence for this message, or none of them received any valuable information [10]. A fair non-repudiation protocol using an on-line TTP is proposed in [18] and a number of protocols have been developed to improve fairness and security with respect to exchanges of electronic goods [1], [6], [16], [17]. NorCIS sets fairness into relation with a business risk when choosing between proving before or proving after a service usage.

Proving service requests and access granting determines a more general case, whereas proving service provisioning and service usage is usually application-dependent and sometimes hard to decide without human intervention. [7] shows how non-repudiation methods can be used to prove service requests and access granting for a lease service using public-key cryptography.

[11] proposes a protocol to provide mutual entity authentication, identity privacy, and a limited version of non-repudiation service to secure mobile communications. The protocol is based on the use of conventional secret-key techniques in combination with modern public-key techniques. Upon subscription a mobile user holds the public encryption key of the HP and a secret-key k shared between the mobile user and the HP. HP's public-key is used to securely transmit a user's identity and credentials to the HP for authentication. Some shortcomings exist. Since evidences are generated based on k, they cannot be generated by the FP, but they can be generated by the HP. Hence, a user is not protected against his malicious HP who generates fake evidences. Furthermore, evidences are generated for service requests not for service consumption. Addressing privacy, a temporary user identity is assigned and sent securely to the user, but done by the FP, and hence, the real identity is known to the FP.

An additional scheme is proposed in [19] by using a combination of a digital signature and a one-way hash chain technique to provide non-repudiation of billing, when a mobile user roams into foreign networks. This scheme aims at improving the abovementioned non-repudiation mechanism by providing evidence to prove a service duration. Mobile users need to submit a digital signature when requesting a call and release chained hash values during the session so that the call and its duration are undeniable. The scheme proposed in [19] is, however, limited to time-based accounting schemes, whereas the proposed composition of evidence in the NorCIS approach allows for the support of many different accounting schemes. Moreover, the signature and verification keys to be used by the user and FPs respectively are generated by the HP. Therefore, a malicious HP is able to fake evidences. The following table summarizes and compares important related approaches.

Table 1: Comparison of Different Related Approaches.

Approach	Non-repudiation	Mobility	Privacy	Fairness	Security
ISO	communication event	not considered	not considered	weakly considered	addresses only non-repudiation protocols; some types of evidence contain timestamps, which are not fully defined, thus allows for a fake evidence attack.
E.g., Zhou-Gollmann, Asokan, et.al., Zhou-Deng-Bao	communication event	not considered	not considered	supported (main focus)	addresses only the non-repudiation protocol
Hasselmeyer, et.al.	lease request	not considered	by encrypting all communication channels; by using proxy services between the client and the server	not supported	attacks are not considered
Lin-Harn	call request	considered	temporary user identities are generated by the FP; the real identity of a mobile user is known to the FP	not considered	malicious HP can generate undetectable fake evidences
Zhou-Lam	call request and duration	considered	the real identity of a mobile user is known only to the HP (and the user)	not supported	malicious HP can generate undetectable fake evidences
NorCIS	consumption of differentiated services (QoS class, duration, volume)	considered	the real identity of a mobile user is known only to the HP (and the user)	supported for consumption of services with time-based accounting	fake evidences are detectable; users cannot falsely deny consumption by key revocation

3 The Design of the NorCIS Model: Roles and Functions

This section designs the model which identifies all parties involved in the non-repudiation process as well as their roles and functions (cf. Figure 2). A provider can take the role of an FP or an HP in an interaction with users and with the other providers. Hence, a provider must implement all functionality assigned to both the HP and the FP. Three roles participate in a non-repudiation process: the user, the FP, and the HP. A party is identified by the role it currently assumes.

3.1 Parties Keeping Evidences

A signed statement on service consumption is called an "evidence of service consumption" or henceforth, "evidence" in short. In principle, a party issuing an invoice must be able to prove service consumption. Generally, it is the HP which charges a user for his consumption of services either in a home or in a foreign domain. The HP is interested in having this evidence. However, when a user is in a foreign domain, FPs deliver services to the user. Normally, an FP does not charge the visiting user directly, instead, he will charge the HP of this user for all services consumed by him in FPs' domain. Therefore, an FP is interested in having this evidence. NorCIS proposes that FPs and HPs keep those evidences.

3.2 Generation and Verification of Evidences

Obviously, the evidence must be generated by the user who has consumed or is consuming the service. This evidence which is called the user evidence, proves that the user agrees to be responsible for paying the consumed service. If the HP of a user is responsible for services consumed by this user in a foreign domain, then the HP must generate an evidence based on the user evidence. This new evidence is called the HP evidence, which proves that the HP agrees to be responsible for paying to the FP for the services consumed by the user.

There exist two mechanisms to generate an evidence: using symmetric or asymmetric keys. In case of the use of symmetric keys, keys are not shared between the user and the provider to avoid fake evidences generated by the provider. In general, a non-repudiation service using symmetric keys requires the keys to be known only to a TTP and evidences to be generated by the TTP. However, this strict requirement does not apply to non-repudiation of service consumption. Instead, keys are shared between the user and a TTP. To verify evidences the provider has to contact the TTP.

To avoid a TTP's involvement, asymmetric keys, *i.e.*, a public-private key pair of the user, are applied. For security reasons this user's public-private key pair has to be generated by the user himself. The private key is used as a signature

key for evidence generation, whereas the public key is used as a verification key for evidence verification.

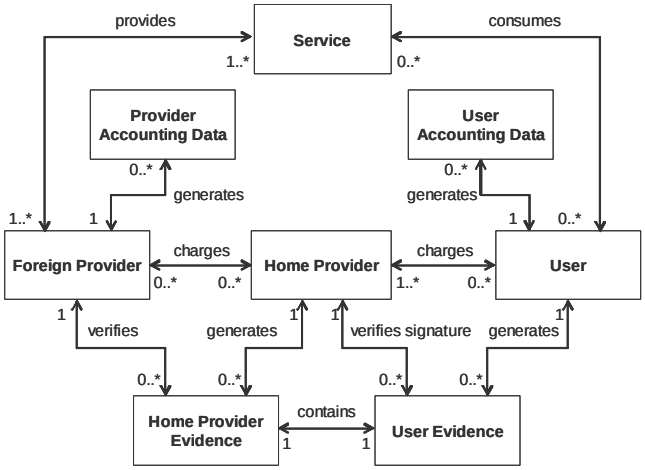

Figure 2: NorCIS Model.

It is in the interest of the provider to verify the correctness of evidences, *i.e.*, the correctness of statements as well as signatures. An HP must be able to verify the correctness of signatures of his users. If an FP does not have the verification key of the visiting user, the signature verification needs to involve the respective HP. Correctness of such statements can be verified only by providers who have access to accounting data. It is in the interest of the provider who delivers services to verify the correctness of this statement in the evidence. Hence, this provider must generate the accounting data. If the verification of an evidence fails, the respective service may need to be terminated.

3.3 Generation and Verification of Non-repudiation Statements

Two issues are of utmost importance with respect to the generation of non-repudiation statements: identification of the party who generates statements and the frequency of such statement generations. Non-repudiation statements can be generated or verified only by an entity which has access to the accounting data. If a statement on service consumption is not generated by the user who has consumed or is consuming the service, the correctness of this statement should be verified by the user before signing it. If this verification fails, the user must not sign such statements while facing the consequence of service termination. Three alternatives exist with respect to the generation of non-repudiation statements as summarized in Table 2.

Table 2: Different Alternatives of Statements Generation.

Statements Generated by	Description	Advantages
FP	The FP accounts for service consumption of all users in order to be able to do charging. Based on this information the service consumption statements are generated and sent to the respective users to be signed. The user accounts for service consumption to verify provider's statements.	Precise accounting at mobile terminal can be made optional if users can trust the accounting data from the provider. Besides, it is not always possible to require every mobile terminal to implement metering and accounting.
User	The user accounts for service consumption, generate statements, signs, and sends them either to the FP or the HP.	Simpler interactions, if statements are always correct. Non-repudiation traffic in the network can be reduced.
HP	Accounting data collected by the FP can be sent to the HP. This way, HP can also be the party who generates the non-repudiation statements. The statements are then sent to the user to be signed. The user accounts for service consumption to verify provider's statements.	This approach does not have advantages compared to the other approaches.

The frequency of statement generation is determined by the accounting scheme for the particular service as shown in Table 3. For time-based and volume-based accounting schemes this frequency is determined by a trade-off between communication overhead and business risk.

Table 3: Frequency of Statement Generation.

Accounting Scheme	Frequency of Statement Generation
Time-based	Every *t* unit of time
Volume-based	Every *n* unit of traffic volume consumed by the user
Event-based	Every *item* sent/received

3.4 Composition of an Evidence: Main Fields

To enable a practical use of evidences of service consumption, key data need to be coded into packet fields. Further information to meet different security requirements are described below in Section 4.3. Therefore, a user evidence of service consumption must contain the following fields of information:

- Service Provider Identifier (SPID), who has delivered or is delivering the specified service to the user.
- User Identifier (UID)
- Service Identifier (SvcID)
- Session Identifier (SessID)
- Consumption Interval (CI, start/end time of interval)
- Traffic Volume (TV)
- User's Digital Signature (UDS) over the hash of all the above information fields

An HP evidence must contain the following fields in addition to the abovementioned fields for a user evidence:

- Signature Verifier Identifier (SVID)
- Signature Verifier's Digital Signature (SVDS, the digital signature of the SVID) over the hash of all the abovementioned information fields.

Four fields are used to describe the consumed service: SvcID, SessID, CI, and TV. The SvcID uniquely identifies a service within an administrative domain, in which the service is being or has been consumed. The SessID is a combination of different attributes which characterize a session. For a transport service this may comprise of the DSCP (Differentiated Service Code Point), source and destination addresses, source and port numbers, and the protocol identifier. For a content service this field must uniquely identify the content in that administrative domain. The Consumption Interval defines the time interval during which the service was consumed (session duration). The TV field is required, if the accounting scheme is volume-based. In an event-based accounting, this field is used for the size of the item. In case of time-based accounting, this field is only indicative.

The CI field is useful for the synchronization between the user and the provider in accounting the service consumption. Synchronization of accounting is possible if the same time scale, *e.g.*, UTC (Coordinated Universal Time) is used and the machines' clocks are synchronized, *e.g.*, using NTP (Network Time Protocol). As clocks cannot be fully synchronized in distributed systems, users and providers have to agree on a small deviation.

To identify the participating parties in a service consumption, three fields are defined: SPID, UID, and SVID. The SPID is the globally unique identifier of the provider who is delivering or has delivered the service. This is the FP Identifier (FPID), if the service consumption happens within a foreign domain, otherwise, this is the HP Identifier (HPID). This identifier is required so that the evidence cannot be misused by other providers. The UID identifies the user who is consuming or has consumed the service. The SVID contains the HPID, because it is the HP who verifies the user's digital signature in the user evidence.

The user's digital signature is performed over the one-way hash of the concatenated SPID, UID, and all fields that describe the consumed service. The SVDS is the digital signature of the HP. It is performed over the one-way hash of the concatenation of all the abovementioned fields. Note that it is not sufficient to calculate the hash over the SVID concatenated with the UDS, because a collusion between the HP and the user can attack the FP by saying that the FP modifies *e.g.*, TV — increasing the value of TV —, after the HP evidence is generated.

4 The NorCIS Architecture

Based on the NorCIS model the implementation architecture has been developed, including interactions of its key components to allow for a non-repudiation of service consumption, Main considerations with respect to mobility, security, performance, and scalability are added.

4.1 Implementation Architecture & Interactions

To provide a non-repudiation service it is necessary to move evidences from one party to the other. Three main entities are involved in providing the non-repudiation service as depicted in Figure 3: A Non-repudiation Client (NR Client) in the user's mobile terminal, a Foreign Non-repudiation Server (Foreign NR Server) in the FP's network (foreign administrative domain), and a Home Non-repudiation Server (Home NR Server) in HP's network (home administrative domain).

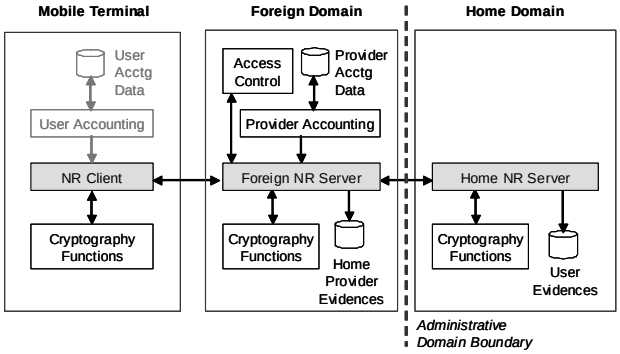

Figure 3: NorCIS Architecture.

Due to the fact that signing a message with a digital signature and verifying a signature belong to general security functions, they are delegated to the Cryptography Functions module. Accounting data which are needed to generate and to verify the non-repudiation statements on service consumption are accessed via the User and Provider Accounting module respectively. Failure in evidence verification may lead to a service termination. Hence, failure notifications need to be sent to the Access Control module. In addition, Table 4 describes all interfaces as shown in Figure 3.

Table 4: Interface Specifications.

Interface	Specification
User Accounting - NR Client	Transfer of session information and accounting data to the NR Client to generate or verify statements on service consumption
NR Client - Cryptograhpy Functions	Signing of non-repudiation statements on service consumption (evidence generation)
Provider Accounting - Foreign NR Server	Transfer of session information and accounting data to the Foreign NR Server to generate or verify statements on service consumption
Foreign NR Server - Cryptography Functions	Verifying the HP's digital signature; verifying user's digital signature, if the FP has the user's verification key
Home NR Server - Cryptography Functions	Verifying the user's digital signature; signing user evidences (generation of HP evidences)
NR Server - NR Evidences	Storing of evidences
NR Client - Foreign NR Server	Transfer of statements to the NR Client (if statements are not generated by the user); transfer of user evidences to the NR Server; transfer of notification in case of verification failure
Foreign NR Server - Home NR Server	Transfer of user evidences to the Home NR Server; transfer of HP evidences from Home NR Server to the Foreign NR Server; transfer of notification in case of verification failure
Foreign NR Server - Access Control	Transfer of notification in case of verification failure

The non-repudiation protocol is responsible for moving evidences from the NR Client to the NR Server. In Figure 4 it is assumed that the FP generates these statements on service consumption, and evidences are sent to the FP. Other

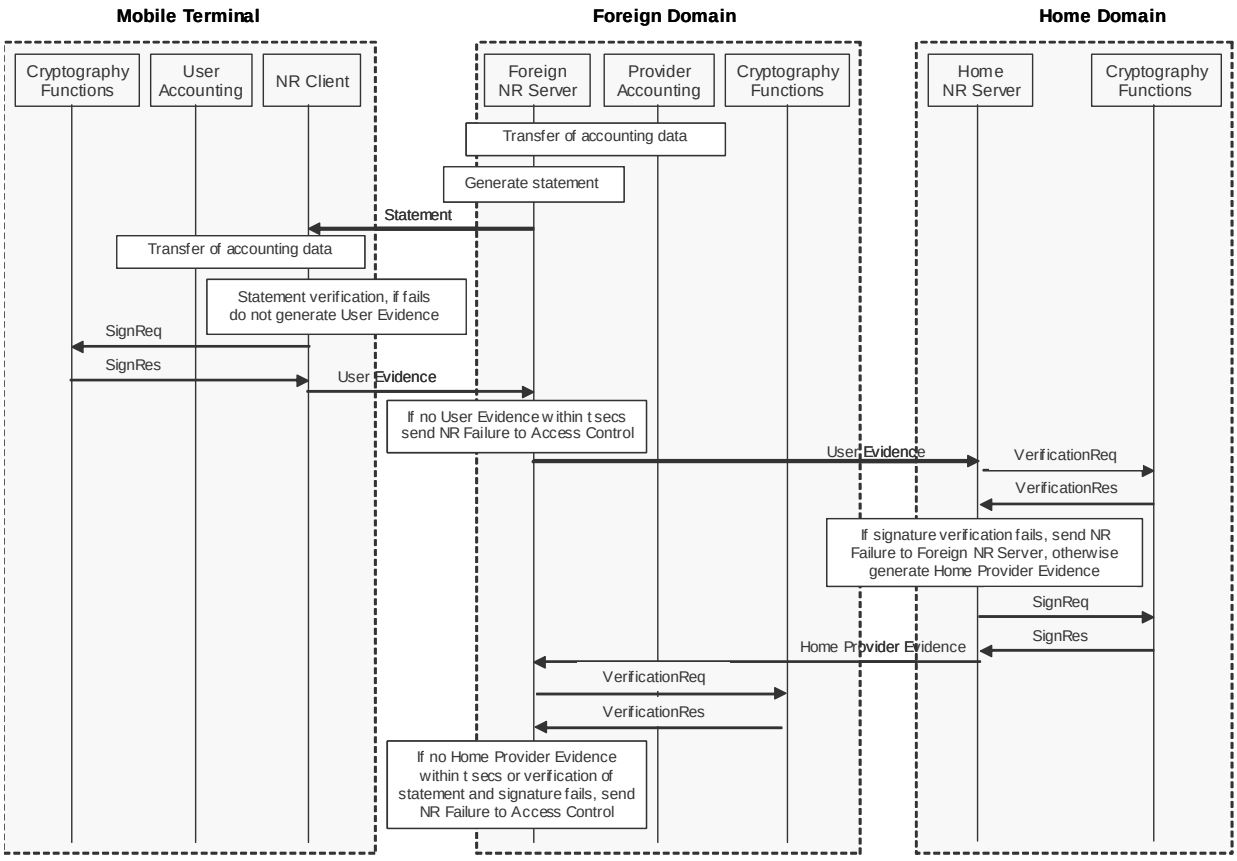

Figure 4: Non-repudiation Interactions.

sequences of interactions are possible, reflecting the different alternatives for statement generation as well as whether FP is capable of verifying users' digital signatures. Note that an error-free transmission of evidences on the transport layer does not guarantee the correctness of evidences. Hence, a response from the NR Server is required after receiving an evidence. However, to reduce traffic, only a NAK (Not Acknowledgement) message is sent to the NR Client, if the NR Server fails to verify the evidence. Furthermore, the NR Server has to inform the Access Control module, if the user does not send any additional or further evidences. In this case, the service will be terminated.

4.2 Mobility Considerations

A clear impact of mobility on the non-repudiation service exists with respect to two areas.

4.2.1 Layer-3 Handover:

The NR Server is located within the HP's and FP's network, while the NR Client is located within the user's Mobile Terminal. Due to the fact, that the terminal is mobile its Care-of Address (CoA) can change. Two kinds of addresses are available for an NR Client to use: Home Address and CoA. Table 5 compares the use of a Home Address with the use of a CoA for the communication between the NR Server and the NR Client.

4.2.2 Inter-Domain Interoperability:

For non-repudiation to be effective it must be applicable across administrative domains. To achieve this inter-domain interoperability, a standard transport protocol and a standardized format for these statements and evidences have to be used. This allows for the NR Client to communicate with the NR Server of a different administrative domain. In addition, this also allows for NR Servers of different administrative domains to interact with each other across such domain boundaries.

Table 5: Home Address versus CoA.

	TCP	**UDP**
Home Address	The use of Home Address makes mobility transparent to transport layer. However, without route optimization (communicating through the Home Agent) it is inefficient due to triangular routing. With route optimization means that the NR Server machine must have a Mobile IP stack (In Mobile IP terminology, NR Server acts as a Correspondence Node).	
	Recommended, because TCP is reliable.	UDP is not reliable. Since the transfer of non-repudiation evidences must be reliable (the evidences must not be lost in the transmission), message retransmission and duplicate detection mechanisms must be implemented if using UDP.

Table 5: Home Address versus CoA.

Care-of Address	In order to send a packet to an NR Client, the NR Server must know the current CoA of the Mobile Terminal. This CoA can be obtained from: • Entities in the network that know this information (aware of Mobile Terminals' layer-3 handover), *e.g.*, location server, handover module, access control module. • The previous packet sent by the NR Client to the NR Server. To obtain a fresh information on CoA in this packet, the NR Client is required to send a notification message to the NR Server after each handover. In this case the NR Client needs to be aware of handover, i.e., it must be told of CoA changes. • The accounting data. To obtain a fresh information on CoA in this data, the entity responsible for accounting must send accounting data to the Accounting module after each handover.	
	Not recommended, because of possible TCP session interruption.	Recommended, because there is no such session interruption. However, UDP is not reliable. Since the transfer of non-repudiation evidences must be reliable (evidences must not be lost in transmission), message retransmission and duplicate detection mechanisms must be implemented if using UDP.

4.3 Security Considerations

Several security aspects have to be considered in providing a non-repudiation service, in particular the integrity of non-repudiation statements, the privacy of user identities during transfer, and a malicious use of keys. Before going into detail of security considerations, it is necessary to define the notations applied:

Ks_A = (Private) signature key of A

Kv_A = (Public) verification key of A

Ke_A = (Public) encryption key of A

Kd_A = (Private) decryption key of A

$\{X\}_{Ks_A}$ = (The one-way hash of) X is signed using Ks_A

$\{X\}_{Ke_A}$ = X is encrypted using Ke_A

X1, X2 = X1 concatenated with X2

f_m = Message of type m

U = User, HP = Home Provider, FP = Foreign Provider

4.3.1 Statement Integrity

Suppose that statements are generated by the NR Server and sent to the NR Client to be singed. If these statements are intercepted and modified by a malicious node in-between the NR Server and the NR Client, the statement verification by the NR Client will fail and no evidence will be generated. This situation leads to a service termination and it can be seen as a Denial-of-Service attack.

In order to solve this threat, it is necessary to secure the communication path between the NR Server and the NR Client. A straightforward analysis shows that the problem is based in a possible modification of the statement and not in the disclosure of information. Hence, message confidentiality is not required. The NR Client has to be ensured that the NR Server has sent the statement, thus, authentication is needed. Furthermore, to be sure that the statement has not been changed, message integrity is required also. To achieve statement integrity, the NR Server must produce a hash of the statement and encrypts it. There exist two alternatives: The first one uses symmetric keys to encrypt the hash, and the second one uses public cryptography. If symmetric keys are applied, they can be negotiated between the NR Server and the NR Client, using the Diffie-Hellman key agreement algorithm with authentication steps. In this respect the HP can act as a key distributor, as he has a security association with the FP and he holds the public encryption key of the user.

If asymmetric keys are used the NR Server signs the hash of the non-repudiation statement with its private key and sends the statement together with the signed hash to the NR Client. The NR Client checks the integrity of the statement and authenticates the NR Server by verifying its signature.

4.3.2 Protection Against Fake Evidence Attacks

Fake evidences are evidences which try to prove fictitious service consumption of a user. An attack by generating fake evidences is called fake evidence attack. Obviously, a user is not interested in generating fake evidences of his own service consumption, because by doing so, he will harm himself. A malicious provider can generate a public-private key pair, then generates fake evidences using the private key of this pair, and claims that the public key of this pair is a verification key of a certain user. This allows for the provider to charge this user for services he did not consume.

To cope with this problem two alternatives are proposed:

- With TTP: the verification key of a user must be certified by a Certification Authority (CA) or
- Without TTP: upon conclusion of a contract with an HP a user signs a paper document specifying his verification key.

Fake evidences can be generated also, if the signature key of a user is compromised. Therefore, the user must be allowed to revoke the validity of the signature key (*e.g.*, by revoking the verification key certificate). This has the consequence that an evidence generated after the revocation of the signature key is invalid. This also means that an evidence must carry information about its generation time. This information must be generated by a TTP. The approaches presented in [15] using an online or offline Time-Stamping Authority (TSA) is highly appropriate for this purpose. In both approaches, a malicious provider cannot generate fake evidences with a time-stamp before the revocation time of a user's signature key. However, the involvement of an online TSA makes mass evidence generation not too efficient. The offline approach is less secure but more efficient.

In the online approach, an evidence is time-stamped by a TSA before being sent to the HP. Two additional fields are needed to include this information within the evidence: Evidence Generation Time-stamp (EGTS) and TSA Digital Signature (TSADS). The TSADS is the digital signature of the TSA over the hash of the EGTS and the UDS.

In the offline approach two types of signature keys are defined: revocable and irrevocable signature keys. The revocable signature key is a long-term master key which is used to issue a temporary (short-term) verification key certificate

of the irrevocable signature key. The temporary certificate is time-stamped by a TSA and the respective irrevocable signature key is used to generate evidences. The expiry time of the evidence is defined the same as the expiry time of the temporary certificate which must be greater than the time-stamp of the TSA. The temporary certificate C'_{user} contains the following information: Kv'_U, Te', Tg', $\{Kv'_U, Te'\}_{Ks_U}$, and $\{\{Kv'_U, Te'\}_{Ks_U}, Tg'\}_{Ks_{TSA}}$, where

Kv'_U = temporary verification key of the user

Te' = expiry time of C'_{user}

Tg' = generation time of C'_{user}

The evidence sent to the HP must contain this temporary certificate or a reference to this certificate.

Eliminating the involvement of a TSA entirely means to take a risk of this kind of attack. The possibility of having such an attack from an FP is, however, relatively small. A malicious FP must own the compromised signature key of both the HP and the user in order to be able to perform this attack. A malicious HP might lose his customers, once it is known that he has cheated one of his customers. The risk of the attack can be reduced by making the attack less attractive *e.g.*, by restricting the cost spent by a customer in a billing period.

4.3.3 Protection Against Denial of Consumption Attacks

Eliminating time-stamp information on evidence generation allows for another type of attack: Denial-of-Service Consumption. A user or an HP can argue that the evidence is not valid as it is generated after revocation of his signature key. To protect against this attack a time-stamp for evidence generation is needed. Unlike fake evidence attack, the time-stamp needs not be generated by a TSA. It is enough if the time-stamp is generated by the user and the HP themselves.

Suppose that the Consumption Interval is between T1 and T2, and Tr_A and Te_A determine the revocation time and the expiry time of A's signature key, respectively. The following interactions between the FP, the user (U), and the HP are needed to protect against the abovementioned attack.

1. FP => U: Statements, User Signing Deadline (Td_U)
   If a User Evidence (UE) should be generated before service usage, then Td_U < T1+delta < T2, otherwise Td_U < T2+delta.
2. U => FP: UE including the generation time Tg_U
3. FP => HP: UE, Td_U, HP Signing Deadline (Td_{HP})
4. HP accepts UE if $Tg_U < Now \le Td_U$ and $Tg_U < Tr_U < Te_U$
5. HP => FP: HP Evidence (HPE) including UE and the generation time Tg_{HP}
6. FP accepts HPE, if $Tg_{HP} < Now \le Td_{HP}$ and $Tg_{HP} < Tr_{HP} < Te_{HP}$

4.3.4 Identity Privacy

Normally, billing is done by the HP. Hence, the HP needs to know the identity of the user. However, an FP does not need to know who is consuming his service. He only needs to know the HP of the user who is responsible for the service consumption.

The user's identity generated upon contract establishment (subscription) is identified by a Registration Identifier (RegID). This identifier must not be disclosed to the FP when the user consumes services in the foreign domain. But the FP needs to assign accounting data correctly to a user. For this purpose a temporary Virtual User Identifier (VID) is required. The mapping of a VID to a RegID is known only to the HP (and of course the user). The concept of VID is based on the research work performed within the Daidalos project [4], [5].

A VID is generated either by the HP or the user, however the mapping between a VID and a RegID must be signed by the user prior to its usage. In order to be unique a VID must contain an HPID and an identifier which is unique in the HP's domain, *e.g.*, the one-way hash of the concatenation of a timestamp with the user's RegID.

Before services can be consumed in a foreign domain, the user has to be authenticated and authorized by the HP. In the authentication process an Authentication Request message is sent to the FP's Access Control module. This message contains VID, HPID, and the encrypted signed mapping of VID to RegID:

User => FP: $f_{AuthenticationReq}$, VID, HPID, X, where

$X = \{VID, RegID, T, L, Cred, \{VID, RegID, T, L\}_{Ks_U}\}_{Ke_{HP}}$,

T = Timestamp, L = Lifetime, and Cred = Credentials.

X will be forwarded to the HP. As the HP holds the verification key of the user, the mapping can be verified. Having this mapping, non-repudiation statements and evidences can use VID instead of RegID in the UID field.

4.3.5 Protection Against Denial of NR Service Attacks

Having direct communications between an NR Server and the NR Clients can endanger the NR Server. One or some malicious NR Clients can generate many fake evidences which overload the NR Server. This can lead to a Denial of NR Service attack. To protect against this kind of attack, three solutions are envisaged:

- Employing a set of NR Agents, each of which mediates between a subset of NR Clients and the NR Server. NR Clients are not directly connected to the NR Server anymore. Surely each NR Agent can be under attack, but this is less dangerous than an attack on the NR Server.
- Deploying a set of co-operating NR Servers within a domain to "back up" each other in case of failures.
- A combination of the above two solutions.

4.4 Fairness

A major problem in proving service consumption is the generation of an evidence which contains the *real* consumption, if the user or the provider may not play fair. On one hand, if an evidence has to be generated *after* a service consumption (pay after use), there is a risk that the user may not send the evidence after consuming the service. On the other hand, if an evidence has to be generated *before* a service consumption (pay before use), the provider may not deliver the service after obtaining the evidence. Therefore, NorCIS defines fairness by relating evidences with real consumption: a protocol for non-repudiation of service consumption is fair,

if in case a service is delivered or consumed it provides for the service provider and the service user a valid evidence containing the real service consumption after completion of the protocol.

One way to reduce the abovementioned risk is by dividing the whole duration of service consumption into smaller intervals depending on the accounting scheme, and requiring an evidence to be generated in each interval. However, this does not really solve the problem of fairness in each interval.

Different “types” of services require different approaches to achieve fairness. Approaches for fair non-repudiation protocols described in [15] can be used to deliver content services which are not streaming, and allow for a provider to obtain from the user the evidence of receipt. However, modifications and extensions to these approaches are needed in order to apply them to services consumed with a time-based accounting. NorCIS proposes a fair non-repudiation protocol with online TTP for consumption of such services.

In order to show the main idea and to not complicate the protocol description, service consumption in the home domain is assumed. The extension for service consumption in foreign domain is straightforward. To obtain a concise description of the protocol, the following abbreviations and notations are applied in addition to previously defined notations:

SOSC = Statements on Service Consumption
EOSC = (User) Evidence of Service Consumption
K = Symmetric encryption key
K^{-1} = Symmetric decryption key
Td_{sub} = Decryption key submission deadline = begin of CI
Td_{rev} = Decryption key revelation deadline = end of CI
H(X) = One-way hash of X

The following five steps describe the proposed protocol. Note, that a secure communication channel is assumed.

1. HP => U: $f_{EvidenceReq}$, SOSC, Td_{sub}, Td_{rev}
2. U => HP: $f_{EvidenceRes}$, H(SOSC), $\{EOSC\}_K$,
   $\{f_{EvidenceRes}, H(SOSC), \{EOSC\}_K\}_{Ks_U}$
3. HP => TTP: U, H(SOSC), Td_{sub}, Td_{rev}
   If U does not deliver decryption key K^{-1} by Td_{sub},
   TTP => HP: $f_{KeySubmissionTimeout}$, H(SOSC),
   and the requested service is not delivered or is discontinued. Otherwise, it will be delivered or continued at Td_{sub}
4. U => TTP: $f_{KeyPublishReq}$, HP, H(SOSC), K^{-1}, Td_{rev},
   $\{f_{KeyPublishReq}, HP, H(SOSC), K^{-1}, Td_{rev}\}_{Ks_U}$
   If the Td_{rev} sent by U does not match the one sent by HP,
   TTP => HP: $f_{KeyRevelationNotMatch}$, H(SOSC),
   and the requested service is not delivered or is discontinued. Otherwise, it will be delivered or continued at Td_{sub}.
   If the service being consumed is terminated before the end of the pre-defined Consumption Interval,
   U => TTP: $f_{ServiceTerminated}$, H(SOSC),
   $\{f_{ServiceTerminated}, H(SOSC)\}_{Ks_U}$
   TTP => HP: $f_{TerminateService}$, H(SOSC),
   $\{f_{TerminateService}, H(SOSC)\}_{Ks_{TTP}}$
5. U, HP <= TTP: $f_{Publish}$, U, HP, H(SOSC), T, K^{-1},
   $\{f_{Publish}, U, HP, H(SOSC), T, K^{-1}\}_{Ks_{TTP}}$, where
   T = Td_{rev} or the time TTP received $f_{ServiceTerminated}$.

The value of T marks the real end of the service consumption. Therefore, this protocol is able to provide evidence of real service consumption, if the communication channels are reliable, and the communication delay is negligible.

4.5 Performance and Scalability Consideration

The non-repudiation service employs a communication protocol that increases the number of messages exchanged through the access network. Messages belonging to the non-repudiation service are an overhead to network traffic, and reduce the effective bandwidth for users' data traffic. In a wideband access, this overhead is not critical, but in narrow band access technologies with smaller access speeds, the overhead of these messages can be relevant.

During the consumption of a particular service, several evidences may need to be generated and sent in regular intervals. These intervals between two successive evidences depend on the pricing, hence, the accounting scheme. One way to reduce the impact of these messages is to control the interval between these messages. On one hand, the larger the interval between two successive evidences, the less traffic they generate. On the other hand, a smaller interval has a lower risk of loss for both the provider and the user. Hence, defining the right interval is a trade-off between communication overhead and business risk.

Furthermore, with respect to scalability considerations, an NR Server is connected to a number of NR Clients and to the respective NR Servers of other administrative domains. The number of NR Clients which are connected to an NR Server can be large. To cope with this problem similar solutions as presented for the security considerations are proposed:

- Employing a set of NR Agents, each of which mediates between a subset of NR Clients and the NR Server. An NR Agent can be located in each Access Router. This reduces the number of connections to the NR Server.
- Deploying a set of co-operating NR Servers within an administrative domain to balance the load.
- A combination of the above two solutions.

5 NorCIS Optimization

Having described the NorCIS model and implementation architecture, useful optimizations of the architecture and the protocol with respect to security, scalability, and non-repudiation traffic are considered and evaluated.

5.1 Employment of NR Agents

Employing a set of NR Agents adds to advantages as described above: a better scalability and a higher security. Additionally, if an NR Server should not rely on Mobile IP and should not be aware of mobility, an NR Agent can be designed to make mobility transparent to the NR Server. Assume that the entity in the network which is aware of CoA changes is called an Handover-aware Entity (HO-aware

Entity). To hide mobility from this NR Server an NR Agent needs an interface to the HO-aware Entity and to other NR Agents. The specification of these interfaces are summarized in Table 6. For easier management, each Access Router hosts an NR Agent. In turn, each NR Agent supports a TCP connection with the NR Server and a UDP connection with each of the NR Clients of the Mobile Terminals connected to the respective Access Router.

Table 6: NR Agent Interfaces.

Interface	Specification
HO-aware Entity - NR Agent	Communicating the new CoA of a Mobile Terminal to the NR Agent.
NR Agent - NR Agent	Forwarding messages from NR Server to NR Clients having changed their points of attachment.
NR Agent - NR Client	Forwarding messages from NR Server or NR Agent to NR Client. Receiving messages from NR Client to be forwarded to NR Server.
NR Agent - NR Server	Receiving messages from NR Server to be forwarded to NR Clients. Forwarding messages from NR Client to NR Server.

5.2 Unsolicited Evidence Generation

As described, statements can be generated by the user. This possibility leads to unsolicited evidence generation, which is possible, when Mobile Terminals deploy accounting of service consumption. Unsolicited evidence generation allows for an efficient and simple interaction between the provider and the user. For example, for a time-based or volume-based pricing scheme, a user can sign the service consumption statement and send it unsolicitedly each time a certain time has elapsed or a certain volume has been achieved.

For time-based and volume-based accounting, unsolicited evidence generation can make use of chained hashes, which work as follows:

1. The user generates nonce r; calculates $H_0 = H(r)$ and chained hashes $H_n = H(H_{n-1})$.
2. The user sends to the FP the user evidence containing the last hash in the chained hashes: Y, $\{H(Y)\}_{Ks_U}$, where Y = SPID, UID, SvcID, SessID, CI, H_m.
3. The FP sends the user evidence to the HP, who signs the user evidence with his signature key Ks_{HP}, and sends it back to the FP. This signed user evidence contains HPID as SVID and the verification key Kv_U of the user: $\{\{H(Y)\}_{Ks_U}, SVID, Kv_U\}_{Ks_{HP}}$.
4. For all subsequent evidences of the same session: i=1...n, $n \le m$, the user sends to the FP: SPID, UID, SvcID, SessID, CI, H_{m-i}.
5. If the session is not yet terminated after the first hash H_0 in the chained hashes has been sent, repeat from step 1.

An evidence stored in the database comprises of: SPID, UID, SvcID, SessID, First CI, Last CI, H_m, H_n, SVID, Kv_{HP}, and $\{\{H(Y)\}_{Ks_U}, SVID, Kv_U\}_{Ks_{HP}}$.

This protocol improves the protocol proposed in [19] by extending the evidence with service consumption information, binding the signature with this information, and involving the HP in the evidence approval.

6 Summary and Conclusions

The NorCIS architecture proposed and its interactions allow for the secure generation and transfer of irrefutable evidences of a consumption of differentiated services in an inter-domain mobility environment offering Internet services. The structure of these evidences allows for capturing information required for a wide variety of accounting schemes as well as information to protect the user and the provider against a number of various attacks. This paper has shown additionally different alternative solutions for the statement generation process. Obviously, an implementation of NorCIS which supports different alternative solutions, hence also different message sequences, requires configurable non-repudiation entities.

Concluding, the NorCIS model and architecture provides for a real-world case of multiple service providers in an Internet-based environment for a value-added service of non-repudiated consumption of Internet services. In due course, this determines an essential support of business-critical and high risk services, which are required for a number of commercial applications. Only if such a service is scalable and widely offered between multiple providers, technological prerequisites can be integrated into tomorrow's networks. As shown in this proposal, the non-repudation protocol and its underlying architecture provide for this technology and flexibility required, and the set of necessary configuration potentials as well as interfaces has been designed.

With respect to future work, this proposal's focus should be extented beyond user mobility, to the impact of session mobility on the non-repudiation of service consumption. This as well as the design of a fair and efficient non-repudiation protocol for a general service consumption — including services consumed with volume-based accounting scheme — is subject to further study.

Acknowledgment

This work has been performed partially in the framework of the EU IST project Daidalos (FP6-2002-IST-1 Contract No. 506997), where ETH Zürich has been funded by the Swiss Bundesministerium für Bildung und Wissenschaft, Bern under grant No. 03.0141.

The authors would like to extend many thanks to their project partners within Daidalos, especially to José Miguel Fernandes, and to Egon Burgener for supporting the initial design of non-repudiation of service consumption.

References

[1] N. Asokan, V. Shoup, and M. Waidner: "Asynchronous protocols for optimistic fair exchange"; IEEE Symposium on Security and Privacy, Oakland, California, May 1998.

[2] C. de Laat, G. Gross, L. Gommans, J. Vollbrecht, D. Spence: "Generic AAA Architecture"; IETF, RFC 2903, August 2000.

[3] E. Burgener: "Non-repudiation of MobyDick Service Consumptions"; Diploma Thesis, ETH Zürich, 2002.

[4] Daidalos — Designing Advanced Interfaces for the Delivery and Administration of Location independent Optimised personal Services; European Union 6th Framework Program IST, FP6-2002-IST-1, URL: http://www.ist-daidalos.org/.

[5] Daidalos: "A4C Framework Design Specification"; Deliverable D341, September 2004.

[6] S. Gürgens, C. Rudolph, and H.Vogt: “On the Security of Fair Non-repudiation Protocols”; International Conference on Information Security, Springer Verlag, 2003.

[7] P. Hasselmeyer, M. Schumacher, M. Voß: “Pay As You Go — Associating Costs With Jini Leases”; 4th International Enterprise Distributed Object Computing Conference (EDOC 2000), Makuhari, Japan, IEEE Publishing, pp. 48-57, September 2000.

[8] ISO/IEC 13888-1: “Information Technology — Security techniques — Non-repudiation — Part 1: General”; ISO/IEC, 1997.

[9] D. Johnson, C. Perkins, J. Arkko: “Mobility Support in IPv6”; IETF, RFC 3775, June 2004.

[10] S. Kremer, O. Markowitch, and J. Zhou: “An Intensive Survey of Non-repudiation Protocols”; Computer Communications Journal, Vol. 25, 2002.

[11] H. Y. Lin and L. Harn.: “Authentication protocols for personal communication systems”; ACM SIGCOMM'95, pp. 256-261, Cambridge, Massachusetts, August 1995.

[12] O. Markowitch, D. Gollmann, and S. Kremer: “On Fairness in Exchange Protocols”; Int. Conf. on Information Security and Cryptology, 2002.

[13] C. Perkins (Ed.): “IP Mobility Support for IPv4”; IETF, RFC 3344, August 2002.

[14] R. Shirey: “Internet Security Glossary”; IETF Informational RFC 2828, May 2000.

[15] J. Zhou: “Non-repudiation in Electronic Commerce”; Artech House, 2001.

[16] J. Zhou, R. H. Deng, F. Bao: “Evolution of Fair Non-repudiation with TTP”; Australasian Conference on Information Security and Privacy, pp. 258-269, University of Wollongong, Australia, April 1999.

[17] J. Zhou, R. Deng, and F. Bao: “Some Remarks on a Fair Exchange Protocol”; Public Key Cryptography, Springer-Verlag, 2000.

[18] J. Zhou, D. Gollmann: “A Fair Non-repudiation Protocol”; IEEE Symposium on Security and Privacy, Oakland, California, May 1996.

[19] J. Zhou, K.Y. Lam: “Undeniable Billing in Mobile Communication”; Proceedings of 4th ACM/IEEE International Conference on Mobile Computing and Networking, pp. 284-290, ACM Press, Dallas, Texas, October 1998.