Unwanted Traffic and Information Disclosure in VoIP Networks

Threats and Countermeasures

Ge Zhang

Faculty of Economic Sciences, Communication and IT

Computer Science

DISSERTATION | Karlstad University Studies | 2012:28

Unwanted Traffic and Information Disclosure in VoIP Networks

Threats and Countermeasures

Ge Zhang

DISSERTATION | Karlstad University Studies | 2012:28

Unwanted Traffic and Information Disclosure in VoIP Networks - Threats and Countermeasures

Ge Zhang

DISSERTATION

Karlstad University Studies | 2012:28

ISSN 1403-8099

ISBN 978-91-7063-432-1

Distribution:
Karlstad University
Faculty of Economic Sciences, Communication and IT
Computer Science
SE-651 88 Karlstad, Sweden
+46 54 700 10 00

Print: Universitetstryckeriet, Karlstad 2012

WWW.KAU.SE

"Know your enemy, know yourself and you can
fight a hundred battles without disaster..."

"Thus, what is of supreme importance in war
is to attack the enemy's strategy...."

The Art of War
Sun Tzŭ

Unwanted Traffic and Information Disclosure in VoIP Networks

Threats and countermeasures

GE ZHANG
D*epartment of Computer Science, Karlstad University*

Abstract

The success of the Internet has brought significant changes to the telecommunication industry. One of the remarkable outcomes of this evolution is Voice over IP (VoIP), which enables realtime voice communications over packet switched networks for a lower cost than traditional public switched telephone networks (PSTN). Nevertheless, security and privacy vulnerabilities pose a significant challenge to hindering VoIP from being widely deployed. The main object of this thesis is to define and elaborate unexplored security & privacy risks on standardized VoIP protocols and their implementations as well as to develop suitable countermeasures. Three research questions are addressed to achieve this objective:

- **Question 1:** *What are potential unexplored threats in a SIP VoIP network with regard to availability, confidentiality and privacy by means of unwanted traffic and information disclosure?*
- **Question 2:** *How far are existing security and privacy mechanisms sufficient to counteract these threats and what are their shortcomings?*
- **Question 3:** *How can new countermeasures be designed for minimizing or preventing the consequences caused by these threats efficiently in practice?*

Part I of the thesis concentrates on the threats caused by "unwanted traffic", which includes Denial of Service (DoS) attacks and voice spam. They generate unwanted traffic to consume the resources and annoy users. Part II of this thesis explores unauthorized information disclosure in VoIP traffic. Confidential user data such as calling records, identity information, PIN code and data revealing a user's social networks might be disclosed or partially disclosed from VoIP traffic. We studied both threats and countermeasures by conducting experiments or using theoretical assessment. Part II also presents a survey research related to threats and countermeasures for anonymous VoIP communication.

Keywords: Voice over IP, Session Initiation Protocol, Realtime Transport Protocol, Network Security, Denial of Service, Timing Attack, Privacy and Anonymity.

Acknowledgments

First and foremost I would like to express my sincere gratitude to my supervisor, Prof. Simone Fischer-Hübner, who has supported me throughout my thesis with her patience, knowledge and experience whilst giving me so much freedom to explore and discover new areas of Voice over IP security & privacy. I also would like to thank my co-advisor, Prof. Andreas J. Kassler, for many encouraging discussions regarding research and education.

I had the great fortune to work with those nice people in Computer Science Department at Karlstad University. For creating a great work environment, a sincere thank goes to all my colleagues. I am especially grateful to Inger Bran for the helpful administrative support. I also want to thank Stefan Alfredsson and Jonas Karlsson for helping me on vulnerability analysis lab configurations. Further, special appreciation is given to all and previous members of the Privacy and Security Group (PRISEC). They are Prof. Simone Fischer-Hübner, Prof. Stefan Lindskog, Dr. Thijs J. Holleboom, Dr. Leonardo A. Martucci, Hans Hedbom, Reine Lundin, Stefan Berthold, Tobias Pulls and Philipp Winter. I really enjoyed collaborating, discussing and exchanging ideas with such inspiring researchers. Moreover, great thanks go to the people who make my life in Karlstad so enjoyable. I would like to thank Peter Dely, Marcel Cavalcanti de Castro, Stefan Berthold, Mohammad Rajiullah, Andreas Lavén and Julio Angulo for organizing sports and games (gym, innebandy, soccer and Super.Mario on Wii). In addition, I also would like to thank those who provided delicious cakes on Fredags Fika.

Before coming to Karlstad, I have had good opportunities to work with other research institutes. My stay in Berlin at Fraunhofer FOKUS institute and Hasso-Plattner Institute (Potsdam) is a valuable experience because I started my early academic life there. My sincere thanks must go to Prof. Thomas Magedanz, Dr. Sven Ehlert, Dr. Yacine Rebahi, Prof. Christoph Meinel and Dr. Feng Cheng for supporting me with valuable advice and co-authoring papers.

I would like to thank my family in China for letting me pursue my dream for so long and so far way from home. My deepest appreciation must go to my parents, Jinyi Zhang and Ming Qian. I am forever indebted to them for their endless love and support. Furthermore, I would like to thank Mandi Zhang and Hui Xie for accompanying and encouraging me.

Finally, I want to thank C-BIC (Compare Business Innovation Centre) and .SE (Stiftelsen för Internetinfrastruktur) for their financial sponsorship of my research.

Karlstad, June 2012 Ge Zhang

List of Appended Papers

This thesis is comprised of the following 10 peer-reviewed papers. References to the papers will be made using the Roman numbers associated with the papers such as Paper I.

Part I – Unwanted traffic:

I. **Ge Zhang**, Simone Fischer-Hübner and Sven Ehlert: Blocking attacks on SIP VoIP proxies caused by external processing, *Special Issue on Secure Multimedia Services of Journal of Telecommunication Systems, 45(1)*, pp.61-76, Sep, 2010, Springer.

II. **Ge Zhang**, Jordi Jaen Pallares, Yacine Rebahi and Simone Fischer-Hübner: SIP proxies: new reflectors? attacks and defenses, in proceedings of *the* 11^{th} *Joint IFIP TC6 and TC11 Conference on Communications and Multimedia Security (CMS' 10)*, LNCS 6109, pp.142-153, Linz, Austria, May, 2010, Springer.

III. **Ge Zhang** and Simone Fischer-Hübner: Detecting near-duplicate SPITs in voice mailboxes using hashes, in proceedings of *the* 14^{th} *Information Security Conference (ISC' 11)*, LNCS 7001, pp. 152-167, Xi'an, China, Oct, 2011, Springer.

Part II – Information disclosure:

IV. **Ge Zhang**, Simone Fischer-Hübner, Leonardo Martucci and Sven Ehlert: Revealing the calling history on SIP VoIP system by timing attacks, in proceedings of *the* 4^{th} *International Conference on Availability, Reliability and Security, (ARES' 2009)*, pp.135-142, Fukuoka, Japan, March, 2009, IEEE computer society.

V. **Ge Zhang** and Simone Fischer-Hübner: Peer-to-peer VoIP communications using anonymisation overlay networks, in proceedings of *the* 11^{th} *Joint IFIP TC6 and TC11 Conference on Communications and Multimedia Security (CMS' 10)*, LNCS 6109, pp.130-141, Linz, Austria, May, 2010, Springer.

VI. **Ge Zhang** and Stefan Berthold: Hidden VoIP calling records from networking intermediaries, in proceedings of *the* 4^{th} *Principles, System and Applications of IP Telecommunications (IPTCOMM' 10)*, pp.15-24, Munich, Germany, Aug, 2010, ACM.

VII. **Ge Zhang**: Timing attacks on a centralized presence model, in proceedings of *IEEE International Conference on Communications 2011 (ICC' 11)*, pp.1-5 Kyoto, Japan, June, 2011, IEEE communication society.

VIII. **Ge Zhang**: Analyzing keystroke patterns of PIN code input for recognizing VoIP users, in proceedings of *the* 26^{th} *IFIP conference on Future Challenges in Security and Privacy for Academia and Industry (IFIP SEC' 11)*, AICT 354, pp.247-258, Lucerne, Switzerland, June, 2011, Springer.

IX. **Ge Zhang** and Simone Fischer-Hübner: Timing attacks on PIN input in VoIP networks (short paper), in proceedings of *the* 8^{th} *Conference on Detection of Intrusions*

and Malware & Vulnerability Assessment (DIMVA' 11), LNCS 6739, pp.75-84, Amsterdam, The Netherlands, July, 2011, Springer.

X. **Ge Zhang** and Simone Fischer-Hübner: A survey on anonymous Voice over IP communications: Attacks and defences, *under submission*.

Some of the papers have been subjected to some minor editorial changes.

Comments on my Participation

I am the principal contributor to all papers listed above. The ideas of these papers stem from the discussion from co-authors and myself. In addition, my contributions also include design of the testbed, implementation of the prototypes, writing proof-of-concept programs and conducting experiments for evaluations.

Most parts of the written material was done by myself. Paper II is a joint effort between Karlstad University and Fraunhofer FOKUS Institute and the discussion part of Section 4.3 in Paper II was done by Jordi Jaen Pallares.

All the coauthors offered their help on proof-reading, commenting and revising the papers.

Other Papers

Apart from the papers included in this thesis, I have also authored the following papers.

1. **Ge Zhang**, Sven Ehlert Thomas Magedanz and Dorgham Sisalem: Denial of service attack and prevention on SIP VoIP infrastructures using DNS flooding, in proceedings of *the* 1^{st} *Principles, System and Applications of IP Telecommunications (IPTCOMM' 07)*, pp. 57-66, New York City, USA, July, 2007, ACM.

2. Sven Ehlert, **Ge Zhang**, Dimitris Geneiatakis, Tasos Dagiuklas, Georgios Kambourakis, Jiri Markl and Dorgham Sisalem: Two layer denial of service prevention on SIP VoIP infrastructures, *the International Journal for the Computer and Telecommunications Industry (COMCOM), 31(10)*, pp.2443-2456, Jun, 2008, Elsevier.

3. **Ge Zhang**, Feng Cheng and Christoph Meinel: Towards secure mobile payment based on SIP, in proceedings of *the* 15^{th} *IEEE International Conference on Engineering of Computer-Based Systems (ECBS' 08)*, pp.96-104, Belfast, UK, March, 2008, IEEE computer society.

4. **Ge Zhang**, Feng Cheng and Christoph Meinel, SIMPA: a SIP-based mobile payment architecture, in proceedings of *the* 7^{th} *International Conference on Computer and Information Science (ICIS' 08)*, pp.287-292, Portland, USA, May, 2008, IEEE computer society.

5. Sven Ehlert, **Ge Zhang** and Thomas Magedanz: Increasing SIP firewall performance by ruleset size limitation, in proceedings of *the* 19^{th} *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC' 08), VoIP Technologies Workshop*, pp. 1-6, Cannes, France, Sep, 2008, IEEE communication society.

6. **Ge Zhang**: An analysis for anonymity and unlinkability for a VoIP conversation, in proceedings of *the* 5^{th} *IFIP Primelife Summer School*, AICT 320, pp.198-212, Nice, France, Sep, 2009, Springer.

7. **Ge Zhang** and Yacine Rebahi: Side effects of identity management in SIP VoIP environment, *Information Security Technical Report, 16(1)*, pp.29-35, Sep, 2011, Elsevier.

Contents

Introductory Summary

1 Introduction

For almost one hundred years already before the emergence of the Internet, realtime voice communication across long distances has been implemented using Public Switched Telephone Network (PSTN). In recent decades, the Internet has quickly established itself as an excellent platform for data and multimedia content distribution. This development has motivated telephony service providers to consider new business models that can take advantage of Internet technologies and the protocols. Voice over IP (VoIP), the transmission of voice traffic using the Internet Protocol, is designed to provide telephony equivalent functions with additional benefits like cost saving and flexibility [1]. For these benefits, more and more people nowadays begin to make phone calls using VoIP (e.g., Skype [2]) on their computers. In addition, many companies also consider to setup VoIP infrastructures to reduce telephony costs. Nevertheless, the benefits of VoIP come along with some problems, one of which is the impact brought by security and privacy threats in the cyber space. The environment of VoIP makes it more vulnerable than PSTN: Many VoIP deployments use an interconnected network environment (e.g., the Internet) with standardized protocols, while PSTN is a closed network environment with proprietary protocols. Thus, it is easily possible for adversaries to find and exploit the vulnerabilities of VoIP implementations as well as to access the VoIP network infrastructures to launch attacks. Concrete examples have been presented in the real world: in 2006, an attacker was arrested and charged with making more than $1 million by breaking into VoIP services and illegally routing calls through their lines [3]. Thus it is unlikely for current VoIP to completely replace PSTN considering the security and privacy issues. Therefore, reducing security and privacy threats for VoIP is an important task and many research projects have been conducted towards this goal [4–8].

VoIP has two kinds of vulnerabilities, as summarized in [9]: One kind is the vulnerabilities of VoIP protocols and their implementations. The vulnerabilities of VoIP protocols need to be assessed by analyzing standardized VoIP protocols and understanding the vulnerabilities present in the protocols. This will be helpful for the development of security mechanisms that can be incorporated into the protocol specifications. The countermeasures against attacks can be easier to design if the potential flaws and their exploitations are understood. The other kind is inherited vulnerabilities coming from existing network infrastructures (e.g., a router on the network layer, or a DNS server on the application layer). VoIP applications rely on these infrastructures and thus the vulnerabilities on them may affect VoIP as well. In this way, all components in a VoIP network need to be protected, no matter it is a VoIP components or not. In addition, the interface between VoIP and non-VoIP components needs to be carefully designed. Compared to many popular applications on the Internet like email and web surfing, VoIP has higher requirement on real time. Thus, security mechanisms need to be designed carefully in regard to performance tradeoff.

1.1 Scope

In this thesis, we have studied the VoIP services using IETF standardized protocols rather than those using proprietary ones. Although Skype [2], a VoIP service using proprietary protocols, currently gains a significant success on business, it is difficult to analyze its details. Also it is not easy to setup a testbed to validate assumptions. In contrast, IETF standardized protocols are well documented and published for analysis. Moreover, there are a lot of open source VoIP products using standardized protocols available for both services and endpoints. Therefore, it has been feasible for us to implement test environments which allow us to evaluate our ideas and hypothesis in practice. In addition, standardized protocols have also been widely adopted by industry [10]. For these reasons, our research is only focused on IETF standardized protocols. Two kinds of protocols are especially important for VoIP. One is the signaling protocol which does not transmit voice packets, but is designed for establishing, controlling, modifying and terminating communications. Another is the voice transport protocol which transports voice traffic with realtime features. In this thesis, the scope of our research focuses on the Session Initiation Protocol (SIP) for signaling and the Realtime Transport Protocol (RTP) for voice transport.

The types of security and privacy studies in this thesis are unwanted traffic and information disclosure. Unwanted traffic may be caused by Denial of Service (DoS) attacks and Spam. These threats significantly reduce the availability of services and the user quality of experience. Information disclosure impacts confidentiality and privacy. VoIP traffic may disclose information about a user's secrets like PIN code or her social relations. Furthermore, there are instances where one also would like to make calls while staying anonymous. This thesis investigates both issues.

1.2 Objective

The objective of this thesis is to define and elaborate unexplored security and privacy risks in VoIP networks, especially including (1) unwanted traffic, like Denial of Service (DoS) attacks and Spam; and (2) information disclosure, like profiling Call Detail Records (CDR), buddy lists and users' PIN codes. We are also interested in finding out how easily these risks can be exploited by attackers in the real world, and in consequence, how high the risks are for VoIP services and users. We furthermore follow the objective to elaborate and develop several defending solutions to eliminate or minimize the impacts caused by the risks. Finally, we would like to know which solutions are the most practical and efficient.

1.3 Structure

This thesis presents an introductory summary and a collection of 10 papers in the area of security & privacy of VoIP services that were either authored or co-authored by the author of this thesis. The rest of the introductory summary is organized as follows. Section

2 presents fundamental background on standardized VoIP protocols. Section 3 lists some general security and privacy considerations for VoIP services. Further, Section 4 outlines the research questions of this thesis and Section 5 discusses the research methodologies that we applied. Related work to this research is outlined in Section 6, followed by a summary of our contributions in Section 7. Section 8 summarizes the contents of the 10 included papers which are divided into two parts: unwanted traffic and information disclosure. Finally, Section 9 provides the main conclusions of this thesis with an outlook on future work in this direction.

2 Background

This section presents with the basic concepts of VoIP and its standardized protocols.

2.1 VoIP

The Internet Protocol (IP) [11] is the standard for data transactions. Content of email and web pages are data and thus those applications can be built on the top of IP. When voice transport is combined with IP, two basic functions are required: (1) A *signaling* function for establishment, modification and termination of a voice conversation, and (2) a *voice transmission* function that carries voice traffic. There are both *standard protocols* and *proprietary protocols* for the implementation of the two functions. The Internet Engineering Task Force (IETF) standardizes protocols like Session Initiation Protocol (SIP) [12] for signaling function and Realtime Transport Protocol (RTP) [13] for voice transmission function. The specifications of the protocols are published as Request for Comments (RFC) on the IETF's homepage [14]. In contrast, the details of proprietary protocols, like Skype [2], are not available to the public, but are maintained internally.

Built on IP networks, VoIP has its own advantages and disadvantages compared with PSTN. The major advantages of VoIP include:

1. **Low costs:** By allowing voice to be converged on existing IP networks (e.g., Internet), the costs for deploying and operating VoIP services are lower than traditional PSTN. Another reason for low costs is that VoIP infrastructures can be software-based. Especially, there are a number of open source VoIP products (e.g., kphone [15], kcall [16], X-Lite [17], SER [18], OpenSIPS [19]). Installing these free software products on computers or smart phones, instead of buying expensive PSTN equipments, saves money for both users and service providers.

2. **Flexible and extendable features:** VoIP can be converged on IP networks along with other applications. Therefore it can provide more features than traditional PSTN services. For example, it is easy to integrate audio, video, game and email features

together with VoIP. In addition, software-based infrastructures can be easily updated. In contrast, PSTN infrastructures are not flexible since they are mostly hardware-based and are designed for specific purposes.

Nevertheless, there are also some disadvantages with VoIP as well:

1. **Security & privacy threats:** VoIP is deployed on interconnected IP networks (e.g., the Internet), an open network environment with public protocols. Thus, it is possible for unauthorized users to access the network infrastructure. In addition, VoIP applications need support from other applications (e.g., DNS servers), the vulnerabilities of those can pose a risk to VoIP services as well. In comparison, within a closed network environment and proprietary protocols, the costs of intruding PSTN networks is higher than the costs of intruding VoIP networks.

2. **Quality loss:** There are many quality issues for VoIP that do not exist for PSTN. Since PSTN infrastructures are implemented for specific purpose and the resources are reserved for each individual, acceptable quality can be guaranteed for PSTN. Nevertheless, IP networks are designed to support multiple purposes without resource reservation. Thus, load of the network varies from time to time. As a result, the quality of data transmission cannot be assured, which may lead to unexpected latency and packet loss. As telephony services are time-sensitive, these quality problems become the most essential barriers for VoIP applications. In addition, it is challenging to tackle both the security & privacy threats and the quality loss. For instance, a security mechanism applied to a VoIP system may introduce more latency since additional operations (e.g., encryption/decryption) are needed. This might lead to quality loss.

2.2 The Session Initiation Protocol (SIP) [12]

2.2.1 SIP message format

SIP is a text-encoded signaling protocol based on some elements inherited from HTTP [20] and SMTP [21]. SIP users are identified by a Uniform Resource Identifier (URI) [22], a universal string with a pair of domain name and user name registered for this domain (e.g., sip:ge.zhang@kau.se). SIP transactions follow a *request* and a *response* manner. Some examples of SIP requests for voice conversation are given below:

- **INVITE:** Initiates a SIP transaction to setup a session.
- **ACK:** Acknowledgement of final response to INVITE.
- **BYE:** Terminate an undergoing session.

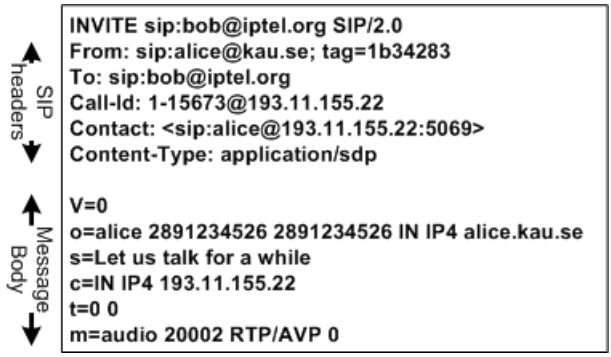

Figure 1: An example of SIP INVITE request

There are a variety of SIP response types. SIP response codes are divided into six classes by the first digit of the code:

- **1xx:** Provisional – the request has been received but the processing is unfinished.
- **2xx:** Success – the request has been received and accepted.
- **3xx:** Redirection – the request should be delivered to another place.
- **4xx:** Client error – the request cannot be processed due to error in the request.
- **5xx:** Server error – the request cannot be processed due to server's failure.
- **6xx:** Global failure – the request cannot be processed at any server.

Both SIP requests and responses are following the message format including three elements: (1) the *first line*, containing either a request method or a response code, a request URI and the SIP version; (2) the *headers*, containing a list of message headers with values for SIP transactions; (3) the *message body*, can be text-based content for different purposes. An example message for the SIP INVITE request is shown in Figure 1, indicating that the caller of "sip:alice@kau.se" invites the callee of "sip:bob@iptel.org" for a VoIP conversation. Several message headers dedicated to routing purposes are explained as follows:

- *To:* indicates the URI of the message recipient.
- *From:* indicates the URI of the message originator.
- *Contact:* indicates one or more SIP URIs of the originator by which the recipient can contact with the originator directly. They can be different from the one in the *From* header.

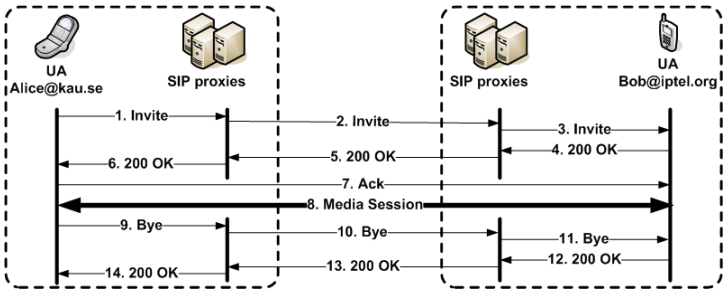

Figure 2: SIP traffic flows across 2 domains

2.2.2 SIP architecture

A SIP network consists of different entities. For simplification, we classify them into two types: *Server* and *User Agent (UA)*. A server provides services (e.g., registering users, locating users and relaying traffic) to the users within a service domain. A UA is a user's equipment connected to the networks to make or answer calls. With regard to SIP network topology, the architecture can be either Client/Server (C/S) or Peer-to-Peer (P2P), as follows:

- Client/Server (C/S): In this architecture, there are centralized servers deployed to provide different services (e.g., user location, traffic relay, session management, etc). The users rely on the servers to build conversations. In this case, a service provider can easily manage calling requests and traffic. However, the requirement on the capability of servers is high and the whole systems will be brought down in case of the failure of servers.
- Peer-to-Peer (P2P): In a P2P architecture, a user relies on other peer nodes for the provision of services. There might be still servers deployed, but the servers only provide limited functions (e.g., enable users to login to the network).

A SIP call may cross several SIP domains, an example of which is shown in Figure 2. Alice, the caller, is located in the domain kau.se and Bob, a callee, is located in iptel.org. Initially, Alice sends an INVITE request to one of the local proxies. This INVITE request indicates that she wants to talk with Bob at iptel.org. Then, the local proxy forwards this INVITE request to the remote proxy at iptel.org. The request is finally delivered to the UA of Bob. If Bob wants to accept the call, his UA will reply with a 200 OK response back through the proxies. After Alice has sent an ACK message to confirm the request, the signaling handshaking is accomplished. Thus, Alice and Bob will build a peer-to-peer voice session in which they can talk with each other by means of exchanging voice packets. Note that voice packets are not routed by SIP proxies. When Alice wants to tear down the

conversation, her UA will send a BYE request to Bob, and Bob's UA will reply with a 200 OK response. Then the call is terminated.

Sometimes SIP proxies require support from other application servers. In Figure 2, the proxies in "kau.se" domain need to resolve the domain name "iptel.org" before forwarding the INVITE message. Thus, the proxy may contact a DNS server for a recursive DNS query of "iptel.org". The message cannot continue to be processed until a DNS answer is received. It is named an *external processing* since the processing depends on a non-VoIP component. Papers I and II discuss the security risks related to external processing.

2.2.3 SIP presence message

The primary function of SIP is to deal with voice sessions, but it also has extensive functions such as conveying presence information, which is to express whether someone is available or not. When a user changes status (e.g., from "idle" to "busy"), his/her UA will multicast the up-to-date presence information to his/her buddies. In this way, a caller can get a brief overview of a callee's willingness before launching a call. Therefore, presence services provide a better quality of experience to users. Paper VII discussed the problem of information disclosure from presence traffic.

2.3 The Realtime Transport Protocol (RTP) [13]

Once a session is established, voice packets will be transmitted using a voice delivery scheme. For instance, the Realtime Transport Protocol (RTP) [13] is the IETF standardized voice delivery protocol. RTP provides end-to-end delivery services for data with real-time characteristics over IP networks. In a session, the communication partners constantly send RTP packets to each other in a fixed time interval (e.g., 20 ms). The payloads of RTP packets are encoded and decoded from analog audio signals by a codec algorithm (e.g., G.711 [23] and Speex [24]). Speech signal is sampled at 8-64k samples per second (Hz) by a user-agent. As a performance requirement, the RTP packet inter-arrival time is fixedly selected between 10 and 50 ms, with 20 ms being the common case. Given a 8kHz voice source, we have 160 samples per packet with 20 ms packets interval. In addition, RTP not only can carry payload for encoded voice, but also is used to indicate a user's keystrokes on the phone pad. It is used by Interactive Voice Response (IVR) which enables a telephone user to interact with an automatic answering machine. For example, a user can be given a menu and asked to make a choice by pressing a number button after her call is established. Paper VIII and IX discuss the problem of information disclosure on keystroke from RTP traffic.

The communication partners must have already negotiated the same codec for encoding/decoding. Two codec properties are related to Paper III, V and VI in this thesis:

- Silence suppression: It allows discontinuous voice packets transmission [25], which

is a capability to recognize the silent periods and to stop producing voice packets during these periods. Thus bandwidth can be significantly saved with little performance impact. If silence suppression is not applied, the voice packets are generated constantly with a fixed time interval (e.g., 20 ms).

- Coding bit rates: Two types of coding bit rates can be distinguished: *Fixed Bit Rate (FBR)* and *Variable Bit Rate (VBR)*. FBR codec (e.g., G.711) employs a fixed codebook with constant bit rate. Thus the generated voice packets have the same packet size. On the other hand, VBR codec (e.g., Speex) can employ an adaptive codebook with variable bit rate. It exploits the fact that some sounds are easier to represent than others. For instance, fricative sounds require lower bit rates than vowels. Thus the fricative sounds need fewer bits to be encoded to save bandwidth. In this way, UAs produce voice packets with different packet sizes.

The transmission of RTP packets is Quality of Service (QoS) sensitive with three issues that are frequently taken as criteria to evaluate:

- End-to-end delay: It indicates the time interval between encoding a voice packet at the sender and decoding it at the recipient. The delay will affect the quality of experience when it reaches a certain threshold. According to [26], users will notice a significant hesitation in their partners' response if the end-to-end delay is above 250 ms.

- Delay jitter: It refers to the variation of packet interarrival time. It is caused by network congestion and improper routing during the transmission of voice packets. As a solution, the receiver can buffer packets to eliminate delay jitter. In return, however, buffering packets introduces more end-to-end delay and impacts interactivity.

- Packet loss: Voice packets might be accidently dropped during the transmission. Retransmission of lost voice packets is not helpful since the packets are time-sensitive. Fortunately, VoIP applications can endure a certain level of packet loss. The level of endurance depends on the codec design.

3 Security requirements and mechanisms for VoIP

VoIP is applied in an open and insecure environment. Therefore, additional security enhancements for VoIP are necessary. This section introduces basic security requirements and threats to VoIP with its security mechanisms.

3.1 Security requirements for VoIP

The security requirements for SIP services are examined according to basic security components (confidentiality, integrity, availability) defined in [27].

- **Confidentiality** means that secret information should not be disclosed to unauthorized parties. When it comes to VoIP, secret information includes content of signaling and voice traffic. Further, VoIP service providers often wish to conceal their network configurations as well as to withhold user information (e.g., calling history).

- **Integrity** means that data should not be modified without authorization or in an improper manner. There are two types of integrity: *data integrity* and *data source integrity*. Regarding VoIP services, data integrity means that VoIP traffic should not be modified by unauthorized intermediaries. Data source integrity means that the source of traffic should not be an impersonated one.

- **Availability** indicates that the services should be accessible upon demand by authorized users. Considering the costs for infrastructures, it is unlikely for a VoIP service provider to offer services with unlimited capacity. Similar to other online applications, VoIP service providers deploy servers by assuming a statistic model for the future usage. However, sophisticated attackers may manipulate their use to break the assumed statistic model on purpose. In this way, legitimate users may be unable to access the service which they should get. Protection on availability aims at preventing that the statistical model is broken.

- **Privacy** refers to the human rights of individual users. More than one hundred years ago, the two US lawyers Warren and Brandeis defined that privacy is the right to be let alone [28]. Alan Westin has defined privacy as "the claim of individuals, groups and institutions to determine for themselves when how and to what extent information about them is communicated to others" [29]. Westin also believes that emerging technologies can impact on privacy. In the context of VoIP, it includes VoIP anonymity which enables users to withhold their identities when placing calls to hind who is communicating with whom. Thus an adversary cannot identify the users for a given call. Anonymous VoIP is important for many users not only for private purposes, but also journalists, human rights workers and the military would often like to keep secret with where they are communicating. Yet another privacy requirement is about users' right to refuse unsolicited calls, for instance, from telemarketers.

3.2 Classic attacks on VoIP

Classic attacks on VoIP are summarized as follows.

- **Eavesdropping:** It is a threat against confidentiality. For instance, network intermediaries can sniff VoIP traffic including voice packets and signaling messages by using some traffic capture tools (e.g., wireshark [30]). From eavesdropped traffic, an attacker may be able to read the conversation content or signaling credentials.

- **Traffic tampering:** It is a threat against data integrity. It refers to a situation where an unauthorized intermediary alters traffic. For example, given that Alice sends an INVITE request to call Bob, an intermediary can alter the first line and *To* header in the request to make Alice talk with another person instead of Bob.

- **Replay:** It is a threat against data source integrity. A replay attack refers to capturing traffic and re-sending them again after a period of time. If the traffic contains credentials, it may enable unauthorized users to access VoIP services. A replay attack that leads to financial loss of legitimate users is called billing attack [31].

- **Identity spoofing:** It is a threat against data source integrity. Attackers may initiate calling requests with fake identities to avoid being traceable or attackers may present spoofed identities to cheat callees.

- **Denial of Service:** It is a threat against availability. Denial of Service (DoS) aims at preventing legitimate users to access VoIP services or at making the services temporarily unavailable. An attacker can mount attacks on SIP services by depleting resources (e.g., CPU, memory and bandwidth) of corresponding SIP proxies [32]. As VoIP is time-sensitive, services may suffer more from DoS than other non-realtime services (e.g., email).

- **Spam:** It is a threat against privacy in the sense of the right to be alone. Spam is foreseen to appear on VoIP due to its low cost and programmable VoIP UAs. It is named as SPam over Internet Telephony (SPIT). A SPIT client could automatically launch calls to a number of VoIP users and then play a pre-recorded audio in a conversation. There might be two kinds for a SPIT, namely *online SPIT* and *offline SPIT*. In online SPIT, the callee of a SPIT is available and thus the callee needs to decide whether to answer it or not. Therefore, the online SPITs annoy users by continuously drawing their attentions. In contrast, the offline SPIT means that the callee of a SPIT is not available and cannot make an answer personally. In this case, the SPIT will be redirected to and answered by the callee's voice mailbox server. As a result, a user's voice mailbox might be filled up with junk voice messages and leaves no room for useful ones.

- **Traffic analysis:** It is a threat against privacy or confidentiality of communications, depending on whether the victim is an individual or an organization. Attackers can profile secret information (e.g., call patterns, calling records and social relations) from VoIP traffic using data mining algorithms.

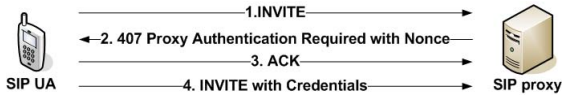

Figure 3: The HTTP Digest authentication adapted in SIP

3.3 Security mechanisms in VoIP-related RFCs

There are some security and privacy specifications that have been standardized by the IETF, classified as protections of signaling, voice traffic and underlayers respectively.

3.3.1 Security mechanisms for signaling traffic

In RFC 3261 [12], RFC 3329 [33] and RFC 4474 [34], several security mechanisms are recommended to secure SIP services. These security mechanisms are summarized as follows:

- **HTTP Digest authentication:** HTTP Digest [35], a stateless, challenge-based mechanism, provides source authentication and anti-replay protection. It can be used for both proxy-to-user authentication and user-to-user authentication. An example of this mechanism is illustrated in Figure 3. The user first sends an INVITE request to a proxy. Then, the proxy responses with a "407" message containing a unique nonce (a random number). The user receives this message and generates a hash digest for the combination of the nonce and owned password. Thus, the user sends the INVITE request again including the generated hash digest. Since the secret knowledge (the nonce and user's password) is shared with the proxy, the hash digest can be re-generated by the proxy to authenticate the source of this request. It is difficult for an eavesdropper to capture the plain text of a password since only a hash digest is transmitted over the network. Furthermore, by hashing a password with a nonce, replay attack can be efficiently prevented.

- **S/MIME:** Both message integrity and confidentiality can be ensured by carrying S/MIME [36] bodies. A signature for part of the message will be generated and attached in order to ensure that the content is not modified during the transmission. Moreover, the message payload and some headers can be encrypted to protect against eavesdropping attacks. However, the integrity of some header fields, which are allowed to be modified by intermediaries (e.g., *Via*), cannot be protected by the signature. Similarly, the fields of some message headers for routing purpose (e.g., *To*, *Via*) must be kept in plain text during transmission.

- **Inter-domain authentication:** To prevent identity fraud problems, an inter-domain authentication has been proposed in RFC 4474 [34]. Its purpose is to authenticate

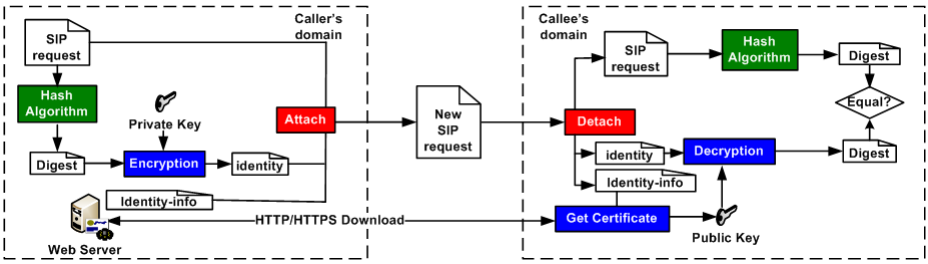

Figure 4: The mechanism of inter-domain authentication for message source

the originator of an inter-domain SIP message. The method is shown in Figure 4. For each outgoing message to other domains, the SIP proxy in the caller's domain generates a hash digest for this message. The digest is signed by the caller's proxy with its private key. The generated signature is encoded in a new header field *Identity* added to the original SIP message. Furthermore, the SIP proxy attaches another new header field *Identity-info*, which contains the Uniform Resource Locator (URL) [37] where the certificate can be fetched from. Then, this SIP request is forwarded to the callee's domain. It is regulated that each certificate must be provided by its domain itself. That is, in a SIP message, the URL indicated in the *Identity-info* field and the originator domain indicated in the *From* field should be matched. For each incoming message from other domains, the SIP proxy in the callee's domain first downloads the certificate according to the URL given in the *Identity-info* field. The public key extracted from the certificate is used to decrypt the signature contained in the *Identity* field. Then, a hash digest will be recomputed for the request. The result is used to be compared to the newly generated hash digest. The proxy will continue to process the message only if the two values are equal. This mechanism is recommended to be deployed in future SIP services to prevent SPAM [38].

3.3.2 Security mechanisms on voice traffic

- **Secure RTP (SRTP):** SRTP [39] is a common security scheme in which communication parties share keys and encrypt/decrypt payloads of RTP packets. However, RTP headers are not encrypted since RTP headers are required to be in clear text to allow for billing purposes and header compression. This means that RTP header fields are still available to intermediaries despite of the protection.

- **Zimmermann RTP (ZRTP):** ZRTP [40] is a mechanism to agree on a session key for building SRTP sessions based on the Diffie-Hellman key exchange algorithm. The scheme does not rely on PKI. Instead, the communicating parties verbally cross-check a shared value displayed at both UAs to ensure the end-to-end security.

3.3.3 Security mechanisms at lower layers

Since SIP and RTP operate on the application layer of the TCP/IP model, the security mechanisms from lower layers can also protect them.

- **TLS:** Transport Layer Security (TLS) [41], working at the transport layer, provides source authentication, message authentication and message confidentiality, based on a Public Key Infrastructure (PKI). There are three phrases to build a TLS connection between two end-points: First, two end-points negotiate for supported cryptographic algorithms. Second, two end-points exchange a symmetric key and authenticate each other. Finally, they communicate with each other using the symmetric key for encrypting messages.
- **IPSec:** IPSec [42] stands for IP Security, which is designed by IETF to provide security at the network layer using a collections of techniques (authentication header (AH), encryption security protocol (ESP) and internet key exchange (IKE)). AH provides authentication to IP packets, while ESP offers security services including both authentication and confidentiality.

4 Research questions

In this section, three underlying research questions of this thesis are outlined.

- **Question 1:** *What are potential unexplored threats in a SIP VoIP network with regard to availability, confidentiality and privacy by means of unwanted traffic and information disclosure?*

  This question is discussed in all listed papers except Paper V and X. Paper I and II proposed two kinds of DoS attacks in a SIP VoIP network, one against SIP infrastructures and another against web servers. Paper III discusses a VoIP spam threat which automatically generates SPITs using a Text-to-Speech synthesis engine. Paper IV, VI and VIII present threats to profile calling detail records, such as “who called whom”. Paper VII and IX discuss methods to profile a victim VoIP user’s social network and PIN code from traffic respectively. None of these threats has been explored in theory or in practice before.

- **Question 2:** *How far are existing security and privacy mechanisms sufficient to counteract these threats and what are their shortcomings?*

  This question is investigated in each of the papers mentioned above. The protection schemes from both IETF standardization and related research work have been discussed to see whether they can solve the problems. In addition, Paper X provides a comprehensive overview of the state of the art on anonymous VoIP communication research including threats and countermeasures.

- **Question 3:** *How can new countermeasures be designed for minimizing or preventing the consequences caused by these threats efficiently in practice?*

  The main objective of this thesis is to improve the security and privacy of current VoIP services. Therefore, it is necessary to develop novel solutions to eliminate the threats explored under Question 1 if the answer to Question 2 is negative. To realize this goal, we design, discuss and evaluate solutions for each vulnerability. The proposed countermeasures are presented in each paper.

5 Research methodology

The research methods conducted in this thesis include: *literature review*, *theoretical analysis*, *ideas development*, *quantitative experiments* and *data analysis*. The detail of each method is described in [43]. Below, we describe how we applied these research methodologies to address the questions above.

Question 1: Firstly, we studied literatures including VoIP specifications and other related papers on network security (*literature review*). We found several vulnerabilities that may exist on VoIP networks (*hypothesis formulation*). To validate our hypothesis, we adopted different methods.

- Paper I and II: We built testbeds with open source VoIP software to simulate the real VoIP networks. We also prototyped attacking tools. Since Papers I and II aim to investigate DoS attacks, we measure the attacking impact in terms of service performance. This method is defined in [44] as a *performance evaluation*. First, we measure the maximum load which a service can handle without being attacked and take this load as a benchmark. Then, we restart the measurement with attacks under different parameters to find out how much the load will be reduced. The amount of reduction reveals the attacking impact.

- Paper IV: We setup a testbed directly connected to the Internet. We collected the round trip times of several SIP requests which may trigger the testbed to download certificate from the Inernet, and then we compared on the round trip times. Note that the data we collected is public information and we did not risk or interrupt the services in the real world: Our experiments only generated a reasonable amount of well-formed http/https traffic to the Internet.

- Paper III, VIII and IX: In Paper III, we generated artificial spam traffic based on an online corpus of message spam. In paper VIII and IX, we invited around 40 students and collected their keystroke flows using pseudonyms. We applied several learning algorithms and measured the recognition accuracy.

- Paper IV and VII: We performed theoretical analysis on the attacks described in

Paper IV and VII. We employed *inductive approach* to propose formalized models for the threats. According to [44], this method is defined as *analytical modeling*.

Question 2: We reviewed RFCs, books and papers with regard to defensive solutions (*literature review*). Based on *analytical modeling* and *quantitative experiments*, we found that proposed methods are mostly insufficient to counteract the attacks that we had explored. In addition, Paper X conducted a *survey* on anonymous VoIP communications and identified the open problems as well as the research challenges.

Question 3: We applied the *evaluation research method* [43] to answer this question. We enumerated possible defensive solutions for the aforementioned threats. Then, we analyzed advantages and disadvantages of each solution theoretically (*theoretical analysis*). In Paper I and II-V, we also implemented and deployed solution prototypes in our testbed. Thus we repeated the attacks and analyzed whether the impacts can be reduced with these countermeasures (*quantitative experiments* and *data analysis*). Due to the constraints on data collection, we only discussed the countermeasures theoretically in Paper VI-IX.

6 Related work

This section summarizes research work which is related to this thesis, on unwanted traffic and information disclosure respectively, and briefly shows how this thesis has advanced state of the art.

6.1 Unwanted traffic

Denial of Service: Sisalem et al. [32] classified different types of DoS attacks targeting SIP infrastructures. They developed a taxonomy of attacks by different exploitable resources, including CPU, memory and bandwidth, to reduce the performance of a victim SIP entity. A stateful SIP proxy has to consume memory resources to keep the transaction states of unfinished SIP transactions. Therefore, stateful SIP proxies are especially vulnerable to INVITE flooding attack, in which an attacker floods SIP proxies with only INVITE messages to create a large number of broken transactions. Sengar et al. [45] proposed a machine learning method to detect INVITE flooding attacks. It is an anomaly detection system, in which they firstly characterized the distributions of INVITE and other messages in legitimate SIP traffic. They can later detect attacks if the distributions become abnormal. Moreover, [46–48] proposed specification-based detection methods using SIP state machine models to counter INVITE flooding attacks. Similarly Geneiatakis et al. [49] proposed a detection methods to prevent INVITE flooding using Bloom filter, which generates a dynamic whitelist to filter suspect traffic. Conner et al. [50] proposed a ringing-based DoS attack to consume memory resources of stateful SIP proxies. Different to INVITE flooding, this attack exploits the potential long time 180 Ringing state. Further details of DoS attacks

on SIP VoIP and preventions are available in the survey by Ehlert [51]. The DoS attacks in our research are different with previous work. One of the threats we investigated is to block SIP proxies by exploiting external infrastructures (e.g., a DNS server and a HTTP server). These risks have not been studied before. Another is to take SIP proxies as reflectors to attack HTTP servers in the network. In the context of this thesis, our countermeasures are on two levels: one is on the protocol level in which we revise the protocols to eliminate this vulnerability and another is on the implementation level in which we implement a cache scheme to enhance the performance of a SIP proxy. Our solutions are based on *prevention*, instead of *detection*. One of the benefits of *prevention* is that *prevention* can minimize the attacking impacts during the attacks. Therefore, the damage caused by attacks has been eliminated from the beginning, whereas, in most cases, *detection* works only after the damage actually happens. It is too late as the damage already occurs and easily cause detection false alarm.

Spam: A popular scheme to combat VoIP spam is to use black lists or white lists to indicate the trust level. For instance, Skype users [2] can customize their configurations to allow being called by anyone or only by the users in their buddy lists. The unclassified calls from the users in a gray-list can be temporarily rejected [38, 52]. Balasubramaniyan et al. [53] generate reputations for VoIP users based on the call durations of their previous calls. It is motivated by the observation that a legitimate user typically makes longer calls than a SPITer. Similarly, Zhang et al. [54] use cumulative online duration to calculate the reputation value. The less time a VoIP user is online, the less calls he/she can launch. This scheme prevents the SPITers who register new accounts for SPITing. A Turing test tells whether the caller is a human or an automatic spam generator. Markkola et al. [55] implemented a prototype of audio CAPTCHA [2]. It says 5 random digits and requires a caller to correctly input them for the call being processed. Quittek et al. [56] proposed a hidden Turing test based on the factor that people usually greet each other at the beginning of a telephone conversation, which results in alternative short periods of silent and speech. SPITers typically do not react to greeting, and then can be detected. Different to previous work, we examined a VoIP spam method which can morph itself to avoid being detected. Our countermeasure takes VoIP flow features to cluster spam traffic.

6.2 Information disclosure

Information hiding on signaling flows: As SIP messages are not mandatorily encrypted, Peterson [57] and Shen et al. [46] summarized privacy-sensitive message fields in SIP. The identities can be replaced by the a trusted third party with randomized pseudonyms. Karopoulos et al. [58] proposed a framework to encrypt caller and callee's identities. Unfortunately, it is not enough to only protect signaling traffic: Voice flows can also reveal secrets.

Traffic analysis attacks on voice flows: Some VoIP users make calls over commercial relays for anonymity. Wang et al. [59] demonstrated that such a solution is vulnerable: An

attacker can embed watermarks into the encrypted VoIP flow. In this way, an attacker can find out who called whom by encoding and decoding the watermarks on both side of the relay. Verscheure et al. [60] proposed an attack to reveal calling records by exploiting the human conversation pattern: When one speaks, the other usually listens. This "alternate in speaking and silence" represents a probabilistic rule of VoIP communication. Taking this into account, the caller and the callee's flows are probabilistically linkable if the attackers can detect the silence and voice period for a flow. This attack is mainly against those VoIP systems which support silence suppression. These two papers are only focused on attacks and no countermeasure solutions are given. We provided a solution to counteract against both of the attacks in Paper V, which is based on packet dropping.

Information leaking of VBR codec: Variable BitRate (VBR) codec allows the codec to change its bit rate dynamically according to the input speech signal. Thus, the user-agents generate voice packets with different sizes if they apply VBR codec. Wright et al. [61, 62] demonstrated attacks to identify the spoken language or partial conversation content of encrypted voice packets by using the packet-length information. Moreover, the packet-length information may also enable attackers to recognize the speaker [63]. Different with this work, we proposed attacks using another side-channel: the RTP header information that indicates keystroke. This side-channel may reveal a users's identity or what the user has typed.

Solution based on unlinkable identity: Munakata et al. [64] proposed a user-driven privacy mechanism by introducing Globally Routable User Agent URIs (GRUU) [65] and Traversal Using Relays around NAT (TURN) [66]. The users can obtain a SIP URI (temp GRUU) and a IP address (IP address of a TURN server) which are unlinkable to their real identities. The proposed mechanism in [64] enables VoIP users themselves to achieve anonymity by using unlinkable identities that are functional yet anonymous. However, this method does not mitigate traffic analysis as well: Intermediaries on both side of a TURN server can still profile the mapping relationship of its relayed flows.

Protection of presence traffic: Loesing et al., [67] proposed a P2P architecture based on Distributed Hash Table (DHT) with existing anonymous overlay networks (e.g., Tor [68]) to provide anonymity services for presence and message users. A user first configures his/her rendezvous address in an anonymous overlay network and later registers this address on a DHT. To extract the rendezvous address from DHT for communication, the user's buddies need to share a common knowledge with him/her. Danezis et al., [69] proposed Drac, a system designed to provide anonymity for instant message, presence and VoIP communications. Drac employs buddies as relays to connect to untrusted contacts. The communications between buddies are unobservable by applying heartbeat packets at a constant rate. However, the buddy-relationship in Drac is public. Our work is different to theirs: First, we consider a centralized presence model, rather than a P2P presence architecture; and our goal is to hide the buddy-relationship among users.

7 Contributions

The objective of the thesis was to improve the security and privacy of current VoIP services. The main contributions of Part I are summarized in Section 7.1, and the contribution of Part II are summarized in Section 7.2. We also summarize the limitations of our research.

7.1 Part I: Unwanted traffic in VoIP networks

We have defined several potential threats with unwanted traffic: (1) An attack can take advantage of the communication latency between a SIP proxy and other external infrastructures (e.g., DNS servers, Web servers) to decrease the throughput of the proxy (Paper I). We named this type of attacks as blocking attacks and a formalized model of such blocking attacks was given in our work. (2) Another kind of DoS attack is against web servers in VoIP networks by taking SIP proxies as reflectors. It takes advantage of the unbalanced traffic volume between the server side and the client side. (3) The last one is the spam problem, in which one spam can be automatically duplicated with different flow patterns while keeping similar information. This makes SPITs hard to be detected.

We also proposed countermeasures to these threats. For the first threat, we applied a cache mechanism with a priority scheme for external information processing (e.g., DNS query). For the second threat, we modified the certificate distribution scheme defined in RFC 4474 [34]. For the third threat, we employed two local-sensitive hash algorithms to cluster spam flows. Different to common hash algorithm, the local-sensitive hash algorithm generates digests with a close distance for similar inputs. Our experiments show the strength and limitations of the proposed solutions.

7.2 Part II: Information disclosure in VoIP networks

We did a comprehensive survey on anonymous VoIP communications. We surveyed some of the proposed attacks for intermediary attackers to identify the communication partners and also reviewed the existing research done to design anonymous VoIP communication services. We also discussed the major open problems in anonymous VoIP communication and possible directions for further research.

We also proposed several threats to Information disclosure. (1) We have identified and analyzed a timing attack aiming at extracting the calling history between domains. An attacker can send spoofed SIP requests to a victim proxy and observe the Round Trip Time (RTT) between the request and its response. With caches widely deployed, the RTTs for recently contacted domain should be relatively lower. Thus, the calling history of a domain can be profiled by a comparison of RTTs. We named this a SIP timing attack in Paper IV, which has not been studied before. (2) Paper VI and VIII discussed two possible methods to profile the calling records (e.g., who called whom) from wiretapped traffic. The first

takes advantage of starting and ending time of different conversations, while another takes advantage of users' keystroke patterns. (3) Paper VII proposed a threat where attackers may profile a user's social network by observing presence traffic. This attack is based on the assumption that the targets of presence traffic are the user's friends. (4) Paper IX presents a method to guess the DTMF PIN code of VoIP users from VoIP traffic. The method analyzes the time intervals between each pair of keystrokes and then uses a learning algorithm to predict the probability of a given keystroke pair.

In addition, we designed and discussed countermeasures against these threats. We applied padding methods to equalize flow patterns including packet sizes, packet interarrival time and round trip time to hide information (Paper IV, VI, VIII and IX). We also applied a VoIP "defensive dropping" method, in which some RTP packets for silence will be randomly dropped the transmission. Thus the flow patterns can be obscured to disclose information. Those countermeasures are also evaluated practically or theoretically.

7.3 Limitations

Although this thesis contributes an updated understanding of VoIP security and privacy, we note that some constraints in the research.

- We have proposed countermeasure solutions for most studied threats. Those solutions are designed for independent threats. Whether these solutions can be well integrated together has not been studied. It is possible that different countermeasure solutions interrupt each other when they are implemented in the same products. For instance, the batch solution in papers VI and VII increase the delay of signaling transmission and thus may cause a Denial of Service. However, it is difficult to have a comprehensive defending architecture to protect VoIP services against all threats. We believe that different service providers have different priorities for their security goals depending on the usage. For instance, a commercial VoIP service provider may care about the availability of their services against DoS and spam, while a military VoIP service provider may more concern on information disclosure.

- We generate artificial VoIP traffic to conduct our experiments instead of using the real VoIP traffic traces. One reason is that real VoIP traffic is confidential data and we do not want to compromise users' privacy. Another reason is that nowadays most VoIP service providers only deployed fundamental services (e.g., calling, messaging) so far. For instance, although the inter-domain authentication scheme [34] is implemented in SER [18], it still has not been widely used. Nevertheless, our goal is to locate the threats in those proposed schemes and protocols before they are widely deployed. Thus, lacking of real VoIP traffic traces does not impact our research goals.

- Papers II, VI and VII focuses more on threats and attacks, rather than countermeasures. We merely discussed possible countermeasure solutions in those papers.

Those solutions have obvious shortcomings. So far we have not found better ways to handle the threats. Thus in those paper our major contribution is studying the problems.

8 Summary of Papers

This section contains short summaries of the papers included in this thesis.

Paper I – Blocking attacks on SIP VoIP proxies caused by external processing

As VoIP applications become increasingly popular, they are more and more facing security challenges that have not been present in the traditional PSTN. One of the reasons is that VoIP applications rely heavily on external Internet-based infrastructures (e.g., DNS server, web server), so that vulnerabilities of these external infrastructures have an impact on the security of VoIP systems as well. This article presents a Denial of Service (DoS) attack on VoIP systems by exploiting long response times of external infrastructures. This attack can lead the whole VoIP system in a blocked state thus reducing the availability of its provided signalling services. The results of our experiments prove the feasibility of blocking attacks. Finally, we also discuss several defending methods and present an improved protection mechanism against blocking attacks.

Paper II – SIP Proxies: New Reflectors in the Internet

To mitigate identity theft in SIP networks, an inter-domain authentication mechanism based on certificates is proposed in RFC 4474 [34]. Unfortunately, the design of the certificate distribution in this mechanism yields some vulnerabilities. In this paper, we investigate an attack which exploits SIP infrastructures as reflectors to bring down a web server. Our experiments demonstrate that the attacks can be easily mounted. Finally, we discuss some potential methods to prevent this vulnerability.

Paper III – Detecting Near-Duplicate SPITs in Voice Mailboxes Using Hashes

Spam over Internet Telephony (SPIT) is a threat to the use of Voice of IP (VoIP) systems. One kind of SPIT can make unsolicited bulk calls to victims' voice mailboxes and then send them a prepared audio message. We detect this threat within a collaborative detection framework by comparing unknown VoIP flows with known SPIT samples since the same audio message generates VoIP flows with the same flow patterns (e.g., the sequence of

packet sizes). In practice, however, these patterns are not exactly identical: (1) a VoIP flow may be unexpectedly altered by network impairments (e.g., delay jitter and packet loss); and (2) a sophisticated SPITer may dynamically generate each flow. For example, the SPITer employs a Text-To-Speech (TTS) synthesis engine to generate a speech audio instead of using a pre-recorded one. Thus, we measure the similarity among flows using local-sensitive hash algorithms. A close distance between the hash digest of flow x and a known SPIT suggests that flow x probably belongs the same bulk of the known SPIT. Finally, we also experimentally study the detection performance of the hash algorithms.

Paper IV – Revealing the calling history of SIP VoIP systems by timing attacks

To provide high-level security assurance to SIP VoIP services, an inter-domain authentication mechanism is defined in RFC 4474. However, this mechanism introduces another vulnerability: a timing attack which can be used for effectively revealing the calling history of a group of VoIP users. The idea here is to exploit the certificate cache mechanisms supported by SIP VoIP infrastructures, in which the certificate from a caller's domain will be cached by the callee's proxy to accelerate subsequent requests. Therefore, SIP processing time varies depending whether the two domains had been into contact beforehand or not. The attacker can thus profile the calling history of a SIP domain by sending probing requests and observing the time required for processing. The result of our experiments demonstrates that this attack can be easily launched. We also discuss countermeasures to prevent such attacks.

Paper V – Peer-to-Peer VoIP Communications Using Anonymisation Overlay Networks

Nowadays, Voice over Internet Protocol (VoIP) which enables voice conversation remotely over packet switched networks gains much attentions for its low costs and flexible services. However, VoIP calling anonymity, particularly to withhold "who called whom", is difficult to achieve since VoIP infrastructures are usually deployed in an open networking environment (e.g., the Internet). Our work studies an anonymisation overlay network (AON) based solution to prevent surveillance from external attackers, who are able to wiretap the communication channels as well as to manipulate voice packets in the channels. However, it has been demonstrated that the VoIP combined with traditional AONs are vulnerable to two attacks, namely watermark attack and complementary matching attack. Taking these two attacks into account, we investigate the "defensive dropping" method in VoIP: A VoIP user-agent sends packets to an AON in a constant rate, but packets during periods of silence are marked. Then, the AON drops some silence packets and forwards the remaining ones to their destinations. The result of our experiments shows that the dropping rate must be

carefully selected to counteract both of the two attacks. Finally, we discuss further threats in terms of this solution.

Paper VI – Hidden VoIP Calling Records from Networking Intermediaries

While confidentiality of telephone conversation contents has recently received considerable attention in Internet telephony (VoIP), the protection of the caller–callee relation is largely unexplored. From the privacy research community we learn that this relation can be protected by Chaum's mixes. In early proposals of mix networks, however, it was reasonable to assume that high latency is acceptable. While the general idea has been deployed for low latency networks as well, important security measures had to be dropped for achieving performance. The result is protection against a considerably weaker adversary model in exchange for usability. In this paper, we show that it is unjustified to conclude that low latency network applications imply weak protection. On the contrary, we argue that current Internet telephony protocols provide a range of promising preconditions for adopting anonymity services with security properties similar to those of high latency anonymity networks. We expect that implementing anonymity services becomes a major challenge as customer privacy becomes one of the most important secondary goals in any (commercial) Internet application.

Paper VII – Timing Attacks on a Centralized Presence Model

Presence information (PI) represents the updated status, context and willingness of communication partners in Voice over IP systems. For instance, the action that Alice switches her status (e.g., from "idle" to "busy") will trigger PI messages to notify her buddies this change. In a centralized presence service system, presence communications are managed by a presence server based on users' buddylists. The privacy concern in this paper is that networking intermediaries, as adversaries, might be able to profile the buddy-relationship among the users by utilizing message arrival time. We found that the threat cannot be totally eliminated even if the server processes messages in batches. Attackers might observe the traffic in several rounds and thus profile the results. In this paper, we introduce the attacks and discuss potential countermeasures.

Paper VIII – Analyzing Key-click Patterns of PIN Input for Recognizing VoIP Users

Malicious intermediaries are able to detect the availability of VoIP conversation flows in a network and observe the IP addresses used by the conversation partners. However, it is insufficient to infer the calling records of a particular user in this way since the linkability

between a user and a IP address is uncertain: users may regularly change or share IP addresses. Unfortunately, VoIP flows may contain human-specific features. For example, users sometimes are required to provide Personal identification numbers (PINs) to a voice server for authentication and thus the key-click patterns of entering a PIN can be extracted from VoIP flows for user recognition. We invited 31 subjects to enter 4-digital PINs on a virtual keypad of a popular VoIP user-agent with mouse clicking. Employing machine learning algorithms, we achieved average equal error rates of 10-29% for user verification and a hitting rate up to 65% with a false positive rate around 1% for user classification.

Paper IX – Timing attacks on PIN input in VoIP networks (Short paper)

To access automated voice services, Voice over IP (VoIP) users sometimes are required to provide their Personal Identification Numbers (PIN) for authentication. Therefore when they enter PINs, their user-agents generate packets for each key pressed and send them immediately over the networks. This paper shows that a malicious intermediary can recover the inter-keystroke time delay for each PIN input even if the standard encryption mechanism has been applied. The inter-keystroke delay can leak information of what has been typed: Our experiments show that the average search space of a brute force attack on PIN can be reduced by around 80%.

Paper X – A survey on anonymous Voice over IP communication: Attacks and defences

Anonymous Voice over IP (VoIP) communication is important for many users, in particular, journalists, human rights workers and the military. Recent research work has shown an increasingly interest in methods of anonymous VoIP communication. This survey starts by introducing and identifying the major concepts and challenges in this field. Then we review anonymity attacks on VoIP and the existing work done to design defending strategies. Finally, we discuss possible directions for the future work in this field.

9 Conclusions and Outlook

Considering the increasing degree of VoIP deployment, we claim that it is crucial for the Internet community including venders and service providers to fully understand the potential security and privacy vulnerabilities of VoIP. A deeper understanding supports design of countermeasures that can be incorporated into the protocol specifications and products. In this thesis, we listed several potential security and privacy vulnerabilities, especially by means of unwanted traffic and information disclosure. Moreover, we have studied the vulnerabilities and countermeasures by conducted experiments.

This final section shows overall conclusions and lessons learned. We also suggest future research possibilities. Some of the main conclusions of this thesis are:

- *Achieving security and privacy for VoIP is more difficult than for traditional PSTN.* VoIP infrastructures are deployed in a relatively open environment (e.g., the Internet). It is easy for potential attackers to access these infrastructures for launching attacks. In contrast, PSTN, a closed network using independent communication protocols, requires more cost for attackers to access. Moreover, since VoIP services heavily rely on assistance from external servers (e.g., DNS server, web server), the communication between VoIP servers and these external servers can be exploited to impact the confidentiality and availability of VoIP users. Such risks have never been reported in PSTN since it does not employs shared infrastructures.

- *Unbalanced resource consumption between client side and server side leads to DoS attacks.* In a client/server SIP architecture, a SIP proxy processes and forwards SIP messages between users. However, a message may consume little resource (e.g., bandwidth, CPU time, etc) on client side but heavy resource on server side under some circumstances. An attacker can thus take advantage of this fact to continuously attack the server side by depleting its resources.

- *It is more difficult to classify voice spam than text spam.* Classification of spam needs machine learning algorithms and content recognition. Currently, the technique of processing text is more mature than that of processing of voice. Thus it is still rather challenging to manage VoIP spam.

- *Side-channels in VoIP traffic disclose secret information.* Depending on the design, different side-channels may be present in VoIP traffic. For instance, the round trip time of a SIP message can reveal the content of local cache which can be used to guess the calling records between two SIP service domains. Also, header fields, packet sizes and packet interarrival time may disclose a user's identity, PIN code and social networks. To break side-channels, these traffic patterns must be equalized or distorted.

- *Security products for VoIP should be designed taking efficiency into account.* In contrast to other services (e.g., email, web), VoIP is time-sensitive. Therefore, complicated and time consuming security mechanisms are not suitable to apply for VoIP. If a designer fails to consider efficiency, attackers can easily manipulate the performance of SIP services to launch a Denial of Service attack.

Current VoIP and Internet technology are still being developed. New functions, features and implementations are likely to introduce further vulnerabilities. Hence in future it will be important to continuously conduct vulnerability assessment on the VoIP protocols and implementations. Another interesting future step is to improve defending alternatives for those found vulnerabilities. There are limitations on our proposed countermeasures,

which may work better if integrated with other alternatives (defense in depth). Moreover, there is no implementation for anonymous VoIP so far, although there are some implementations for low-latency anonymous communications (e.g., Tor [68], An.On [70]) which are however not scalable for VoIP applications. Especially, an optimization scheme to enhance anonymity with an acceptable communication quality is a necessity. Future work also should include implementation and testing of proposed defensive architectures in practice.

References

[1] U. Varshney, A. Snow, M. McGivern, and C. Howard. Voice over IP. *Commun. ACM*, 45(1):89–96, 2002.

[2] Skype. http://www.Skype.com, visited at 11th-Oct-2011.

[3] VoIP Security Alert: Hackers Start Attacking For Cash. http://www.informationweek.com/news/188702963, visited at 16th-Jan-2012.

[4] D. Butcher, X. Li, and J. Guo. Security challenge and defense in VoIP infrastructures. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, 37(6):1152–1162, 2007.

[5] T.J. Walsh and D.R. Kuhn. Challenges in securing Voice over IP. *Security & Privacy, IEEE*, 3(3):44–49, 2005.

[6] S. M. Bellovin, M. Blaze, and S. Landau. The real national-security needs for VoIP. *Commun. ACM*, 48(11):120, 2005.

[7] D. C. Sicker and T. Lookabaugh. VoIP security: Not an afterthought. *Queue*, 2(6):56–64, 2004.

[8] E. A. Blake. Network security: VoIP security on data network–a guide. In *InfoSecCD '07: Proceedings of the 4th annual conference on Information security curriculum development*, pages 1–7, New York, NY, USA, 2007. ACM.

[9] P. Park. *Voice over IP Security*. Cisco Press, 2009.

[10] C. J. Dawson. http://sip-trunking.tmcnet.com/topics/sip-trunking/articles/67855-report-sip-trunking-increasing-popularity.htm, visited at 16th-Jan-2009.

[11] J. Postel. Internet Protocol (IP), 1981. RFC 791.

[12] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, and E. Schooler. SIP: Session Initiation Protocol, 2002. RFC 3261.

[13] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson. RTP: A transport protocol for real-time applications, 2003. RFC 3550.

[14] The Internet Engineering Task Force (IETF). http://www.ietf.org/, visited at 16th-Feb-2012.

[15] Kphone. http://sourceforge.net/projects/kphone, visited at 16th-Feb-2009.

[16] Kcall. http://www.basyskom.de/index.pl/kcall, visited at 16th-Feb-2009.

[17] X-Lite. http://www.counterpath.net/x-lite.html, visited at 16th-Feb-2009.

[18] SIP Express Router. http://www.iptel.org, visited at 16th-Sep-2008.

[19] OpenSIPS. http://www.opensips.org/, visited at 16th-Feb-2009.

[20] R. Fielding, J. Gettys, J. Mogul, H. Frystyk, L. Masinter, P. Leach, and T. Berners-Lee. Hypertext Transfer Protocol – HTTP/1.1, 1999. RFC 2616.

[21] J. B. Postel. Simple Mail Transfer Protocol (SMTP), 1982. RFC 821.

[22] T. Berners-Lee, R. Fielding, and L. Masinter. Uniform Resource Identifier (URI): Generic syntax, 2005. RFC 3986.

[23] G.711. http://www.itu.int/rec/T-REC-G.711/e, visited at 21th-Oct-2011.

[24] Spexx. http://www.speex.org/, visited at 11th-Oct-2011.

[25] R. Zopf. Real-time Transport Protocol (RTP) Payload for Comfort Noise (CN), 2002. RFC 3389.

[26] Recommendation G.114 - One-way Transmission Time. http://www.itu.int/itudoc/itu-t/aap/sg12aap/history/g.114/index.html, visited at 21th-Oct-2011.

[27] M. Bishop. *Introduction to computer security*. Addison-Wesley, 2005.

[28] S. D. Warren and L. D. Brandeis. The Right to Privacy. *Harvard Law Review*, 4(5), 1890.

[29] A. Westin. *Privacy and Freedom*. Atheneum, 1968.

[30] Wireshark. http://www.wireshark.org/, visited at 16th-Feb-2009.

[31] R. Zhang, X. Wang, X. Yang, and X. Jiang. Billing attacks on SIP-based VoIP systems. In *WOOT '07: Proceedings of the first USENIX workshop on Offensive Technologies*, pages 1–8, Berkeley, CA, USA, 2007. USENIX Association.

[32] D. Sisalem, J. Kuthan, and S. Ehlert. Denial of service attacks targeting a SIP VoIP infrastructure: attack scenarios and prevention mechanisms. *IEEE Network*, 20(5):26–31, 2006.

[33] J. Arkko, V. Torvinen, G. Camarillo, A. Niemi, and T. Haukka. Security mechanism agreement for the Session Initiation Protocol (SIP), 2003. RFC 3329.

[34] J. Peterson and C. Jennings. Enhancements for authenticated identity management in the Session Initiation Protocol (SIP), 2006. RFC 4474.

[35] J. Franks, P. Hallam-Baker, J. Hostetler, S. Lawrence, P. Leach, A. Luotonen, and L. Stewart. HTTP authentication: Basic and digest access authentication, 1999. RFC 2616.

[36] B. Ramsdell. Secure/Multipurpose Internet Mail Extensions (S/MIME) version 3.1 message specification, 2004. RFC 3851.

[37] T. Berners-Lee, L. Masinter, and M. McCahill. Uniform Resource Locators (URL), 1994. RFC 1738.

[38] J. Rosenberg and C. Jennings. The Session Initiation Protocol (SIP) and Spam, 2008. RFC 5039.

[39] M. Baugher, D. McGrew, M. Naslund, E. Carrara, and K. Norrman. The Secure Real-time Transport Protocol (SRTP), 2004. RFC 3711.

[40] P. Zimmermann, A. Johnston, and J. Callas. ZRTP: Media Path Key Agreement for Unicast Secure RTP, 2011. RFC 6189.

[41] T. Dierks and E. Rescorla. The Transport Layer Security (TLS) protocol version 1.2, 2008. RFC 5246.

[42] S. Kent and R. Atkinson. Security architecture for the Internet Protocol, 1998. RFC 1825.

[43] C. Robson. *Real world research*. Blackwell Publishing, 2002.

[44] R. Jain. *The art of computer systems performance analysis: Techniques for experimental design, measurement, simulation, and modeling*. Wiley- Interscience, 1991.

[45] H. Sengar, D. Wijesekera, H. Wang, and S. Jajodia. Fast detection of denial of service attacks on IP telephone. In *14th IEEE International Workshop on Quality of Service*, New Haven, USA, June 2006. IEEE.

[46] E. Y. Chen. Detecting DoS attacks on SIP system. In *1st IEEE Workshop on VoIP Management and Security*, Vancouver, Canada, April 2006. IEEE.

[47] H. Sengar, D. Wijesekera, H. Wang, and S. Jajodia. VoIP intrusion detection through Interacting protocol state machines. In *DSN '06: the International Conference on Dependable Systems and Networks*, pages 393–402, Washington, USA, June 2006. IEEE Computer Society.

[48] S. Ehlert, C. Wang, T. Magedanz, and D. Sisalem. Specification-based denial-of-service detection for SIP Voice-over-IP networks. In *3rd International Conference on Internet Monitoring and Protection*, Bucharest, Hungary, July 2008. IEEE.

[49] N. Vrakas D. Geneiatakis and C. Lamrinoudakis. Utilizing Bloom Filters for Detecting Flooding Attacks against SIP Based Services. *Computer and Security*, 28(7):578–591, 2009.

[50] W. Conner and K. Nahrstedt. Protecting SIP proxy servers from ringing-based denial-of-service attacks. In *the Tenth IEEE International Symposium on Multimedia (ISM)*, Berkeley, USA, December 2008. IEEE.

[51] S. Ehlert, D. Geneiatakis, and T. Magendaz. Survey of Network Security Systems to Counter SIP-based Denial-of-Service Attacks. *Computer and Security*, 29(2):225–243, 2010.

[52] D. Shin, J. Ahn, and C. Shim. Progressive multi gray-leveling: a voice spam protection algorithm. *IEEE networks*, 20(5):18 – 24, 2006.

[53] V. A. Balasubramaniyan, M. Ahamad, and H. Park. Callrank: Using call duration, social networks and pagerank to combat SPIT. In *Proceedings of CEAS '07*. ACM.

[54] R. Zhang and A. Gurtov. Collaborative reputation-based voice spam filtering. In *Proceedings of DEXA workshop '09*. IEEE Computer Society.

[55] A. Markkola and J. Lindqvist. Accessible voice CAPTCHAs for internet telephony. In *Proceedings of SOAPS '08*. ACM.

[56] J. Quittek, S. Niccolini, S. Tartarelli, M. Stiemerling, M. Brunner, and T. Ewald. Detecting SPIT calls by checking human communication patterns. In *Proceedings of ICC '07*. IEEE Communication Society.

[57] J. Peterson. A Privacy Mechanism for the Session Initiation Protocol (SIP), 2002. RFC 3323.

[58] G. Karopoulos, G. Kambourakis, S. Gritzalis, and E. Konstantinou. A framework for identity privacy in SIP. *Journal of Network and Computer Applications*, 33:16–28, January 2010.

[59] X. Wang, S. Chen, and S. Jajodia. Tracking anonymous peer-to-peer VoIP calls on the Internet. In *CCS '05: Proceedings of the 12th ACM conference on Computer and communications security*, pages 81–91, New York, NY, USA, 2005. ACM.

[60] O. Verscheure, M. Vlachos, A. Anagnostopoulos, P. Frossard, E. Bouillet, and P. S. Yu. Finding "who is talking to whom" in VoIP networks via progressive stream clustering. In *ICDM '06: Proceedings of the* 6^{th} *International Conference on Data Mining*, pages 667–677, Washington, DC, USA, 2006. IEEE Computer Society.

[61] C. V. Wright, L. Ballard, F. Monrose, and G. M. Masson. Language identification of encrypted VoIP traffic: Alejandra y roberto or alice and bob? In *SS'07: Proceedings of* 16^{th} *USENIX Security Symposium on USENIX Security Symposium*, pages 1–12, Berkeley, CA, USA, 2007. USENIX Association.

[62] C. V. Wright, L. Ballard, S. E. Coull, F. Monrose, and G. M. Masson. Spot me if you can: Uncovering spoken phrases in encrypted VoIP conversations. In *SP '08: Proceedings of the 2008 IEEE Symposium on Security and Privacy*, pages 35–49, Washington, DC, USA, 2008. IEEE Computer Society.

[63] L.A. Khan, M.S. Baig, and A. M. Youssef. Speaker Recognition from Encrypted VoIP Communications. *Digital Investigation*, 7(1-2):65 – 73, 2010.

[64] M. Munakata, S. Schubert, and T. Ohba. User-agent-driven privacy mechanism for SIP, 2010. RFC 5767.

[65] J. Rosenberg. Obtaining and using globally routable user agent uris (GRUUs) in the session initiation protocol (SIP), 2009. RFC 5627.

[66] R. Mahy, P. Matthews, and J. Rosenberg. Traversal using relays around NAT (TURN): Relay extensions to session traversal utilities for NAT (STUN), 2010. RFC 5766.

[67] K. Loesing, M. Dorsch, M. Grote, K. Hildebrandt, M. Roglinger, M. Sehr, C. Wilms, and G. Wirtz. Privacy-aware presence management in instant message systems. In *Proceedings of the* 20^{th} *Parallel and Distributed Processing Symposium, 2006. (IPDPS '06)*. IEEE, 2006.

[68] R. Dingledine, N. Mathewson, and P. Syverson. Tor: the second-generation onion router. In *SSYM'04: Proceedings of the 13th conference on USENIX Security Symposium*, pages 21–21, Berkeley, CA, USA, 2004. USENIX Association.

[69] G. Danezis, C. Diaz, C. Troncoso, and B. Laurie. Drac: An architecture for anonymous low-volume communications. In *PETS '10: Proceedings of the 10th Privacy Enhancing Technologies Symposium*, pages 202–219, Berlin, Heidelberg, 2010. Springer-Verlag.

[70] An.On. http://anon.inf.tu-dresden.de/index_en.html, visited at 10th-May-2011.