A Comprehensive Survey of Voice over IP Security Research

Angelos D. Keromytis, *Senior Member, IEEE*

***Abstract*—We present a comprehensive survey of Voice over IP security academic research, using a set of 245 publications forming a closed cross-citation set. We classify these papers according to an extended version of the VoIP Security Alliance (VoIPSA) Threat Taxonomy. Our goal is to provide a roadmap for researchers seeking to understand existing capabilities and to identify gaps in addressing the numerous threats and vulnerabilities present in VoIP systems. We discuss the implications of our findings with respect to vulnerabilities reported in a variety of VoIP products.**

We identify two specific problem areas (denial of service, and service abuse) as requiring significant more attention from the research community. We also find that the overwhelming majority of the surveyed work takes a black box view of VoIP systems that avoids examining their internal structure and implementation. Such an approach may miss the mark in terms of addressing the main sources of vulnerabilities, *i.e.,* implementation bugs and misconfigurations. Finally, we argue for further work on understanding cross-protocol and cross-mechanism vulnerabilities (emergent properties), which are the byproduct of a highly complex system-of-systems and an indication of the issues in future large-scale systems.

***Index Terms*—VoIP, SIP, security**

I. Introduction

VoIP refers to a class of products that enable advanced communication services over data networks. While voice is a key aspect in such products, video and other capabilities (*e.g.,* collaborative editing and whiteboard sharing, file sharing, calendaring) are supported. The key advantages of VoIP are flexibility and low cost. The former derives from the (generally) open architectures and software-based implementation, while the latter is due to new business models, equipment and network-link consolidation, and ubiquitous consumer-grade broadband connectivity.

Due to these benefits, VoIP has seen rapid uptake in both the enterprise and consumer markets. An increasing number of enterprises are replacing their internal phone switches with VoIP-based implementations, both to introduce new features and to eliminate redundant equipment. Consumers have embraced a slew of technologies with different features and costs, including P2P calling (Skype), Internet-to-PSTN network bridging, and wireless VoIP. These new technologies and business models are being promoted by a new generation of startup companies that are challenging the traditional status quo in telephony and personal telecommunications. As a result, a number of PSTN providers have already completed or are in the process of transitioning from circuit-switched networks to VoIP-friendly packet-switched backbones. Finally, as the commercial and consumer sectors go, so do governments and militaries due to cost reduction concerns and the general dependence on Commercial Off The Shelf (COTS) equipment for the majority of their computing needs.

Because of the need to seamlessly interoperate with the existing telephony infrastructure, the new features, and the speed of development and deployment, VoIP protocols and products have been repeatedly found to contain numerous vulnerabilities [1], [2], [3] that have been exploited [4], [5], [6]. As a result, a fair amount of research has been directed towards addressing some of these issues. However, the effort is unbalanced, with little effort spent on some highly deserving problem areas.

This comprehensive survey covers 245 VoIP security research papers and books, complementing our previous work that analyzed known vulnerabilities [1], [2], [3]. Our primary goal is to create a roadmap of existing work in securing VoIP, towards reducing the start-up effort required by other researchers to initiate research in this space. A secondary goal is to identify gaps in existing research, and to help inform the security community of challenges and opportunities for further work. Finally, in the context of the VAMPIRE project[1] we seek to provide guidance as to what further work in needed to better understand and analyze the activities of attackers. We classify these papers according to the class of threat they seek to address, using an extended version of the VoIP Security Alliance (VoIPSA) [7] threat taxonomy. We discuss our findings, and contrast them with our previous survey on VoIP vulnerabilities.

Paper Organization: Section II provides a brief overview of SIP, perhaps the most popular VoIP technology currently in use. Section III summarizes the threat taxonomy as defined by the VoIP Security Alliance. Our survey of the research literature is given in Section IV. We then discuss our findings in Section V.

II. SIP Overview

We focus our attention on Session Initiation Protocol (SIP) [8], a popular and widely deployed technology. Most research has focused on SIP, primarily because of its wide use and the

Author's address: Angelos D. Keromytis, Department of Computer Science, Mail Code 0401, Columbia University in the City of New York, New York, NY 10027, USA.

The bulk of this work was conducted while the author was on sabbatical leave with Symantec Research Labs, France. This work was supported in part by the French National Research Agency (ANR) under Contract ANR-08-VERS-017 and by the US National Science Foundation under Grant CNS-09-14845.

[1] http://vampire.gforge.inria.fr/

availability of a number of free and open-source implementations.

SIP is a protocol standardized by the Internet Engineering Task Force (IETF), and is designed to support the setup of bidirectional communication sessions including, but not limited to, VoIP calls. It is similar in some ways to HTTP, in that it is text-based, has a request-response structure, and even uses a mechanism based on the HTTP Digest Authentication [9] for user authentication. However, it is an inherently stateful protocol that supports interaction with multiple network components (*e.g.*, middleboxes such as PSTN bridges), and asynchronous notifications. While its finite state machine is seemingly simple, in practice it has become quite large and complicated — an observation supported by the fact that the main SIP RFC [8] is one of the longest ever defined, with additional RFCs further extending the specification.

SIP is a signaling protocol, relying on RTP [10] for media transfer. There exists an RTP profile (named Secure RTP, or SRTP [11]) that supports encryption and integrity, but is not yet widely used. The RTP protocol family also includes RTCP, which is used to control certain RTP parameters between communicating endpoints.

SIP can operate over a number of transport protocols, including TCP [12], UDP [13] and SCTP [14]. UDP is generally the preferred method due to simplicity and performance, although TCP has the advantage of supporting TLS protection of call setup. However, recent work on Datagram TLS (DTLS) [15] may render this irrelevant. SCTP, on the other hand, offers several advantages over both TCP and UDP, including DoS resistance [16], multi-homing and mobility support, and logical connection multiplexing over a single channel.

In the SIP architecture, the main entities are endpoints (whether softphones or physical devices), a proxy server, a registrar, a redirect server, and a location server. Figure 1 shows a high-level view of the SIP entity interactions. The registrar, proxy and redirect servers may be combined, or they may be separate entities operated independently. Endpoints communicate with a registrar to indicate their presence. This information is stored in the location server. A user may be registered via multiple endpoints simultaneously.

During call setup, the endpoint communicates with the proxy which uses the location server to determine where the call should be routed to. This may be another endpoint in the same network (*e.g.*, within the same enterprise), or another proxy server in another network. Alternatively, endpoints may use a redirect server to directly determine where a call should be directed to; redirect servers consult with the location server in the same way that proxy servers operate during call setup. Once an end-to-end channel has been established (through one or more proxies) between the two endpoints, SIP negotiates the actual session parameters (such as the codecs, RTP ports, *etc.*) using the Session Description Protocol (SDP) [17].

Figure 2 shows the message exchanges during a two-party call setup. Alice sends an INVITE message to the proxy server, optionally containing session parameter information encoded within SDP. The proxy forwards this message directly to Bob, if Alice and Bob are users of the same domain. If Bob is registered in a different domain, the message will be relayed

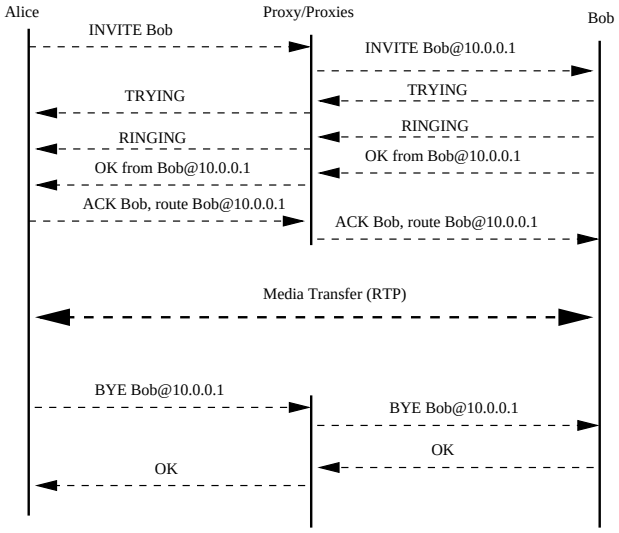

Fig. 2. Message exchanges during a SIP-based two-party call setup.

to Bob's proxy, and from there to Bob. Note that the message may be forwarded to multiple endpoints, if Bob is registered from multiple locations. While these are ringing (or otherwise indicating that a call setup is being requested), RINGING messages are sent back to Alice. Once the call has been accepted, an OK message is sent to Alice, containing his preferred parameters encoded within SDP. Alice responds with an ACK message. Alice's session parameter preferences may be encoded in the INVITE or the ACK message.

Following this exchange, the two endpoints can begin transmitting voice, video or other content (as negotiated) using the agreed-upon media transport protocol, typically RTP. While the signaling traffic may be relayed through a number of SIP proxies, the media traffic is exchanged directly between the two endpoints. When bridging different networks, *e.g.*, PSTN and SIP, media gateways may disrupt the end-to-end nature of the media transfer. These entities translate content (*e.g.*, audio) between the formats that are supported by the different networks.

Because signaling and media transfer operate independently, the endpoints are responsible for indicating to the proxies that the call has been terminated, using a BYE message which is relayed through the proxies along the same path as the call setup messages.

There are many other protocol interactions supported by SIP, that cover many common (and uncommon) scenarios including call forwarding (manual or automatic), conference calling, voicemail, *etc.* Typically, this is done by semantically overloading SIP messages such that they can play various roles in different parts of the call. We shall see in Section III examples of how this flexibility and protocol modularity can be used to attack the system.

SIP traffic is typically transmitted over port 5060 (UDP or TCP), although the port can vary based on configuration parameters. The ports used for the media traffic, however, are dynamic and negotiated via SDP during call setup. This poses some problems when Network Address Translation (NAT) or

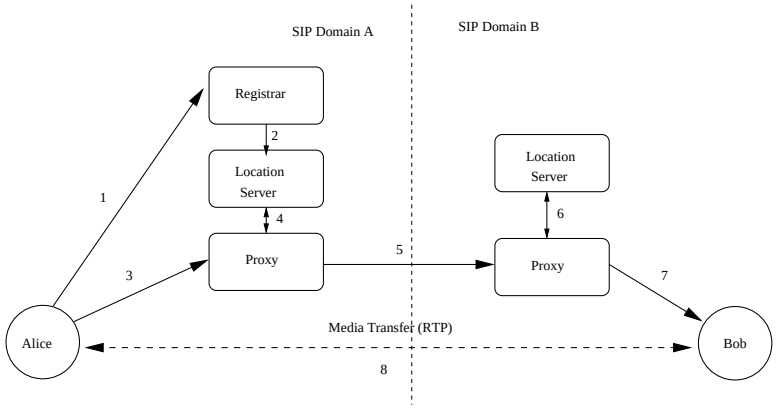

Fig. 1. Session Initiation Protocol (SIP) entity interactions. User Alice registers with her domain's Registrar (1), which stores the information in the Location Server (2). When placing a call, Alice contacts her local Proxy Server (3), which may consult the Location Server (4). A call may be forwarded to another Proxy Server (5), which will consult its domain Location Server (6) before forwarding the call to the final recipient. After the SIP negotiation terminates, RTP is used directly between Alice and Bob to transfer media content. For simplicity, this diagram does not show the possible interaction between Alice and a Redirection Server (which would, in turn, interact with the Location Server).

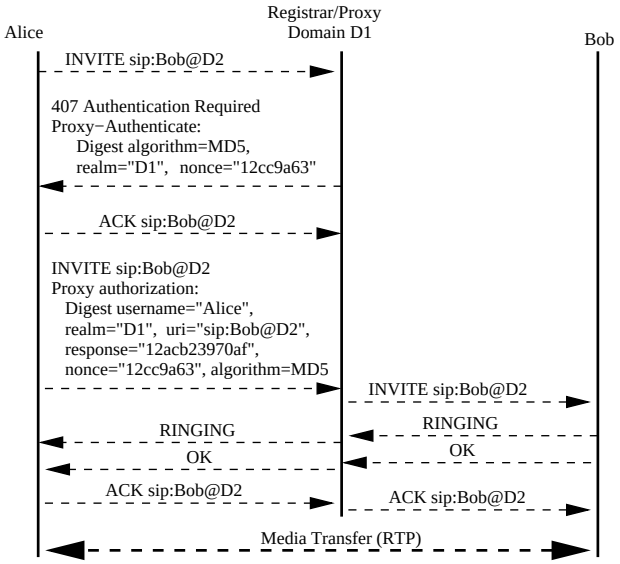

Fig. 3. SIP Digest Authentication

firewalls are traversed. Typically, these have to be stateful and understand the SIP exchanges so that they can open the appropriate RTP ports for the media transfer. In the case of NAT traversal, endpoints may use protocols like STUN to enable communication. Alternatively, the Universal Plug-and-Play (uPnP) protocol [2] may be used in some environments, such as residential broadband networks consisting of a single subnet behind a NAT gateway.

For authenticating endpoints, the registrar and the proxy typically use HTTP Digest Authentication, as shown in Figure 3. This is a simple challenge-response protocol that uses a shared secret key along with a username, domain name, a nonce, and specific fields from the SIP message to compute a cryptographic hash. Passwords are not transmitted in plaintext form over the network. It is worth noting that authentication may be requested at almost any point during a call setup. We shall later see an example where this can be abused by a malicious party to conduct toll fraud in some environments.

For more complex authentication scenarios, SIP can use S/MIME encapsulation [18] to carry complex payloads, including public keys and certificates. When TCP is used as the transport protocol for SIP, TLS can be used to protect the SIP messages. TLS is required for communication among proxies, registrars and redirect servers, but only recommended between endpoints and proxies or registrars. Alternatively, IPsec [19] may be used to protect all communications, regardless of the transport protocol. However, because few implementations integrate SIP, RTP and IPsec, it is left to system administrators to setup and manage such configurations.

III. VoIP Threat Classification

To classify the surveyed work, we use the taxonomy provided by the Voice over IP Security Alliance (VoIPSA)[3]. VoIPSA is a vendor-neutral, not for profit organization composed of VoIP and security vendors, organizations and individuals with an interest in securing VoIP protocols, products and installations. The VoIPSA security threat taxonomy [7] aims to define the security threats against VoIP deployments, services, and end users. The key elements of this taxonomy are:

1) **Social threats** are aimed directly against humans. For example, misconfigurations, bugs or bad protocol interactions in VoIP systems may enable or facilitate attacks that misrepresent the identity of malicious parties to users. Such attacks may then act as stepping stones to further attacks such as phishing, theft of service, or unwanted contact (spam).
2) **Eavesdropping, interception, and modification threats** cover situations where an adversary can unlawfully and without authorization from the parties concerned listen in on the signaling (call setup) or the

[2]http://www.upnp.org/

[3]http://www.voipsa.org/

content of a VoIP session, and possibly modify aspects of that session while avoiding detection. Examples of such attacks include call re-routing and interception of unencrypted RTP sessions.

3) **Denial of service threats** have the potential to deny users access to VoIP services. This may be particularly problematic in the case of emergencies, or when a DoS attack affects all of a user's or organization's communication capabilities (*i.e.*, when all VoIP and data communications are multiplexed over the same network which can be targeted through a DoS attack). Such attacks may be VoIP-specific (exploiting flaws in the call setup or the implementation of services), or VoIP-agnostic (*e.g.*, generic traffic flooding attacks). They may also involve attacks with physical components (*e.g.*, physically disconnecting or severing a cable) or through computing or other infrastructures (*e.g.*, disabling the DNS server, or shutting down power).
4) **Service abuse threats** covers the improper use of VoIP services, especially (but not exclusively) in those situations where such services are offered in a commercial setting. Examples of such threats include toll fraud and billing avoidance [5], [6].
5) **Physical access threats** refer to inappropriate/unauthorized physical access to VoIP equipment, or to the physical layer of the network (following the ISO 7-layer network stack model).
6) **Interruption of services threats** refer to non-intentional problems that may nonetheless cause VoIP services to become unusable or inaccessible. Examples of such threats include loss of power due to inclement weather, resource exhaustion due to over-subscription, and performance issues that degrade call quality.

In our discussion and classification of related work that follows, we focus on the first four elements of the taxonomy, since the last two are largely outside the scope of computer security research. In addition to these four categories, we also use the following:

- **Overviews and Surveys** covers work that does not offer any original technical research, but rather summarize attacks and defenses in VoIP. While valuable in helping understand the problem space, such works are generally (but not always) fairly narrow in scope and do not typically suggest solutions to the problems surveyed; at best, they summarize existing/known techniques and mechanisms for mitigating those problems.
- **Field Studies and System/Protocol Analysis** covers work that analyzes software, protocols, and systems using a variety of techniques.
- **Performance Analysis** covers work that measures the performance impact of security mechanisms, both on call setup (authentication costs) and on media transfer.
- **Authentication Protocols** covers work that proposes extensions or variants of authentication mechanisms and algorithms with SIP. Typically, these papers have a strong cryptographic element, with VoIP used primarily as a motivating environment.
- **Architectures** covers work that defines cross-cutting approaches to secure VoIP. In the surveyed work, a significant portion of these papers revolves around intrusion detection systems.
- **Middleboxes** covers work that describes new firewall architectures or mechanisms for enabling VoIP to work with the current generation of firewalls.
- **Intrusion Detection** covers other intrusion and anomaly detection work that could not be easily classified in any of the previous categories.
- **Miscellaneous** includes other work that does not fit in the remainder of the classification.

Figure 4 graphically depicts our overall classification scheme, annotated with the number of items in each category.

IV. Survey of VoIP Security Research

A. Collection Methodology

We used a structured approach to compiling the list of papers. While we do not claim to have all VoIP security papers, we have identified as many as was possible using our methodology. The process we used was:

- Compile an initial collection of papers, based on:
  - Personal knowledge (direct and through recommendation) of specific papers.
  - Searches on CiteSeer, IEEE Xplore, ACM Digital Library and Google Scholar (keywords used were "VoIP security", "SIP security", "VoIP vulnerabilities", "SIP vulnerabilities", "SIP attacks", "VoIP attacks").
  - Browsing the proceedings of top security conferences and journals (IEEE Security & Privacy Symposium, ISOC Symposium on Network and Distributed Systems Security, ACM Computer and Communications Security, USENIX Security, RAID, ACM Transactions on Information and Systems Security, IEEE Transactions on Dependable and Secure Computing), and some area-specific workshops (*e.g.*, VoIP Security Workshop) from the past 5 years.
- Expand this selection by:
  - Fetching all relevant cited papers not already in the collection.
  - Browsing the proceedings of conferences or journals in which these cited papers appeared, to identify other relevant papers.
  - Searching for other VoIP security papers by the authors of these cited papers.
- Iterate until no new papers are added to the collection.

In addition, a few more papers were suggested by anonymous reviewers, as part of the review process for this paper. The same algorithm was used, and expanded the 3 initial suggestions to 12 total additional papers. In the process, we discovered a case of plagiarism for papers that were published 6 years apart; we notified the authors and journal editors involved.

In order to avoid a lopsided distribution of papers (and an infinite expansion), we did not include in the collection

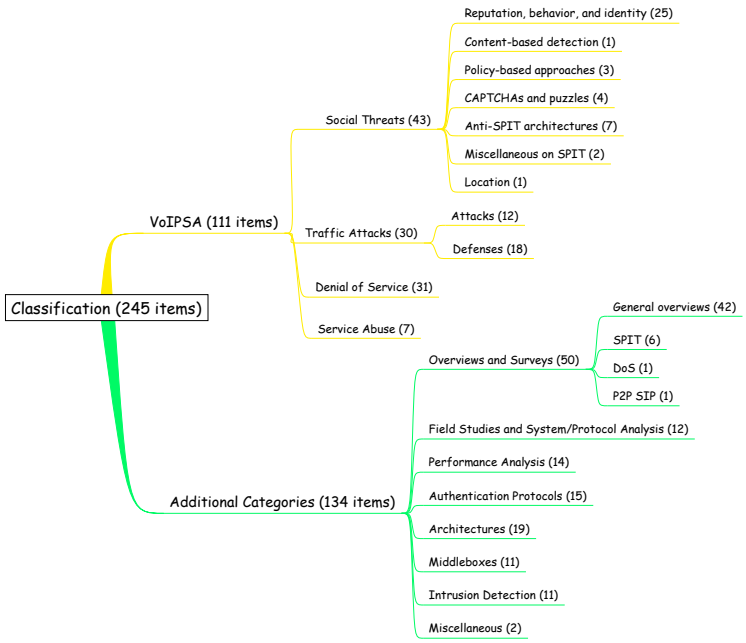

Fig. 4. Classification tree

papers that were deemed of only peripheral relevance to VoIP or VoIP security. The result of this process (modulo any papers inadvertently missed) was 245 publications.

In the following two sections, we discuss the related work using the extended VoIPSA taxonomy, as described in Section III. For each classification area, we give the paper count as a crude indication of the level of activity.

B. VoIPSA-based Classification (111 items)

We now discuss the work that fits naturally within the first four categories of the VoIPSA taxonomy, which constitutes 45% of the surveyed papers. All other work is further classified and discussed in Section IV-C.

1) Social threats (43 items): The majority of work in this area focuses on SPam over Internet Telephony (SPIT) detection and prevention, although there are other items included in this category as well (*e.g.*, secure principal binding). We have broken down the work based on the general technical approach taken, and discuss the work in rough chronological order within each thrust; we use the same approach in the remainder of the text. As we can see, the majority of work has focused on reputation and behavior-based approaches.

a) Reputation, behavior, and identity (25 items): Srivastava and Schulzrinne [20] describe DAPES, a system for blocking SPIT calls and instant messages based on several factors, including the origin domain of the initiator (caller), the confidence level in the authentication performed (if any), whether the call is coming through a known open proxy, and a reputation system for otherwise unknown callers. They give an overview of other reputation-based systems and compare them with DAPES.

Dantu and Kolan [21] show that it is possible to use as a detection mechanism for high-volume SPIT the velocity and acceleration (first- and second-order derivative of the number) of incoming calls from a user, host or domain. Once either of these values exceeds a threshold, related calls can be dropped. The same method can also mitigate against certain VoIP-based denial of service attacks.

MacIntosh and Vinokurov [22] propose a statistical detection algorithm for SPIT that can be implemented at the receiver's server. For each external entity that communicates with local users, their system keeps track of the number of call setups and terminations in both directions (incoming and outgoing). Simultaneous deviation of two or more of these counters from their assumed long-term averages supposedly indicates spam activity, with confidence increasing as the deviation widens. The approach assumes that attackers cannot rapidly change their identity.

Croft and Olivier [23] propose extending the call setup process by adding a "call me back" scheme using a Verifying Authority (VA) and a Mediator. The Mediator acts as a call bridge, allowing the call to connect only once the VA has approved it (possibly based on policy and on such information as caller/callee identities, location and time of the call, *etc.*). The user receives the "call me back" request from the VA, and decides whether to proceed with picking up the call based on local policy and other information (*e.g.*, CallerID).

Dantu and Kolan [24], [25] describe the Voice Spam Detector (VSD), a multi-stage SPIT filter based on trust, reputation, and feedback among the various filter stages. The primary filter stages are call pattern and volume analysis, black and white lists of callers, per-caller behavior profile based on Bayesian classification and prior history, and reputation information from the callee's contacts and social network. They provide a formal model for trust and reputation in a voice network, based on intuitive human behavior. They evaluate their system in a

laboratory experiment using a small number of real users and injected SPIT calls.

Rebahi and Sisalem [26] develop the concept of the "SIP social network" as a means for managing reputation toward countering SPIT. However, no experimental evaluation or validation of any of these schemes is performed. Rebahi *et al.* [27] extend the previous by proposing two schemes for protecting against SPIT and SPIM (spam over instant messaging). The first uses reputation, with users indicating how much "trust" they have in the persons in their contact lists. These lists (and the trust values) are posted in a directory, where others can access them upon receiving a call from a previously unknown (to them) entity. This scheme requires that every user's contact information be published, and that attackers cannot mask or change their identities. The second scheme is built around the notion of "payment at risk", wherein a caller may be required to deposit a small amount to a SIP server prior to placing a call, depending on the callee's or the SIP proxy's policy. If the user indicates that the call was SPIT, the payment is then forfeit.

Hansen *et al.* [28] present SPIT-AL, an anomaly detection system seeking to identify SPIT calls. Their system takes into consideration information about the caller (such as CallerID, IP address, whitelists/blacklists, *etc.*) and the call (*e.g.*, time), and allows for different responses (grey-listing, audio CAPTCHA, *etc.*). A key element of their architecture is that users manage their own rules and responses, in order to comply with the various German telecommunication laws.

Baumann *et al.* [29] overview SPIT threats and various defense mechanisms. They then propose to prevent Sybil attacks in SIP by binding user identities to biometric information (specifically voice fingerprint) that is stored in global servers. Users wishing to place calls must first prove their identity, thereafter receiving credentials that can be used to place calls.

Madhosingh [30] integrates white and black lists with CAPTCHAs for those callers that are not *a priori* known (and included in a whitelist or a blacklist). If the test is passed, the call is allowed through. However, such callers are not allowed to leave voicemail messages to the callee's system; instead, such messages are stored on the caller's local SIP server, and the callee is sent an indication about the availability of a voicemail and instructions on how to retrieve it.

Bertrand *et al.* [31] propose an anomaly detection technique for identifying and blocking SPIT that creates caller profiles based on their IP address. The criteria used by in their analysis includes number of received error messages, the use of a directory service, whether multiple calls are placed by the same caller, the duration of calls and the variance of call duration across multiple calls, and the number of simultaneous incoming calls (from multiple different users) to the same user. In response to an identified SPIT call, their system rate limits call delivery, temporary blacklists the most aggressive callers, or redirects the call to a voicemail or other automated system that notifies the caller of the problem. They propose implementing this functionality in the network, where it will work together with routers. Because of this choice, their system must do real-time layer 7 reconstruction and analysis of traffic, which in turn requires hardware support to keep up with the volume. Such a system should be able to handle up to 10^6 simultaneous sessions, for 10^5 subscribers, with 10^4 incoming calls per second. They present a prototype Java-based implementation running on Linux, using the `netfilter` and `iptables` components to divert and block traffic. Their performance evaluation shows that this prototype can handle 80 incoming calls per second, adding approximately 5*ms* to the average 5.8 seconds call establishment time.

Yan *et al.* [32] argue for the use of active fingerprinting in SPIT prevention systems. Protocol implementations interpret the standards in slightly different ways, especially with respect to indicating errors. Thus, it is possible to identify the implementation of a peer SIP device by observing its responses to a set of specially crafted messages. These may be either standards-compliant or non-compliant. By creating a number of different tests, it is possible to actively fingerprint a remote SIP device that is trying to initiate a call. Their conjecture is that malicious SIP user agents will not be able to mimic legitimate stacks because of the diversity in possible responses, and because often such tools implement only a subset of SIP. In their analysis, they were able to create unique fingerprints for 20 different SIP devices. The system evaluation was limited to a performance (throughput) oriented experiment using PlanetLab.

Balasubramaniyan *et al.* [33] propose to use call duration and social network graphs to establish a measure of reputation for callers. Their intuition is that users whose call graph has a relatively small fan-out and whose call durations are relatively long are less likely to be spammers. Conversely, users who place a lot of very short calls are likely to be engaging in SPIT. Furthermore, spammers will receive few (if any) calls. Their system works both when the parties in a call have a social network link between them, and when such a link does not exist by assigning global reputation scores. Users that are mistakenly categorized as spammers are redirected to a Turing test, allowing them to complete the call if they answer correctly. In a simulation-based evaluation, the authors determine that their system can achieve a false negative rate of 10% and a false positive rate of 3%, even in the presence of large numbers of spammers.

Ono and Schulzrinne [34] propose the use of weak social ties as a means to label calls with unknown or incomplete caller ID information, in conjunction with a blacklist/whitelist filtering scheme. As one specific mechanism, they describe the use of weakly secret information that a user makes available to potential callers, who must then use that information in future calls. Another similar technique involves callers providing contact/identifying information to potential future callees. Both of these schemes exploit cross-media interactions, leveraging the fact that most calls are associated with some other interaction between the caller and callee entities. For example, an e-commerce web site may accept such a weak secret from a customer, or provide one (depending on the scheme); this secret would then be used when calling the customer in the future.

Guang-Yu *et al.* [35] describe a multi-layer SPIT detection and prevention architecture that takes into consideration the behavioral characteristics of specific types of SPIT campaigns,

starting from the reconnaissance phase.

Patankar *et al.* [36] compare two SPIT detectors derived from the email spam domain. One of these techniques is based on user reputation through a referral social network model, while the other assigns a trust value to incoming SIP messages based on their direct prior interactions with the caller. Their simulations indicate that the referral-based model is more effective, correctly identifying SPIT in over 98% of cases. In an environment with little-to-moderate amounts of SPI, this likely be sufficient by itself. If the level of SPIT approaches the current (*circa* 2010) levels of email spam, then additional filtering/blocking mechanisms would have to be employed.

Wu *et al.* [37] apply semi-supervised clustering to call parameters (with optional user feedback) in order to distinguish SPIT from non-SPIT calls. The evaluation, which was done using manually created call traces, shows that the approach is scalable (in the number of calls) and offers reasonable detection performance. Hyung-Jong *et al.* [38] describe a behavior-based system that seeks to identify likely SPIT callers.

Sorge and Seedorf [39] apply reputation techniques to the SPIT problem, by evaluating the quality of information (tags) attached to outgoing calls by the callers' SIP-based service provider (SSP). Their scheme allows receiving SIP providers to evaluate the likelihood of a call being SPIT using caller-SSP information, providing incentives to honest SSPs to correctly tag their outbound calls. They demonstrate, through analytical means, that the precision of their SPIT detection improves by almost 50% even in a limited trust case, with greater improvements as longer trust chains of SSPs are taken into consideration.

Phithakkitnukoon and Dantu propose the use of user feedback in closed email systems (such as Gmail) to identify spammers [40]. The challenge in their scheme, which envisions a binary "spammer/non-spammer" classification is to choose an appropriate threshold for determining when this transition occurs so as not to misclassify benign users who were accidentally or maliciously tagged as spammers.

b) Content-based detection (1 item): Pörschmann and Knospe [41] propose a SPIT detection mechanism based on applying spectral analysis to the audio data of VoIP calls to create acoustic fingerprints. SPIT calls are identified by detecting a large number of fingerprints across a large number of different calls.

c) Policy-based approaches (3 items): Tschofenig *et al.* [42] propose the use of a SIP Authentication service that uses the Security Assertion Markup Language (SAML) to specify authentication requirements for SIP callers prior to placing a call. The relevant information (*e.g.*, identity) is then forwarded to the receiver, which can use similar SAML policies to determine whether they are willing to receive the call.

d'Heureuse *et al.* [43] describe an anti-SPIT system that integrates user roles and personal preferences in its approach to blocking unwanted calls. Their system allows users to express their requirements and current status using a policy language based on an extended version of the IETF-standardized Call Processing Language (CPL). The authors give some paper examples of how users might use such a system, and briefly describe a prototype implementation.

Soupionis *et al.* [44] propose a policy-based approach for defending against SPIT. They use a rules-based approach to SPIT detection, combined with a number of mitigation strategies and mechanisms that policy can invoke in response to detection.

d) CAPTCHAs and puzzles (4 items): Banerjee *et al.* [45] propose the use of computational puzzles as part of the identity (public/private key pair) generation phase in peer-to-peer VoIP networks to prevent spammers from creating large numbers of disposable identities. Once identities become harder to generate on demand, trust and reputation-based mechanisms can be used to manage SPIT.

Quittek *et al.* [46] propose the use of *hidden* Turing tests to identify SPIT callers. As a concrete approach, they leverage the interaction model in human conversation, which minimizes the amount of simultaneous ("double") talk by the participants, and the fact that there is a short pause at the beginning of an answered call, followed by a statement by the callee that initiates the conversation. By looking for signs of violation of such norms, it is possible to identify naïve automated SPIT callers. The authors implement their scheme and integrate it with a VoIP firewall.

Wang [47] describes an end-point audio CAPTCHA system for countering SPIT, meant to be installed and used by users and system administrators. She conducts a usability study, examining the installation and management overhead of the tool (including the design and recording of challenge questions), the understandability and time-to-answer of the system and the questions by legitimate callers, and correctness in answering. For the latter metric, she focuses specifically on English-as-Second-Language users. Lindqvist and Komu [48] describe a similar approach using image human-interaction proofs in conjunction with SIP.

e) Anti-SPIT architectures (7 items): Niccolini [49] discusses the difficulties in protecting against IP telephony spam (SPIT) and overviews the various approaches for blocking such calls, identifying the technical and operational problems with each. Possible building blocks for SPIT prevention include blacklists/whitelists combined with strong identity verification to provide a reliable CallerID system, referral-based systems among trusted SIP domains [50], [51], pattern or anomaly detection techniques to discriminate SPIT based on training data, multi-level grey-listing of calls based on caller behavior (similar to throttling) [52], [53], computational puzzles and CAPTCHAs, explicit callee consent (a form of capability, required to actually place a call), content filtering on voicemail spam, callee feedback to indicate whether a call was SPIT or legitimate (typically combined with whitelisting/blacklisting, and requiring strong identity), changing one's SIP address as soon as SPIT messages arrive, requiring a monetary fee for the first contact, and legal action. Niccolini argues that none of these methods by itself is likely to succeed, promotes a modular and extensible approach to SPIT prevention, and presents a high-level architecture with these properties that was designed for use in a commercial SIP router.

Schlegel *et al.* [54] describe a framework for preventing SPIT. They argue for a modular approach to identifying SPIT, using hints from both signaling and media transfer. The first

stage of their system looks at information that is available prior to accepting the call, while the second stage interacts with a caller (possibly prior to passing on the call to the callee). The various components integrated in their system include whitelists/blacklists, call statistics, IP/domain correlation, and Turing tests. Their system also allows for feedback from the callee to be integrated into the scoring mechanism, for use in screening future calls. The evaluation focuses on scalability, by measuring the response time to calls as call volumes increase. A similar architecture, with some additional components, is described by Quittek *et al.* [55] and later extended [56].

The SPIDER project (SPam over Internet telephony Detection sERvice) third public report [57] describes an anti-SPIT architectural framework. (The first two reports are described later, in Section IV-C1.) Elements of this architecture include improved authentication, whitelisting/blacklisting, behavior analysis, the use of computational puzzles for challenge/response, reputation management, and audio content analysis.

Mathieu *et al.* [58] describe SDRS, an anti-SPIT system that combines several of detection schemes and takes into consideration users' and operators' preferences. Gritzalis and Mallios [59] survey various defenses against SPIT, and propose an integrated framework for mitigating the various limitations of each individual mechanism.

f) Miscellaneous on SPIT (2 items): Kolan *et al.* [60] use traces of voice calls in a university environment to validate a mathematical model for computing the nuisance level of an incoming call, using feedback from the receivers. The model is intended to be used in predicting SPIT calls in VoIP environments, and is based on the history of prior communications between the two parties involved, which includes explicit feedback from the receiver indicating that a call is unwanted (at a particular point in time). Dritsas *et al.* [61] combine several criteria that they argue define a SPIT call with an ontology for SPIT, towards improving the management of SPIT incidents.

g) Location (1 item): Kong *et al.* [62] propose a scheme for securing the user location information in SIP, *i.e.,* the integrity and authenticity of the binding a principal's SIP URI and a correponding device's contact/network address. The threat addressed is tampering with the location information of a user such that calls to that user are redirected to a malicious party (impersonation) or are dropped (denial of service). In their approach, users create temporary public keys that are bound to their location and identity through the SIP registration process, possibly leveraging the existing SIP authentication mechanism used (or using some out-of-bound mechanism for securing the binding). Users then digitally sign their registration information, which the local registrar verifies before sending to the location server. To allow entities in other domains to verify the location information, the user public key can be conveyed through a secure channel at the domain level, *e.g.,* by leveraging registrar public key certificates, or a pairwise shared secret key between two domains. This approach assumes benign and reliable registrar servers. To mitigate this weakness in the assumptions and to improve overall service reliability, the authors also propose the use of Byzantine Fault Tolerance techniques, adapting their protocols (public key binding & querying, and user registration) to a quorum environment. They conduct an experimental evaluation of their non-replicated scheme, showing that it can achieve the same performance as unsecured SIP and is 3–50 times faster than TLS-protected SIP.

2) Eavesdropping, Interception, and Modification (30 items): Considerable work has been dedicated to protecting and attacking VoIP signaling and data traffic. We divide the work in two sub-categories, attacks and defenses.

a) Attacks (12 items): Wang *et al.* [63], [64] describe a de-anonymization attack against VoIP streams that use low-latency anonymity proxies. Their intuition is to insert a watermark in the encrypted stream, tracking its propagation across the network. The watermark used is a perturbation of the inter-packet delay for selected packets in the stream. With appropriate use of redundancy, they demonstrate a tracking attack against 2-minute Skype calls across the Internet using 3 *ms* delays. Depending on the watermark parameters chosen, they can achieve 99% true positive and 0% false positive rate or 100% true positive and 0.1% false positive rate. Srivatsa *et al.* [65] demonstrate flow-analysis attacks that expose the privacy of peer-to-peer VoIP participants.

Shah *et al.* [66] examine the use of injected jitter into VoIP as a covert channel to exfiltrate keyboard activity of interest (*e.g.,* passwords). This attack would be effective even when the VoIP stream is encrypted.

Takahashi and Lee [67] examine the problem of covert channels in VoIP protocols, identifying and quantifying several ways in which data can be surreptitiously leaked out of a user's system or an enterprise network. As an example, they demonstrate the steganographic insertion of a second voice channel in a SIP-based VoIP conversation. This has the potential of leaking an otherwise secure (encrypted) conversation through a secondary channel, or can be used to hide the true communication content from an eavesdropper. They determine such parameters as channel capacity and perceptual quality of the encoded signal through experimental evaluation. They conclude with a discussion of several possible countermeasures and detection methods.

Weiser *et al.* [68] provide an overview of the security considerations in RTP, the media transfer protocol used in both SIP and H.323. They analyze six different implementations, discovering confidentiality (eavesdropping a call), integrity (injecting voice into an ongoing call) and availability (performing DoS) compromises. This work assumes that no security mechanism (such as SRTP) is used.

Wright *et al.* [69] apply machine learning techniques to determine the language spoken in a VoIP conversation, when a variable bit rate (VBR) voice codec is used based on the length of the encrypted voice frame. As a countermeasure, they propose the use of block ciphers for encrypting the voice. In follow-on work [70] they use profile Hidden Markov Models to identify specific phrases in the encrypted voice stream with a 50% average accuracy, rising to 90% for certain phrases.

Wang *et al.* [71] evaluate the resilience of three commercial VoIP services (AT&T, Vonage and Gizmo) against man-in-the-middle adversaries. They show that it is possible for an

attacker to divert and redirect calls in the first two services by modifying the RTP endpoint information included in the SDP exchange (which is not protected by the SIP Digest Authentication), and to manipulate a user's call forwarding settings in the latter two systems. These vulnerabilities permit for large-scale voice pharming, where unsuspecting users are directed to fake interactive voice response systems or human representatives. The authors argue for the need for TLS or IPsec protection of the signaling.

Verscheure *et al.* [72] exploit the nature of human conversation (*i.e.*, alternating periods of talking and silence for each participant) to reveal communication pairs over a period of time. The technique does not work as well against systems that do not use silence suppression, as these effectively introduce a form of constant (voice) traffic padding in both directions.

Petraschek *et al.* [73] examine the usability and security of ZRTP, a key agreement protocol based on the Diffie Hellman key exchange, designed for use in VoIP environments that lack pre-established secret keys among users or a public key infrastructure (PKI). ZRTP is intended to be used with SRTP, which performs the actual content encryption and transfer. Because of the lack of a solid basis for authentication, which makes active man-in-the-middle attacks easy to launch, ZRTP uses Short Authentication Strings (SAS) to allow two users to verbally confirm that they have established the same secret key. The verbal communication serves as a weak form of authentication at the human level. The authors identify a relay attack in ZRTP, wherein a man-in-the-middle adversary can influence the SAS read by two legitimate users with whom he has established independent calls and ZRTP exchanges. The attacker can use one of the legitimate users as an oracle to pronounce the desired SAS string through a number of means, including social engineering. The authors point out that SAS does not offer any security in some communication scenarios with high security requirements, *e.g.*, a user calling (or being called by) their bank. The authors implement their attack and demonstrate it in a lab environment.

Zhang *et al.* [74] show that, by exploiting DNS and VoIP implementation vulnerabilities, it is possible for attackers to perform man-in-the-middle attacks even when they are not on the direct communication path of the parties involved. They demonstrate their attack against Vonage, requiring that the attacker only knows the phone number and the IP address of the target phone. Such attacks can be used to eavesdrop and hijack the victims' VoIP calls. The authors recommend that users and operators use signaling and media protection, conduct fuzzing and testing of VoIP implementations, and develop a lightweight VoIP intrusion detection system to be deployed on the VoIP phone.

b) Defenses (18 items): Guo *et al.* [75] propose a new scheme for protecting voice content that provides strong confidentiality guarantees while allowing for graceful voice degradation in the presence of packet loss. They evaluate their scheme via simulation and micro-benchmarks. However, Li *et al.* [76] show that the scheme is insecure.

Bellovin *et al.* [77] argue against the enactment of legislation (in the US) mandating the integration of lawful-intercept capabilities into VoIP implementations. Their key concerns is that, based on a history of system compromises and implementation weaknesses, mandating such capabilities would enable or ease attacks against personal communications by adversaries that would otherwise be unable to conduct such attacks. They suggest that lawful interception needs be met either at the application provider or the network link level.

Seedorf [78] proposes the use of cryptographically generated SIP URIs to protect the integrity of content in P2P SIP. Specifically, he uses self-certifying SIP URIs that encode a public key (or, more compactly, the hash of a public key). The owner of the corresponding private key can then post signed location binding information on the peer-to-peer network (*e.g.*, Chord) that is used by call initiators to determine call routing.

Fessi *et al.* [79] propose extensions to P2P SIP that provide location and interaction privacy for participants. They develop a signaling protocol for P2P SIP that uses two different Kademlia-based overlay networks for storing information and forwarding traffic, respectively. Their scheme requires a centralized authentication server, which provides verifiable identities at the application/SIP layer. They consider attacks against their scheme, shared with more general anonymity systems (such as Tor). They use analytical models to estimate communication reliability, cryptographic overhead, and end-to-end signaling latency.

Talevski *et al.* [80] describe the addition of security (in the form of encryption and integrity protection) to a lightweight VoIP protocol suitable for mobile devices. Kuntze *et al.* [81] propose a mechanism for providing non-repudiation of voice content by using digital signatures, taking into consideration packet losses by reporting to the sender which packets were actually received.

Wang *et al.* [82] extend the SIP call setup to include a Diffie Hellman based key exchange that results in multiple shared keys that the parties switch among during the call in a deterministic (but unknown to an adversary) fashion. Their stated goal is to impede cryptanalytic attacks that depend on the same shared secret key being used throughout a call. They conduct a performance evaluation using a prototype implementation of their scheme on software phones, concluding that the overhead is negligible. The likely adoption of DTLS-SRTP would probably supersede this effort. Gurbani and Kolesnikov [83] discuss DTLS-SRTP and SDES (another proposed protocol for media protection), and propose a lightweight scheme that mitigates some of the performance concerns and security weaknesses of DTLS-SRTP.

Hlavacs *et al.* [84] propose the integration of computational puzzles in ZRTP as a way to mitigate the man-in-the-middle attack described earlier [73]. Effectively, their scheme places an upper bound to the amount of time a ZRTP exchange may take, placing the attacker under (hopefully) severe time constraint and making them unable to carry out the independent but parallel calls that are necessary. The authors propose a new puzzle scheme based on computing selected eigenvectors of real symmetric matrices. An additional protection mechanism suggested is to randomly delay (by short amounts of time) the receiving of calls, again trying to make more difficult the attacker's task of orchestrating and playing against each other two independent calls.

Palmieri and Fiore [85] describe an adaptation of SIP to provide end-to-end security using existing and well-known primitives (*e.g.*, digital signatures and efficient encryption mechanisms). The authors developed a prototype implementation and conducted a performance analysis of their scheme. One drawback of this scheme relative to ZRTP is that it requires a PKI. When compared to at least some proposed deployments of DTLS-SRTP, this scheme provides end-to-end non-repudiation and end-to-end authentication while being resistant to man-in-the-middle attacks.

Zhang and Berthold [86] discuss several passive traffic analysis attacks on VoIP systems. These attacks exploit both signaling and media flow information. They also discuss techniques that can be used to mitigate some of these attacks, and conclude with a list of open problems. Many of the attacks and the countermeasures are shared with those in general-purpose anonymity systems. Zhang and Fischer-Hübner [87] and Melchor *et al.* [88] also discuss techniques for protecting the privacy of VoIP calls. The former studies an approach based on using an anonymization overlay network (such as Tor) with traffic padding (where the overlay knows what traffic to drop because it is marked by the sender). The latter discussed and evaluated (using an analytical model) the use of MIXes to provide strong resistance against traffic analysis for VoIP flows. Their scheme uses dummy traffic, broadcasting, and private information retrieval as building blocks. Srivatsa *et al.* [89] examined the problem of on-demand construction of QoS-sensitive routes in anonymizing networks.

Elbayoumy and Shepherd [90] propose the use of TEA (Tiny Encryption Algorithm) as a lightweight confidentiality mechanism. Subsequently, they propose an adaptive scheme where the selection of encryption algorithm to be used in protecting traffic is made with consideration of the CPU capabilities of both communicating parties [91], [92].

3) Denial of Service (31 items): Reynolds and Ghosal [93] describe a multi-layer protection scheme against flood-based application- and transport-layer denial of service (DoS) attacks in VoIP. They use a combination of sensors located across the enterprise network, continuously estimating the deviation from the long-term average of the number of call setup requests and successfully completed handshakes. Similar techniques have been used in detecting TCP SYN flood attacks, with good results. The authors evaluate their scheme via simulation, considering several different types of DoS attacks and recovery models.

Larson *et al.* [94], [95] experimentally analyzed the impact of distributed denial of service (DDoS) attacks on VoIP call quality. They also established the effectiveness of low-rate denial of service attacks that target specific vulnerabilities and implementation artifacts to cause equipment crashes and reboots. They discuss some of the possible defenses against such attacks and describe Sprint's approach, which uses regional "cleaning centers" which divert suspected attack traffic to a centralized location with numerous screening and mitigation mechanisms available. They recommend that critical VoIP traffic stay on private networks, the use of general DDoS mechanisms as a front-line defense, VoIP-aware DDoS detection and mitigation mechanisms, traffic policing and rate-limiting mechanisms, the use of TCP for VoIP signaling (which makes IP spoofing, and hence anonymous/unfilterable DoS attacks, very difficult), extended protocol compliance checking by VoIP network elements, and the use of authentication where possible.

Bremler-Barr *et al.* [96] describe de-registration attacks in SIP, wherein an adversary can force a user to be disassociated with the proxy server and registrar, or to even divert that user's calls to any party (including to the attacker). This attack works even when authentication is used, if the adversary can eavesdrop on traffic between the client and the SIP proxy. They demonstrate the attack against several SIP implementations, and propose a protection mechanism that is similar to one-time passwords.

Chen [97] describes a denial of service detection mechanism that models the SIP transaction state machine and identifies attacks by measuring the number of transaction and application errors, the number of transactions per node, and the traffic volume per transaction. If certain thresholds are exceeded, an alert is generated. Chen does not describe how appropriate thresholds can be established, other than to indicate that historical records can be used.

Sengar *et al.* [98], [99] describe vFDS, an anomaly detection system that seeks to identify flooding denial of service attacks in VoIP. The approach taken is to measure abnormal variations in the relationships between related packet streams using the Hellinger distance, a measure of the deviation between two probability measures. Using synthetic attacks, they show that vFDS can detect flooding attacks that use SYN, SIP, or RTP packets within approximately 1 second of the commencement of an attack, with small impact on call setup latency and voice quality. A similar approach, using Hellinger distance on traffic sketches, is proposed by Tang *et al.* [100], overcoming the limitations of the previous schemes against multi-attribute attacks. Furthermore, their scheme does not require the constant calculation of an accurate threshold (defining "normal" conditions).

Zhang *et al.* [101] describe a denial of service attack wherein adversaries flood SIP servers with calls involving URIs with DNS names that do not exist. Servers attempting to resolve them will then have to wait until the request times out (either locally or at their DNS server), before they can continue processing the same or another call. This attack works against servers that perform synchronous DNS resolution and only maintain a limited number of execution threads. They experimentally show that as few as 1,000 messages per second can cause a well provisioned synchronous-resolution server to exhibit very high call drops, while simple, single-threaded servers can be starved with even 1 message per second. As a countermeasure, they propose the use of non-blocking DNS caches, which they prototype and evaluate.

Fiedler *et al.* [102] present VoIP Defender, an open architecture for monitoring SIP traffic, with a primary focus on high-volume denial of service attacks. Their architecture allows for a variety of detection methods to be integrated, and several different attack prevention and mitigation mechanisms to be used. Key design goals include transparency, scalability, extensibility, speed and autonomous operation. Their evalua-

tion of the prototype implementation consists exclusively of performance measurements.

Conner and Nahrstedt [103] describe a semantic-level attack that causes resource exhaustion on stateful SIP proxies by calling parties that (legitimately or in collusion) do not respond. This attack does not require network flooding or other high traffic volume attacks, making it difficult to detect with simple, network-based heuristics used against other types of denial of service attacks. They propose a simple algorithm, called *Random Early Termination* (RET) for releasing reserved resources based on the current state of the proxy (overloaded or not) and the duration of each call's ringing. They implement and evaluate their proposed scheme on a SIP proxy running in a local testbed, showing that it reduces the number of benign call failures when under attack, without incurring measurable overheads when no attack is underway.

Luo *et al.* [104] experimentally evaluate the susceptibility of SIP to CPU-based denial of service attacks. They use an open-source SIP server in four attack scenarios: basic request flooding, spoofed-nonce flooding (wherein the target server is forced to validate the authenticator in a received message), adaptive-nonce flooding (where the nonce is refreshed periodically by obtaining a new one from the server), and adaptive-nonce flooding with IP spoofing. Their measurements show that these attacks can have a large impact on the quality of service provided by the servers. They propose several countermeasures to mitigate against such attacks, indicating that authentication by itself cannot solve the problem and that, in some circumstances, it can exacerbate its severity. These mitigation mechanisms include lightweight authentication and whitelisting, proper choice of authentication parameters, and binding of nonces to client IP addresses.

Fuchs *et al.* [105] apply anomaly detection techniques to protect against VoIP-originated denial of service attacks at the phone call level at public safety service centers (*e.g.*, 911 or 112 operators). Specifically, they use call traces from normal operations to determine the level of calls coming from the PSTN, GSM and VoIP networks during normal operation and at disaster time. They then use these profiles to discriminate against VoIP-based DoS attacks by limiting the accepted number of calls that can originate from that domain, building on previous work that identified the network of origin as a potential discriminator [106]. Using call traces from a fire department response center, they evaluate the call response rate against the DoS attack intensity. Their analysis shows that it is possible to identify such attacks early and to avoid false positives if VoIP-originated calls under normal scenarios are less than 27% of total call volume.

Hyun-Soo *et al.* [107] propose a detection mechanism for de-registration and other call disruption attacks in SIP that is based on message retransmission: when a server receives an unauthenticated (but possibly legitimate) message M that could disturb a call or otherwise deny service to a user, it asks the user's agent to retransmit the last SIP message sent by that agent, as an implicit authenticator. If the retransmission matches M (*i.e.*, this was a legitimate request), the server proceeds with its processing. If the retransmission does not match M, or if multiple retransmissions are received within a short time window (as may be the case when an attacker can eavesdrop on the network link between the SIP proxy and the user, identifying the request for retransmission), M is discarded. However, the scheme requires a new SIP message to signal that a retransmission is needed. Geneiatakis and Lambrinoudakis [108], [109] consider some of the same attacks, and propose mitigation through an additional SIP header that must be included in all messages and can cryptographically validate the authenticity and integrity of control messages.

Ormazabal *et al.* [110] describe the design and implementation of a SIP-aware, rule-based application-layer firewall that can handle denial of service (and other) attacks in the signaling and media protocols. They use hardware acceleration for the rule matching component, allowing them to achieving filtering rates on the order of hundreds of transactions per second. The SIP-specific rules, combined with state validation of the endpoints, allow the firewall to open precisely the ports needed for only the local and remote addresses involved in a specific session, by decomposing and analyzing the content and meaning of SIP signaling message headers. They experimentally evaluate and validate the behavior of their prototype with a distributed testbed involving synthetic benign and attack traffic generation.

Ehlert *et al.* [111], [112] propose a two-layer DoS prevention architecture for SIP. The first layer is comprised of a bastion host that protects against well-known network-layer attacks (such as TCP SYN flooding) and SIP-flooding attacks. The second layer is located at the SIP proxy, and is composed of modules that perform signature-based detection of malformed SIP messages and a non-blocking DNS cache to protect against attacks involving SIP URIs with irresolvable DNS names [101]. They conduct a series of evaluations in an experimental testbed, where they validate the effectiveness of their architecture to block or mitigate a number of DoS attacks. Ehlert *et al.* [113] separate propose and experimentally evaluate (via a testbed) a specification-based intrusion-detection system for denial of service attacks. Geneiatakis *et al.* [114], [115] use counting Bloom filters to detect messages that are part of a denial of service attack in SIP by determining the normal number of pending sessions for a given system and configuration based on profiling.

Awais *et al.* [116] describe an anti-DoS architecture based on bio-inspired anomaly detection. They compare their scheme against a cryptography-based mechanism using synthetic traffic. Similar work is described by Rebahi *et al.* [117]. Akbar and Farooq [118] conduct a comparative evaluation of several evolutionary and non-evolutionary machine learning algorithms using synthetic SIP traffic datasets with different levels of attack intensities and durations. They conclude that different algorithms and settings are best suited for different scenarios. The same authors subsequently apply anomaly detection techniques to identify RTP fuzzing attacks that seek to cause server crashes through malformed packet headers and payloads [119]. They investigate several different classifiers, analyzing their accuracy and performance using synthetic RTP traces. Nassar *et al.* [120] use support vector machine (SVM) classifiers on 38 distinct features in SIP traffic to identify SPIT and DoS traffic. Their experiments using SIP traffic traces

show good performance and high detection accuracy.

Rafique *et al.* [121] analyze the robustness and reliability of SIP servers under DoS attacks. They launch a number of synthesized attacks against four well-known SIP proxy servers (OpenSER, PartySIP, OpenSBC, and MjServer). Their results demonstrate the ease with which SIP servers can be overloaded with call requests, causing such performance metrics as Call Completion Rate, Call Establishment Latency, Call Rejection Ration and Number of Retransmitted Requests to deteriorate rapidly as attack volume increases, sometimes with as few as 1,000 packets/second. As an extreme case of such attacks large volumes of INVITE messages can even cause certain implementations to crash. While valuable in documenting the susceptibility to such attacks, this work proposes no defense strategies or directions.

Akbar *et al.* [122] conduct an analysis of three anomaly detection algorithms for detecting flood attacks in IMS: adaptive threshold, cumulative sum, and Hellinger distance. They use synthetic traffic data to determine the detection accuracy of these algorithms in the context of a SIP server being flooded with SIP messages.

Battistello [123] introduces a DoS-resistant protocol for authenticated call establishment with key exchange across different domains.

4) Service Abuse (7 items): Truong *et al.* [124] describe a rules-based intrusion detection system for H.323 that uses an FSM model to detect unexpected messages, aimed at identifying illegitimate RAS (Registration, Admission and Status) messages being forwarded to a H.323 gatekeeper.

Kotulski and Mazurczyk [125], [126], [127] propose the use of steganographic and digital watermarking to embed additional information into SIP traffic to provide stronger origin authentication and content integrity guarantees in a bandwidth-sensitive manner. Their scheme encodes the necessary information into unused fields in the IP, UDP and RTP protocol headers, and also into the transmitted voice.

Zhang *et al.* [128] present a number of exploitable vulnerabilities in SIP that can manipulate billing records in a number of ways, showing their applicability against real commercial VoIP providers. Their focus is primarily on attacks that create billing inconsistencies, *e.g.*, customers being charged for service they did not receive, or over-charged for service received. Some of these attacks require a man-in-the-middle capability, while others only require some prior interaction with the target (*e.g.*, receiving a call from the victim SIP phone device).

Abdelnur *et al.* [129] use AVISPA to identify a protocol-level vulnerability in the way SIP handles authentication [130]. AVISPA is a model checker for validating security protocols and applications using a high-level protocol specification and security-goals language that gets compiler into an intermediate format that can be consumed by a number of lower-level checkers. The attack is possible with the SIP Digest Authentication, whereby an adversary can reuse another party's credentials to obtain unauthorized access to SIP or PSTN services (such as calling a premium or international phone line). This attack is possible because authentication may be requested in response to an INVITE message at any time during a call, and the responder may issue an INVITE message during a call either automatically (because of timer expirations) or through a user action (*e.g.*, placing the caller on hold in order to do a call transfer). While the solution is simple, it requires changes possibly to all end-device SIP implementations.

Geneiatakis *et al.* [131] address the problem of billing attacks against telephony service providers and their users. They propose an authentication-based scheme that leverages the existing Authentication, Authorization and Accounting (AAA) infrastructure operated by the service provider to provide the latter with explicit and non-repudiable call confirmation by the call initiator. However, the scheme has not been implemented or evaluated, experimentally or formally.

C. Additional Categories (134 items)

We now classify the remainder of the surveyed work (55% of the total) using the following categories: Overviews (19.7%), Field Studies and Analysis (4.9%), Performance Analysis (5.7%), Authentication Protocols (6.1%), Architecture (7.8%) Middleboxes (4.5%) Intrusion Detection (4.5%), and Miscellaneous (0.8%).

1) Overviews and Surveys (50 items): There is a considerable body of work focusing on surveying and summarizing risks and threats in SIP, and describing existing work on defense mechanisms.

a) General overviews (42 items): Ackermann *et al.* [132] describe threats in VoIP, focusing on specific attacks and vulnerabilities as case studies. Hunter [133], Batchvarov [134], Bradbury [135], and Chau [136] provide summaries of specific security concerns in VoIP.

Sicker and Lookabaugh [137] discuss threats in VoIP and the need for security to be integrated at design and deployment time. Vuong and Bai [138] provide a brief survey of the types of intrusion detection systems that can be used to monitor for specific types of attacks in VoIP.

Geneiatakis *et al.* [139] describe how SQL injection attacks can be launched through SIP, by including partial SQL statements in certain fields of SIP protocol messages that are likely to be used in subsequent database operations (*e.g.*, parts of the SIP URI in the `To:` field may be used to look up the location of the user receiving the call). They demonstrate the attack in a lab experiment, and briefly discuss the applicability of general SQL injection defense mechanisms in a SIP environment.

Tucker [140] gives an overview of SIP and H.323, and briefly mentions some security concerns (with an emphasis on denial of service). Posegga and Seedorf [141] offer a similar threat analysis. Edelson [142] discusses denial of service, SPIT, eavesdropping and security of emergency calls, before talking about the particular requirements of VoIP in wireless. She concludes with a brief discussion of intrusion detection for VoIP. Albers *et al.* [143] gives a high-level overview of the types of vulnerabilities that SIP-based systems may be exposed to, and discusses the capabilities and limitations of a number of commercially available (as of 2005) SIP intrusion prevention and testing systems. In a related publication, McGann and Sicker [144] argue that several of the VoIP security tools available in 2005 did not cover the extent of known

vulnerabilities, do not provide the coverage claimed by the developers, and were not user-friendly. A short overview of some SIP security mechanisms is given by Geneiatakis *et al.* [145].

Cao and Malik [146], [147] examine the vulnerabilities that arise from introducing VoIP technologies into the communications systems in critical infrastructure applications. They examine the usual threats and vulnerabilities, and discuss mitigation techniques. They conclude by providing some recommendations and best practices to operators of such systems.

Allain [148] discusses the security challenges in VoIP environments, focusing on a couple of specific issues to highlight the tradeoffs. Adelsbach *et al.* [149] provide a comprehensive description of SIP and H.323, a list of threats across all networking layers, and various protection mechanisms. A similar analysis was published by the US National Institute of Standards and Technology (NIST) [150]. An updated summary, with practical recommendations to users and operators is provided by Walsh and Kuhn [151]. Anwar *et al.* [152] identify some areas where the NIST report remains incomplete: counter-intuitive results with respect to the relative performance of encryption and hash algorithms, the non-use of the standardized Mean Opinion Score to evaluate call quality, and the lack of anticipation of RTP-based denial of service. They then propose the use of design patterns to address the problems of secure traversal of firewalls and NAT boxes, detecting and mitigating DoS attacks in VoIP, and securing VoIP against eavesdropping.

Geneiatakis *et al.* [153] also survey a number of SIP security vulnerabilities. Geneiatakis *et al.* [154] categorize potential attacks on VoIP services, and provide recommendations and guidelines for protecting the infrastructure. They use ontologies to represent these recommendations, and first-order logic to translate them to a unified security policy for VoIP.

Me and Verdone [155] describe the security threats and high-level vulnerabilities in SIP when used in 802.11 or other similar wireless environments. Singhai and Sahoo [156] describe the risks of VoIP technologies (focusing on SIP and H.323) and compare them with the public switched telephony network (PSTN). Rippon [157] provides a laundry list of threats and mitigation techniques for VoIP systems. Brief descriptions of some VoIP-related threats are given by Hung and Martin [158], [159] and Zandi *et al.* [160].

Xin [161] provides a somewhat more detailed overview of VoIP-related security concerns. Persky gives a very detailed description of several VoIP vulnerabilities [162]. Quinten *et al.* [163] survey the various techniques for preventing and reducing SPIT, offering some suggestions as to possible combinations that increase overall blocking effectiveness. Hansen and Woodward [164] overview threats in VoIP environments and recommend that VoIP and data networks be logically or physically separated. James and Woodward [165] propose a security framework for end users of VoIP technologies, combining a number of commonly available mechanisms and recommendations.

Butcher *et al.* [166] overview security issues and mechanisms for VoIP systems, focusing on security-oriented operational practices by VoIP providers and operators. Such practices include the separation of VoIP and data traffic by using VLANs and similar techniques, the use of integrity and authentication for configuration bootstrapping of VoIP devices, authentication of signaling via TLS or IPsec, and the use of media encryption. They briefly describe how two specific commercial systems implement such practices, and propose some directions for future research.

A comprehensive discussion of threats and security solutions is given by Thermos and Takanen [167]. Kurmus and Garet [168] summarize a number of threats and specific vulnerabilities using actual attack tools.

Sisalem *et al.* [169] provide an in-depth description of SIP and IMS, discussing the security mechanisms available in each part of the architecture. The focus particularly on the DoS and SPIT threats, also describing some available countermeasures.

Gurbani and Kolesnikov [170] discuss in depth and compare SDES, DTLS-SRTP, and ZRTP in terms features supported (*e.g.*, conferencing, PSTN calling) and security features/weaknesses (*e.g.*, susceptibility to man-in-the-middle attacks and key leakage). They conclude that all three are suitable, but they each offer a feature or suppress a vulnerability that the others do not.

Keromytis [1], [2], [3] surveys over 200 vulnerabilities in SIP implementations that were disclosed in the CVE database from 1999 to 2009. He classifies these vulnerabilities along several dimensions, including the VoIPSA threat taxonomy, the traditional Confidentiality/Integrity/Availability concerns, and a Protocol/Implementation/Configuration axis. He finds that the various types of denial of service attacks constitute the majority of disclosed vulnerabilities, over 90% of which were due to implementation problems and 7% due to configuration.

b) SPIT (6 items): The SPIDER project (SPam over Internet telephony Detection sERvice) released a public report [171] providing an overview of SPIT threats and the relevant European legal framework (both on an EU and national basis). The second public report [172] focuses on SPIT detection and prevention, summarizing some of the work done in this space and defining criteria for evaluating the efficiency of anti-SPIT mechanisms. The report classifies prior work according to fulfillment of these criteria, expanding on the relative strengths and weaknesses of each approach.

Dritsas *et al.* [173] and Marias *et al.* [174] survey the risks of SPIT in SIP, the latter also taking into consideration feedback from SIP operators. They then classify a number of previously proposed anti-SPIT mechanisms along a prevent/detect/handle axis. Dritsas *et al.* [175] survey a number of anti-SPIT mechanisms and techniques against a set of criteria that they argue is needed to identify a call as SPIT.

d'Heureuse *et al.* [176] give an overview of the various anti-SPIT efforts in standardization bodies and propose an architecture for dealing with unwanted communications composed of 5 stages: non-intrusive pre-call message analysis, interaction with the caller, pre-connection callee feedback, call content analysis and real-time callee feedback, and post-call callee feedback.

c) Denial of Service (1 item): Sisalem *et al.* [177] give an overview of SIP-based DoS attacks, looking at a couple of specific scenarios. They provide some recommendations to

implementors of VoIP systems that mitigate some of these attacks.

d) P2P SIP (1 item): Seedorf [178] overviews the security challenges in peer-to-peer (P2P) SIP. Threats specific to P2P-SIP include subversion of the identity-mapping scheme (which is specific to the overlay network used as a substrate), attacks on the overlay network routing scheme, bootstrapping communications in the presence of malicious first-contact nodes, identity enforcement (Sybil attacks), traffic analysis and privacy violation by intermediate nodes, and free riding by nodes that refuse to route calls or otherwise participate in the protocol other than to obtain service for themselves.

2) Field Studies and System/Protocol Analysis (12 items): Wieser *et al.* [179] extend the PROTOS testsuite [180] with a SIP-specific analysis fuzzing module. They then test their system against a number of commercial SIP implementations, finding critical vulnerabilities in all of them [181].

Berson [182] conducted an evaluation of the Skype system under contract by Skype itself, allowing him access to the source code. The evaluation focused primarily on the cryptographic protocols and algorithms used, and did not discover any significant issues. Baset and Schulzrinne [183] performed a black-box analysis of Skype, identifying some characteristics of the underlying protocol. Biondi and Desclaux [184] dissected the Skype binary in detail, exposing the extensive anti-reverse engineer and anti-debugging mechanisms built in the program. Their analysis identified a small number of vulnerabilities (including a buffer overflow).

Thermos and Hadsall [185] survey a number of Small Office Home Office (SOHO) VoIP gateways and related equipment, as provided by 3 different commercial VoIP providers with different corporate profiles and customer bases. Their analysis looks at four key factors: manageability, node security, signaling security, and media security. They find numerous problems, including insecure access to the web-based management interface, default passwords and inappropriate services, lack of encryption to protect signaling and media, and low-level implementation issues (*e.g.*, presumed buffer overflow vulnerabilities and fuzzing-induced crashes). A similar survey by Scholz [186] looks at protocol and device problems and vulnerabilities at a medium-size German ISP with high rate of VoIP adoption. He focuses on intentional and unintentional denial of service attacks, problems in customer-premises equipment (*e.g.*, SIP phones), and protocol-independent issues. A number of problems are found, including DoS through call forks, misconfigured devices, and lawful-interception evasion, among others.

INRIA has been conducting a multi-thrust effort to apply testing and fuzzing toward identifying vulnerabilities in SIP protocols [187], implementations [188] and deployed systems [189], [190]. It is worth noting that this work has resulted in a number of vulnerability disclosures in the Common Vulnerabilities and Exposures (CVE) database and elsewhere.

Gupta and Shmatikov [191] formally analyze the security of the VoIP protocol stack, including SIP, SDP, ZRTP, MIKEY, SDES, and SRTP. Their analysis uncovers a number of flaws, most of which derive from subtle inconsistencies in the assumptions made in designing the different protocols. These attacks include a replay attack in SDES that completely break content protection, a man-in-the-middle attack in ZRTP, and a (perhaps theoretical) weakness in the key derivation process used in MIKEY. They also show several minor weaknesses and vulnerabilities in all protocols, primarily enabling denial of service attacks. Floroiu and Sisalem [192] also conduct a comparative analysis of the security aspects of DTLS, ZRTP, MIKEY and SDES. They describe a number of possible attacks against these protocols, and propose mitigation approaches in some cases.

3) Performance Analysis (14 items): Reason and Messerchmitt [193], in one of the earliest works on the subject of the performance impact of security mechanisms on VoIP, looked specifically at the error-expansion properties of encryption and their effect on voice quality. They analytically derive the post-decryption Bit Error Rate (BER) relative to the pre-encryption BER for block and stream ciphers, and analyze the effect of error-expansion mitigation techniques, such as the use of forward error correction, on quality of service. They discuss an error-robust encryption scheme that is analogous to self-synchronizing ciphers.

Elbayoumi and Shepherd [194] conduct a performance comparison of block and stream cipher encryption in the context of securing VoIP calls. They analyze the impact of each on end-to-end delay and subjective quality of perceived voice. A broader view at several performance-impacting parameters is given by the same authors in a concurrent paper at the same journal [195].

Salsano *et al.* [196] give an overview of the various SIP security mechanisms (as of 2002), focusing particularly on the authentication component. They conduct an evaluation of the processing costs of SIP calls that involve authentication, under different transport, authentication and encryption scenarios. They show that a call using TLS and authentication is 2.56 times more expensive than the simplest possible SIP configuration (UDP, no security). However, a fully protected call takes only 54% longer to complete than a configuration that is more representative than the basic one but still offers no security; the same fully-protected call has the same processing cost if it is transported over TCP with no encryption (TLS). Of the overhead, approximately 70% is attributed to message parsing and 30% to cryptographic processing. With the advent of Datagram TLS (DTLS) [197], it is possible that encryption and integrity for SIP can be had for all configurations (UDP or TCP) at no additional cost.

Barbieri *et al.* [198] find that when using VoIP over IPsec, performance can drop by up to 63%; however, it is questionable whether these results still hold, given the use of hardware accelerators and the more efficient AES algorithm in IPsec. Simulation-based work by Ranganathan and Kilmartin [199] shows that the use of IPsec with pre-established Security Associations (SAs) increases SIP call setup time by 1.4% and media (voice) transfer by 1.6%. However, when taking into consideration the delay in establishing SAs for the first time using a dynamic key-agreement protocol such as IKE [200] or IKEv2 [201], the call setup delay can increase dramatically. They identify encryption engine queuing delays as a potential concern, as call volumes increase.

A conclusion similar to Salsano *et al.* [196] is reached by Bilien [202] and Bilien *et al.* [203], [204], who study the overhead in SIP call setup latency when using end-to-end and hop-by-hop security mechanisms. They consider protocols such as MIKEY, S/MIME, SRTP, TLS, and IPsec, concluding that the overall penalty of using full-strength cryptography is low.

Xiao and Zarrella [205] conduct an experimental evaluation of the impact of security mechanisms on VoIP in wireless environments with a specific voice codec. They specifically look at how the use of IPsec and WEP affect the Mean Opinion Score, packet loss, and delay of VoIP calls in 802.11 networks. They find that WEP has a bigger impact on packet loss than IPsec, but the latter can cause larger packet delays and fewer but more extreme voice artifacts (disturbances) in the call.

Also in the context of VoIP for wireless networks, Lakay and Agbinya [206] summarize similar experiments that show SIP security mechanism processing is responsible for 80% of the call setup delay when using stateless proxies, and 45% for stateful proxies.

Eun-Chul *et al.* [207], evaluate via simulation the costs of different security protocols (TLS, DTLS and IPsec) with respect to call setup delay using different transport protocols (TCP, UDP and SCTP). They conclude that the most efficient combinations, DTLS/UDP and IPsec/UDP, approximately double the call setup delay. However, since the analysis is purely simulation-based, their results are sensitive to the configured relative costs for processing the various protocols.

Shen *et al.* [208] also study the performance impact of using TLS as a transport protocol for SIP. In their experiments using a testbed, they use profiling at various system levels (application, library, and kernel), and decompose the costs at a fine level of granularity. They determine that use of TLS can reduce performance by a factor of up to 20 (when compared with the unsecured SIP-over-UDP). The main overhead factor is the cost of RSA signatures during session negotiation, while symmetric key operations impose a relatively small cost. They recommend that operators amortize the setup cost over long-lived connections. Finally, they provide a cost model for provisioning SIP-over-TLS servers, predicting an average performance overhead of 15% under a suggested system configuration.

Rebahi *et al.* [209] analyze the performance of RSA as used in SIP for authentication and identity management (via public-key certificates and digital signatures), and describe the use of Elliptic Curve DSA (ECDSA) within this context to improve performance. Using ECDSA, their prototype can handle from 2 to 8 times as many call setup requests per second, with the gap widening as key sizes increase.

4) Authentication Protocols (15 items): Buschel [210] argues for integrated authentication between User Agents and all elements of a SIP infrastructure. Over the years, a number of authentication schemes aiming to replace Digest Authentication have been proposed, using such basic blocks as Diffie Hellman [211], Elliptic Curve Diffie Hellman (ECDH) [212], Elliptic Curve Discrete Logarithm Problem (ECDLP) [213], nonces [214], PKI [215], [216] hash functions [217], and others [218], not all of them secure [219].

Cao and Jennings [220] propose a new mechanism for authenticating the responding user's identity in SIP without exposing said identity to untrusted intermediate elements. Their scheme requires additional headers in SIP messages, and has not been implemented or evaluated.

Insu and Keecheon [221] propose a secret key based mechanism to reduce the performance requirements of using public key certificates to protect signaling (*e.g.,* with TLS) in an enterprise VoIP environment.

Schmidt *et al.* [222] suggest that administration overheads for implementing strong authentication in SIP could be lowered by grouping users with the same function or role (*e.g.,* agents in a calling center). They propose a proxy-based mechanism for implementing a form of "certificate sharing" among a group of users, without exposing the corresponding private key to any of them. They demonstrate feasibility of the scheme by implementing it in the NIST SIP proxy, with no further evaluation.

Wang and Zhang [223] discuss an authentication and key agreement mechanism for SIP that uses certificate-less public-key cryptography. Certificate-less public-key cryptography [224] is a variant of identity-based cryptography (where the public key of an entity is its public identity); here, the public key for an entity is generated collaboratively between that entity and a trusted third party in such a way that the public key can be verified by any other entity that knows the public parameters under which the trusted third party operates. Compared to previous proposals that used identity-based cryptography [225], their scheme does not require that the trusted third party

5) Architectures (19 items): Singh and Vuong [226] use a mobile agent framework to collect and correlate events from various network components, toward detecting a number of attacks. The stated advantages of their approach are that it does not require a new protocol for exchanging event information and that mitigation and recovery capabilities can be implemented by extending the framework and the agents, with no changes to the VoIP protocols. They also propose using user behavior profiles to detect anomalous behavior. They describe the operation of their system in a number of attack scenarios, including protocol-based denial of service, call hijacking, packet flooding, and abnormal call patterns.

Casola *et al.* [227], [228] suggest the use of a policy-based approach to design secure VoIP infrastructures. The policies express security goals in measurable terms; suggested infrastructure designs can then be evaluated against these policies to determine whether the goals are met to an acceptable degree.

Wu *et al.* [229] design an intrusion detection system, called SCIDIVE, that is specific to VoIP environments. Specifically, SCIDIVE aims to detect different classes of intrusions, can operate with different viewpoints (on clients, proxies, or servers), and takes into consideration both signaling (*i.e.,* SIP) and media-transfer protocols (*e.g.,* RTP). SCIDIVE's ability to correlate cross-protocol behavior, theoretically allows for detection of more complex attacks. However, the system is rules-based, which limits its effectiveness against new/unknown attacks. The primary evaluation (conducted on a small testbed) consists of four simple cross-protocol attacks,

which would have evaded other contemporary, non-specialized intrusion detection systems. In follow-on work, Apte *et al.* [230], [231] develop SPACEDIVE, a VoIP-specific intrusion detection system that allows for correlation of events among distributed rules-based detectors. They demonstrate the ability of SPACEDIVE to detect certain classes of attacks using a simple SIP environment with two domains, and compare it with SCIDIVE.

Martin and Hung [232] discuss a high-level policy for VoIP applications, intended to guide the implementation, configuration, and use of VoIP systems.

SNOCER[4], a project funded by the European Union, is "investigating approaches for overcoming temporal network, hardware and software failures and ensuring the high availability of the offered VoIP services based on low cost distributed concepts." The first public project report [233] provides an overview of VoIP infrastructure components and the threats that must be addressed (staying primarily at the protocol and network level, and avoiding implementation issues with the exception of SQL injection), along with possible defense mechanisms. There is also discussion on scalable service provisioning (replication, redundancy, backups *etc.*), toward providing reliability and fault tolerance. The second public project report [234] describes an architecture for protecting against malformed messages and related attacks using specification-based intrusion detection, protocol message verification, and redundancy. They use ontologies to describe SIP vulnerabilities, to allow for easy updating of the monitoring components (IDS) [235].

Niccolini *et al.* [236] design an intrusion detection/intrusion prevention system architecture for use with SIP. Their system uses both knowledge-based and behavior-based detection, arranged as a series in that order. They develop a prototype implementation using the open-source Snort IDS. They evaluate the effectiveness of their system in an attack scenario by measuring the mean end-to-end delay of legitimate SIP traffic in the presence of increasing volumes of malformed SIP INVITE messages.

Marshall *et al.* [237] describe the AT&T VoIP security architecture. They divide VoIP equipment into three classes: trusted, trusted-but-vulnerable, and untrusted. The latter consists of the customer premises equipment, which is outside the control of the carrier. The trusted domain includes all the servers necessary to provide VoIP service. Between the two sit various border and security elements, that are responsible for protecting the trusted devices while permitting legitimate communications to proceed. They describe the interactions among the various components, and the security mechanisms used in protecting these interactions.

Sher and Magedanz [238] describe a security architecture for IMS service delivery platforms, focusing on time-independent attacks (*e.g.,* software vulnerabilities). The key element of their proposed approach is an intrusion detection and prevention system that inspects all incoming and outgoing SIP messages to the IMS application servers, applying rules that detect and mitigate specific attacks. A brief performance evaluation shows that a prototype can operate with acceptable delay parameters.

Ding and Su [239] propose the combination of specification-based intrusion detection with anomaly detection techniques and attack-specific methods using hierarchical colored Petri nets.

Nassar *et al.* [240] advocate the use of SIP-specific honeypots to catch attacks targeting the Internet telephony systems, protocols and applications. They design and implement such a honeypot system, and explore the use of a statistical engine for identifying attacks and other misbehavior, based on training on legitimate traces of SIP traffic. The engine is based on their prior work that uses Bayesian-based inference. The resulting SIP honeypot effort is largely exploratory, with performance and effectiveness evaluations left for future work. In follow-on work, Nassar *et al.* [241] describe an intrusion detection and prevention architecture for VoIP that integrates SIP honeypots and an application-layer event correlation engine.

Barry and Chan [242] describe a host-based intrusion detection architecture for SIP that combines specification-based and signature-based detection, and allows for the correlation of information across modules to identify cross-protocol attacks. They conduct a simulation-based evaluation using OMNeT++ to determine detection accuracy and performance impact.

Rieck *et al.* [243] apply machine learning techniques to detecting anomalous SIP messages, incorporating a "self-learning" component by allowing for periodic re-training of the anomaly detector using traffic that has been flagged as normal. The features used for clustering are based on n-grams and on tokenization of the SIP protocol. To prevent training attacks, wherein an adversary "trains" the anomaly detector to accept malicious inputs as legitimate, they employ randomization (choosing random samples for the training set), sanitization [244], and verification (by comparing the output of the new and old training models). Their experimental prototype was shown to handle 70 Mbps of SIP traffic, while providing a 99% detection rate with no false positives.

Dantu *et al.* [245] describe a comprehensive VoIP security architecture, composed of components distributed across the media gateway controller, the proxy server(s), the IP PBX, and end-user equipment. These components explicitly exchange information toward better training of filters, and creating and maintaining whitelists/blacklists. Implicit feedback is also provided through statistical analysis of interactions (*e.g.,* call frequency and duration). The architecture also provisions for a recovery mechanism that incorporates explicit feedback and quarantining.

6) Middleboxes (11 items): Reynolds and Ghosal [246] describe a VoIP-aware middlebox architecture that integrates the enterprise firewall, media gateway, and intrusion detection facilities to allow the secure operation of dynamic VoIP applications. The problem of firewall and NAT traversal by VoIP protocols has been the subject of some research [247], [248], [249], [250], generally involving some kind of signaling (whether in-band or out-of-band) between the end-device and the middlebox.

Bessis *et al.* [251] discuss the necessary features of a SIP-specific firewall, juxtaposing them with specific threats to SIP

[4]http://www.snocer.org/

messages at each network layer (data link, network, transport and session). They propose a simple, hardware-accelerated SIP-proxy as a front-end SIP firewall and argue that this approach would block most of the attacks.

Gurbani *et al.* [252] propose an mechanism whereby proxies create an overlay network between user agents. This network is used for rendezvous/coordination purposes only. Once user agents establish a session, the proxies become transparent traffic forwarders, with the user agents communicating over an end-to-end secure session. This approach allows users to communicate without exposing (as much) private information to proxies, at the cost of requiring a PKI and a new message extension.

Sengar *et al.* [253], [254], [255] examine the problem of cross-infrastructure vulnerabilities created by bridging VoIP and PSTN networks. They outline a high-level architecture that integrates firewall-like functionality with trust management, signaling encryption and authentication, and intrusion detection.

Ehlert *et al.* [256] describe a rule-reduction algorithm for improving the performance of firewalls operating in busy VoIP environments, in balance with security requirements. Their algorithm works by merging similar single-mapped rules into a more general rule, then dropping less important rules, and finally calculating the accuracy of the new ruleset. If needed, their algorithm re-iterates until an acceptable solution is achieved.

7) Intrusion Detection (11 items): Mandjes *et al.* [257], [258] describe the use of statistical techniques to identify anomalies in VoIP networks. Their work is primarily directed at non-adversarial anomalies, although certain attacks (such as denial of service) would also be detected by their scheme.

Geneiatakis *et al.* [259], [260] discuss malformed-message attacks against SIP servers and equipment, primarily depending on the PROTOS testsuite for SIP implementations [180]. To detect such attacks, they propose building an intrusion detection system that leverages the SIP syntax grammar [8] to decompose incoming messages, and a grammar for specifying rules that check whether specific constraints are being violated (or specific conditions met) [261], [262]. In subsequent work, Geneiatakis and Keromytis [263] apply entropy theory and "itself information" to the problem of identifying anomalies in a stream of SIP messages.

However, Hantehzadeh *et al.* [264] point out that the most approaches to anomaly detection in SIP use datasets with large differences between anomalous and normal messages, which make them easy to detect. An analysis using a dataset with minimal such differences (while maintaining the distinction between malicious and normal messages) indicates that existing classification schemes do not perform as well. They propose feature reduction techniques to enhance these classification schemes even on "trickier" datasets. Similar results, focusing specifically on the performance of classifiers using Euclidean distance, are discussed by Mehta *et al.* [265].

Sengar *et al.* [255] model the protocol state machine of individual SIP nodes (derived from the SIP specification) and inter-node interactions, in order to have a complete picture of the overall system state towards detecting anomalous behavior and attacks. This is particularly important in VoIP, since nodes can interact in many ways, and with several other nodes during a call and throughout their operation. They conduct a performance evaluation to determine the overhead added to call setup and media transfer by their system, and its overall scalability. While their system can identify known attacks (for which attack patterns can be specified) with high accuracy and low false positives, detecting previously unknown attacks depends on the fidelity of the protocol state machines. This problem is left for future work.

Seo *et al.* [266] develop a stateful intrusion detection system for SIP, modeling SIP state transitions to match the expected state of the monitored SIP entities. Their system allows the specification of rules that match attacks and misbehavior based not only on the content of the communications but also on the state of the SIP call and of the proxies.

8) Miscellaneous (2 items): Cao *et al.* [267] describe how to transparently add information in SIP and H.323 messages such that calls can be tracked across the network. A similar approach, leveraging watermarking of VoIP content, was previously described by Steinebach *et al.* [268].

V. Discussion

In our previous work [1], [2], [3], we surveyed 215 vulnerabilities in SIP implementations that had been disclosed in the CVE database from 1999 to 2009. We classified these vulnerabilities along several dimensions, including the VoIPSA threat taxonomy, the traditional {Confidentiality, Integrity, Availability} concerns, and a {Protocol, Implementation, Configuration} axis. We found that the various types of denial of service attacks constitute the majority of disclosed vulnerabilities, over 90% of which were due to implementation problems and 7% due to configuration.

Considering the research work we have surveyed (some of which was discussed in this paper), we can see that out of a total of 245 publications, almost 20% concern themselves with an overview of the problem space and of solutions — a figure we believe is reasonable, considering the enormity of the problem space and the speed of change in the protocols, standards, and implementations. We also see a considerable amount of effort (roughly 20%) going toward addressing SPIT. While SPIT does not appear to be a major issue with VoIP users at this point, our past and current experiences with email spam and telemarketing seems to provide sufficient motivation for research in this area. Most of the work is focused on identifying SPIT calls and callers based on behavioral traits, although a number of other approaches are under exploration (*e.g.,* CAPTCHAs and real-time content analysis). One of the problems is the lack of a good corpus of data for experimentation and validation of the proposed techniques.

We were also not surprised to see a sizable portion of research (over 15%) directed at design, analysis (both security- and performance-oriented), and attacking of cryptographic protocols as used in VoIP. The cryptographic research community appears to be reasonably comfortable in proposing tweaks and minor improvements to the basic authentication mechanisms, and the systems community appears content with

analyzing the performance of different protocol configurations (*e.g.*, TLS *vs.* IPsec).

Most distressing, however, is the fact that comparatively little research (less than 13%) is going toward addressing the problem of denial of service. Given the numerical dominance of SIP-specific DoS vulnerabilities (as described earlier) and the ease of launching such attacks, it is clear that significantly more work is needed here. What work is being done seems to primarily focus on the server and infrastructure side, despite our finding that half of DoS-related vulnerabilities are present on endpoints. Furthermore, much of the existing work focuses on network-observable attacks (*e.g.*, "obviously" malformed SIP messages), whereas the majority of VoIP DoS vulnerabilities are the result of implementation failures. More generally, additional work is needed in strengthening implementations, rather than introducing middleboxes and network intrusion detection systems, whose effectiveness has been shown to be limited in other domains; taking a black box approach in securing VoIP systems is, in our opinion, not going to be sufficient.

Also disconcerting is the lack of research (2.8%) in addressing service abuse threats, considering the visibility of large fraud incidents [4], [5], [6].

In general, we found little work that took a "big picture" view of the VoIP security problem. What cross-cutting architectures have been proposed focus primarily on intrusion detection. Work is needed to address cross-implementation and cross-protocol problems, above and beyond the few efforts along those lines in the intrusion detection space.

Finally, we note that none of the surveyed works addressed the problem of configuration management. While such problems represent only 7% of known vulnerabilities, configuration issues are easy to overlook and are likely under-represented in our previous analysis due to the nature of vulnerability reporting.

VI. Conclusions

We have presented a survey of 245 publications on the topic of VoIP security, classifying them according to the VoIPSA threat taxonomy. We juxtaposed this survey against our previous analysis on VoIP security vulnerabilities. We identified two specific areas (denial of service and service abuse) as being under-represented in research efforts directed at them (relative to their importance in the vulnerability survey). Furthermore, we identify implementation bugs and misconfigurations as two general problem areas that merit considerably more work than they currently attract. We hope that our work will ease the task of conducting research in VoIP security and help guide the often disjoint research efforts.

References

[1] A. D. Keromytis, "Voice over IP: Risks, Threats and Vulnerabilities," in *Proceedings of the Cyber Infrastructure Protection (CIP) Conference*, June 2009.

[2] A. D. Keromytis, "A Look at VoIP Vulnerabilities," *USENIX ;login: Magazine*, vol. 35, pp. 41–50, February 2010.

[3] A. D. Keromytis, "Voice over IP Security: Research and Practice," *IEEE Security & Privacy Magazine*, vol. 8, pp. 76–78, March/April 2010.

[4] B. Krebs, "Security Fix: Default Passwords Led to $55 Million in Bogus Phone Charges," June 2009.

[5] The Register, "Two charged with VoIP fraud." http://www.theregister.co.uk/2006/06/08/voip_fraudsters_nabbed/, June 2006.

[6] The Register, "Fugitive VOIP hacker cuffed in Mexico." http://www.theregister.co.uk/2009/02/11/fugitive_voip_hacker_arrested/, February 2009.

[7] VoIP Security Alliance, "VoIP Security and Privacy Threat Taxonomy, version 1.0." http://www.voipsa.org/Activities/taxonomy.php, October 2005.

[8] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, and E. Schooler, "SIP: Session Initiation Protocol." RFC 3261 (Proposed Standard), June 2002. Updated by RFCs 3265, 3853, 4320, 4916, 5393.

[9] J. Franks, P. Hallam-Baker, J. Hostetler, S. Lawrence, P. Leach, A. Luotonen, and L. Stewart, "HTTP Authentication: Basic and Digest Access Authentication." RFC 2617 (Draft Standard), June 1999.

[10] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications." RFC 3550 (Standard), July 2003. Updated by RFC 5506.

[11] I. Johansson and M. Westerlund, "Support for Reduced-Size Real-Time Transport Control Protocol (RTCP): Opportunities and Consequences." RFC 5506 (Proposed Standard), Apr. 2009.

[12] J. Postel, "Transmission Control Protocol." RFC 793 (Standard), Sept. 1981. Updated by RFCs 1122, 3168.

[13] J. Postel, "User Datagram Protocol." RFC 768 (Standard), Aug. 1980.

[14] L. Ong and J. Yoakum, "An Introduction to the Stream Control Transmission Protocol (SCTP)." RFC 3286 (Informational), May 2002.

[15] E. Rescorla and N. Modadugu, "Datagram Transport Layer Security." RFC 4347 (Proposed Standard), Apr. 2006.

[16] M. Handley, E. Rescorla, and IAB, "Internet Denial-of-Service Considerations." RFC 4732 (Informational), Dec. 2006.

[17] M. Handley, V. Jacobson, and C. Perkins, "SDP: Session Description Protocol." RFC 4566 (Proposed Standard), July 2006.

[18] B. Ramsdell, "Secure/Multipurpose Internet Mail Extensions (S/MIME) Version 3.1 Message Specification." RFC 3851 (Proposed Standard), July 2004.

[19] S. Kent and K. Seo, "Security Architecture for the Internet Protocol." RFC 4301 (Proposed Standard), Dec. 2005.

[20] K. Srivastava and H. Schulzrinne, "Preventing Spam For SIP-based Instant Messages and Sessions," Technical Report CUCS-042-04, Columbia University, Department of Computer Science, October 2004.

[21] R. Dantu and P. Kolan, "Preventing Voice Spamming," in *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM), Workshop on VoIP Security Challenges and Solutions*, December 2004.

[22] R. MacIntosh and D. Vinokurov, "Detection and Mitigation of Spam in IP Telephony Networks Using Signaling Protocol Analysis," in *Proceedings of the IEEE/Sarnoff Symposium on Advances in Wired and Wireless Communication*, pp. 49–52, April 2005.

[23] N. Croft and M. Olivier, "A Model for Spam Prevention in Voice over IP Networks using Anonymous Verifying Authorities," in *Proceedings of the 5^{th} Annual Information Security South Africa Conference (ISSA)*, July 2005.

[24] R. Dantu and P. Kolan, "Detecting Spam in VoIP Networks," in *Proceedings of the USENIX Workshop on Steps to Reducing Unwanted Traffic on the Internet (SRUTI)*, pp. 31–37, July 2005.

[25] P. Kolan and R. Dantu, "Socio-technical Defense Against Voice Spamming," *ACM Transactions on Autonomous and Adaptive Systems (TAAS)*, vol. 2, March 2007.

[26] Y. Rebahi and D. Sisalem, "SIP Service Providers and the Spam Problem," in *Proceedings of the 2^{nd} VoIP Security Workshop*, June 2005.

[27] Y. Rebahi, D. Sisalem, and T. Magedanz, "SIP Spam Detection," in *Proceedings of the International Conference on Digital Telecommunications(ICDT)*, pp. 29–31, August 2006.

[28] M. Hansen, M. Hansen, J. Möller, T. Rohwer, C. Tolkmit, and H. Waack, "Developing a Legally Compliant Reachability Management System as a Countermeasure against SPIT," in *Proceedings of the 3^{rd} Workshop on Securing Voice over IP*, June 2006.

[29] R. Baumann, S. Cavin, and S. Schmid, "Voice Over IP - Security and SPIT," KryptDet Report FU Br 41, Swiss Army, August/September 2006.

[30] A. Madhosingh, "The Design of a Differentiated SIP to Control VoIP Spam," Masters Thesis Report SPIT, CAPTCHA, Florida State University, Computer Science Department, 2006.

[31] M. Bertrand, Q. Loudier, Y. Gourhant, F. Bougant, and M. Osty, "SPIT Mitigation by a Network-Level Anti-Spit Entity," in *Proceedings of the 3^{rd} Workshop on Securing Voice over IP*, June 2006.

[32] H. Yan, H. Zhang, K. Sripanidkulchai, Z. Shae, and D. Saha, "Incorporating Active Fingerprinting into SPIT Prevention Systems," in *Proceedings of the 3^{rd} Workshop on Securing Voice over IP*, June 2006.

[33] V. Balasubramaniyan, M. Ahamad, and H. Park, "CallRank: Combating SPIT Using Call Duration, Social Networks and Global Reputation," in *Proceedings of the 4^{th} Conference on Email and Anti-Spam (CEAS)*, August 2007.

[34] K. Ono and H. Schulzrinne, "Have I Met You Before? Using Cross-Media Relations to Reduce SPIT," in *Proceedings of the 3^{rd} International Conference on Principles, Systems and Applications of IP Telecommunications (IPTComm)*, pp. 1–7, June 2009.

[35] H. Guang-Yu, W. Y.-Y. Wen, and H. Zhao, "SPIT Detection and Prevention Method Based on Signal Analysis," in *Proceedings of the 3^{rd} International Conference on Convergence and Hybrid Information Technology (ICCIT)*, vol. 2, pp. 631–638, November 2008.

[36] P. Patankar, G. Nam, G. Kesidis, and C. R. Das, "Exploring Anti-Spam Models in Large Scale VoIP Systems," in *Proceedings of the 28^{th} International Conference on Distributed Computing Systems (ICDCS)*, pp. 85–92, June 2008.

[37] Y.-S. Wu, S. Bagchi, N. Singh, and R. Wita, "Spam Detection in Voice-Over-IP Calls through Semi-Supervised Clustering," in *Proceedings of the 39^{th} Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*, pp. 307–316, June 2009.

[38] K. Hyung-Jong, K. M. Joo, K. Yoonjeong, and J. H. Cheol, "DEVS-based Modeling of VoIP Spam Callers' Behavior for SPIT Level Calculation," *Simulation Modeling Practice and Theory*, vol. 17, pp. 569–584, April 2009.

[39] C. Sorge and J. Seedorf, "A Provider-Level Reputation System for Assessing the Quality of SPIT Mitigation Algorithms," in *Proceedings of the IEEE Internation Conference on Communications (ICC)*, pp. 1–6, June 2009.

[40] S. Phithakkitnukoon and R. Dantu, "Defense Against SPIT Using Community Signals," in *Proceedings of the IEEE International Conference on Intelligence and Security Informatics (ISI)*, June 2009.

[41] C. Pörschmann and H. Knospe, "Analysis of Spectral Parameters of Audio Signals for the Identification of Spam Over IP Telephony," in *Proceedings of the 5^{th} Conference on Email and Anti-Spam (CEAS)*, August 2008.

[42] H. Tschofenig, R. Falk, J. Peterson, J. Hodges, D. Sicker, and J. Polk, "Using SAML to Protect the Session Initiation Protocol (SIP)," *IEEE Network*, vol. 20, pp. 14–17, September/October 2006.

[43] N. d'Heureuse, J. Seedorf, and S. Niccolini, "A Policy Framework for Personalized and Role-Based SPIT Prevention," in *Proceedings of the 3^{rd} International Conference on Principles, Systems and Applications of IP Telecommunications (IPTCOMM)*, July 2009.

[44] Y. Soupionis, S. Dritsas, and D. Gritzalis, "An Adaptive Policy-Based Approach to SPIT Management," in *Proceedings of the 13^{th} European Symposium on Research in Computer Security (ESORICS)*, pp. 446–460, October 2008.

[45] N. Banerjee, S. Saklikar, and S. Saha, "Anti-vamming Trust Enforcement in Peer-to-peer VoIP Networks," in *Proceedings of the International Conference on Communications and Mobile Computing (IWCMC)*, pp. 201–206, July 2006.

[46] J. Quittek, S. Niccolini, S. Tartarelli, M. Stiemerling, M. Brunner, and T. Ewald, "Detecting SPIT Calls by Checking Human Communication Patterns," in *Proceedings of the IEEE International Conference on Communications (ICC)*, pp. 1979–1984, June 2007.

[47] T. Wang, "A VoIP anti-Spam System based on Reverse Turing Test," Masters Thesis ETD-05072007-173147, North Carolina State University, May 2007.

[48] J. L. Janne and M. Komu, "Cure for Spam Over Internet Telephony," in *Proceedings of the 4^{th} IEEE Consumer Communications and Networking Conference (CCNC)*, pp. 896–900, January 2007.

[49] S. Niccolini, "SPIT Prevention: State of the Art and Research Challenges," in *Proceedings of the 3^{rd} Workshop on Securing Voice over IP*, June 2006.

[50] M. Haberler and O. Lendl, "Secure Selective Peering with Federations," in *Proceedings of the 3^{rd} Workshop on Securing Voice over IP*, June 2006.

[51] S. Saklikar and S. Saha, "Identity Federation for VoIP-based Services," in *Proceedings of the ACM Workshop on Digital Identity Management*, pp. 62–71, November 2007.

[52] D. Shin and C. Shim, "Voice Spam Control with Gray Leveling," in *Proceedings of the 2^{nd} VoIP Security Workshop*, June 2005.

[53] D. Shin, J. Ahn, and C. Shim, "Progressive Multi Gray-Leveling: A Voice Spam Protection Algorithm," *IEEE Network*, vol. 20, pp. 18–24, September/October 2006.

[54] R. Schlegel, S. Niccolini, S. Tartarelli, and M. Brunner, "SPam over Internet Telephony (SPIT) Prevention Framework," in *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM)*, pp. 1–6, November/December 2006.

[55] J. Quittek, S. Niccolini, S. Tartarelli, and R. Schlegel, "Prevention of Spam over IP Telephony (SPIT)," *NEC Technical Journal*, vol. 1, no. 2, pp. 114–119, 2006.

[56] J. Quittek, S. Niccolini, S. Tartarelli, and R. Schlegel, "On Spam over Internet Telephony (SPIT) Prevention," *IEEE Communications Magazine*, vol. 46, pp. 80–86, August 2008.

[57] Y. Rebahi, S. Ehlert, S. Dritsas, G. F. Marias, D. Gritzalis, B. Pannier, O. Capsada, T. Golubenco, J. F. Juell, and M. Hoffmann, "General Anti-Spam Security Framework for VoIP Infrastructures," Tech. Rep. Deliverable WP2/D2.3, SPIDER COOP-32720, July 2007.

[58] B. Mathieu, S. Niccolini, and D. Sisalem, "SDRS: A Voice-over-IP Spam Detection and Reaction System," *IEEE Security & Privacy Magazine*, vol. 6, pp. 52–59, November/December 2008.

[59] D. Gritzalis and Y. Mallios, "A SIP-oriented SPIT Management Framework," *Computers & Security*, vol. 27, pp. 136–153, October 2008.

[60] P. Kolan, R. Dantu, and J. W. Cangussu, "Nuisance of a Voice Call," *ACM Transactions on Multimedia Computing, Communications and Applications (TOMCCAP)*, vol. 5, pp. 6:1–6:22, October 2008.

[61] S. Dritsas, V. Dritsou, B. Tsoumas, P. Constantopoulos, and D. Gritzalis, "OntoSPIT: SPIT management through ontologies," *Computer Communications*, vol. 32, pp. 203–212, January 2009.

[62] L. Kong, V. B. Balasubramaniyan, and M. Ahamad, "A Lightweight Scheme for Securely and Reliably Locating SIP Users," in *Proceedings of the 1^{st} IEEE Workshop on VoIP Management and Security (VoIP MaSe)*, pp. 9–17, April 2006.

[63] X.Wang, S. Chen, and S. Jajodia, "Tracking Anonymous Peer-to-Peer VoIP Calls on the Internet," in *Proceedings of the 12^{th} ACM Conference on Computer and Communications Security (CCS)*, pp. 81–91, November 2005.

[64] S. Chen, X. Wang, and S. Jajodia, "On the Anonymity and Traceability of Peer-to-Peer VoIP Calls," *IEEE Network*, vol. 20, pp. 32–37, September/October 2006.

[65] M. Srivatsa, A. Iyengar, and L. Liu, "Privacy in VoIP networks: A k-Anonymity Approach," in *Proceedings of the 28^{th} IEEE Conference on Computer Communication (INFOCOM)*, pp. 2856–2860, April 2009.

[66] G. Shah, A. Molina, and M. Blaze, "Keyboards and Covert Channels," in *Proceedings of 15^{th} USENIX Security Symposium*, pp. 59–75, July/August 2006.

[67] T. Takahashi and W. Lee, "An Assessment of VoIP Covert Channel Threats," in *Proceedings off the 3^{rd} International Conference on Security and Privacy in Communications Networks (SecureComm)*, pp. 371–380, September 2007.

[68] C. Wieser, J. Röning, and A. Takanen, "Security analysis and experiments for Voice over IP RTP media streams," in *Proceedings of the 8^{th} International Symposium on Systems and Information Security (SSI)*, November 2006.

[69] C. V. Wright, L. Ballard, F. N. Monrose, and G. M. Masson, "Language Identification of Encrypted VoIP Traffic: Alejandra y Roberto or Alice and Bob?," in *Proceedings of 16^{th} USENIX Security Symposium*, pp. 1–12, August 2007.

[70] C. V. Wright, L. Ballard, S. Coulls, F. N. Monrose, and G. M. Masson, "Spot Me If You Can: Recovering Spoken Phrases in Encrypted VoIP Conversations," in *Proceedings of IEEE Symposium on Security and Privacy*, pp. 35–49, May 2008.

[71] X. Wang, R. Zhang, X. Yang, X. Jiang, and D. Wijesekera, "Voice Pharming Attack and the Trust of VoIP," in *Proceedings of the 4^{th} International Conference on Security and Privacy in Communication Networks (SecureComm)*, pp. 1–11, September 2008.

[72] 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 *Proceedings of the 6^{th} International Conference on Data Mining (ICDM)*, p. 667=677, December 2006.

[73] M. Petraschek, T. Hoeher, O. Jung, H. Hlavacs, and W. N. Gansterer, "Security and Usability Aspects of Man-in-the-Middle Attacks on ZRTP," *Journal of Universal Computer Science*, vol. 14, no. 5, pp. 673–692, 2008.

[74] R. Zhang, X. Wang, R. Farley, X. Yang, and X. Jiang, "On the Feasibility of Launching the Man-In-The-Middle Attacks on VoIP from Remote Attackers," in *Proceedings of the 4^{th} International ACM Symposium on Information, Computer, and Communications Security (ASIACCS)*, pp. 61–69, March 2009.

[75] J.-I. Guo, J.-C. Yen, and H.-F. Pai, "New Voice over Internet Protocol Technique with Hierarchical Data Security Protection," *IEE Proceedings — Vision, Image and Signal Processing*, vol. 149, pp. 237–243, August 2002.

[76] C. Li, S. Li, D. Zhang, and G. Chen, "Cryptanalysis of a Data Security Protection Scheme for VoIP," *IEE Proceedings—Vision, Image and Signal Processing*, vol. 153, pp. 1–10, February 2006.

[77] S. M. Bellovin, M. Blaze, and S. Landau, "The Real National-Security Needs for VoIP," *Communications of the ACM (CACM)*, vol. 48, p. 120, November 2005.

[78] J. Seedorf, "Using Cryptographically Generated SIP-URIs to Protect the Integrity of Content in P2P-SIP," in *Proceedings of the 3^{rd} Workshop on Securing Voice over IP*, June 2006.

[79] A. Fessi, N. Evans, H. Niedermayer, and R. Holz, "Pr2-P2PSIP: Privacy Preserving P2P Signaling for VoIP and IM," in *Proceedings of the 4^{th} Annual ACM Conference on Principles, Systems and Applications of IP Telecommunications (IPTCOMM)*, pp. 141–152, August 2010.

[80] A. Talevski, E. Chang, and T. Dillon, "Secure Mobile VoIP," in *Proceedings of the International Conference on Convergence Information Technology*, pp. 2108–2113, November 2007.

[81] N. Kuntze, A. U. Schmidt, and C. Hett, "Non-Repudiation in Internet Telephony," in *Proceedings of the IFIP International Information Security Conference*, pp. 361–372, May 2007.

[82] C.-H. Wang, M.-W. Li, and W. Liao, "A Distributed Key-Changing Mechanism for Secure Voice over IP (VoIP) Service," in *Proceedings of the IEEE International Conference on Multimedia and Expo*, pp. 895–898, July 2007.

[83] V. K. Gurbani and V. Kolesnikov, "Work in Progress: A secure and lightweight scheme for media keying in the Session Initiation Protocol (SIP)," in *Proceedings of the 4^{th} Annual ACM Conference on Principles, Systems and Applications of IP Telecommunications (IPTCOMM)*, pp. 35–44, August 2010.

[84] H. Hlavacs, W. N. Gansterer, H. Schabauer, J. Zottl, M. Petraschek, T. Hoeher, and O. Jung, "Enhancing ZRTP by using Computational Puzzles," *Journal of Universal Compter Science*, vol. 14, no. 5, pp. 693–716, 2008.

[85] F. Palmieri and U. Fiore, "Providing True End-to-End Security in Converged Voice over IP Infrastructures," *Computers & Security*, vol. 28, pp. 433–449, September 2009.

[86] G. Zhang and S. Berthold, "Hidden VoIP Calling Records from Networking Intermediaries," in *Proceedings of the 4^{th} Annual ACM Conference on Principles, Systems and Applications of IP Telecommunications (IPTCOMM)*, pp. 15–24, August 2010.

[87] G. Zhang and S. Fischer-Hübner, "Peer-to-Peer VoIP Communications Using Anonymisation Overlay Networks," in *Proceedings of the 11^{th} Conference on Communications and Multimedia Security (CMS)*, May/June 2010.

[88] C. A. Melchor, Y. Deswarte, and J. Iguchi-Cartigny, "Closed-circuit Unobservable Voice over IP," in *Proceedings of the 23^{rd} Annual Computer Security Applications Conference (ACSAC)*, pp. 119–128, December 2007.

[89] M. Srivatsa, L. Liu, and A. Iyengar, "Preserving Caller Anonymity in Voice-over-IP Networks," in *Proceedings of the IEEE Symposium on Security and Privacy (S&P)*, pp. 50–63, May 2008.

[90] A. D. Elbayoumy and S. Shepherd, "A High Grade Secure VoIP System Using the Tiny Encryption Algorithm," in *Proceedings of the 7^{th} Annual International Symposium on Advanced Radio Technologies*, pp. 342–350, March 2005.

[91] A. D. Elbayoumy and S. Shepherd, "QoS Control Using an Endpoint CPU Capability Detector in a Secure VoIP System," in *Proceedings of the 10^{th} IEEE Symposium on Computers and Communications*, pp. 175–181, June 2005.

[92] A. D. Elbayoumy and S. Shepherd, "A High Grade Secure VoIP System Using an Endpoint CPU Capability Detector," in *Proceedings of the ITA05 International Conference on Internet Technologies and Applications*, pp. 173–180, September 2005.

[93] B. Reynolds and D. Ghosal, "Secure IP Telephony using Multi-layered Protection," in *Proceedings of the ISOC Symposium on Network and Distributed Systems Security (NDSS)*, February 2003.

[94] J. Larson, T. Dawson, M. Evans, and J. C. Straley, "Defending VoIP Networks from Distributed DoS (DDoS) Attacks," in *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM)*, November/December 2004.

[95] J. Larson, T. Dawson, M. Evans, and J. C. Straley, "Defending VoIP Networks from DDoS Attacks," in *Proceedings of the 2^{nd} Workshop on Securing Voice over IP*, June 2005.

[96] A. Bremler-Barr, R. Halachmi-Bekel, and K. Kangasharju, "Unregister Attacks in SIP," in *Proceedings of the 2^{nd} IEEE Workshop on Secure Network Protocols*, pp. 32–37, November 2006.

[97] E. Y. Chen, "Detecting DoS Attacks on SIP Systems," in *Proceedings of the 1^{st} IEEE Workshop on VoIP Management and Security (VoIP MaSe)*, April 2006.

[98] H. Sengar, H. Wang, D. Wijesekera, and S. Jajodia, "Fast Detection of Denial-of-Service Attacks on IP Telephony," in *Proceedings of the 14^{th} IEEE International Workshop on Quality of Service (IWQoS)*, pp. 199–208, June 2006.

[99] H. Sengar, H. Wang, D. Wijesekera, and S. Jajodia, "Detecting VoIP Floods Using the Hellinger Distance," *IEEE Transactions on Parallel and Distributed Systems*, vol. 19, pp. 794–805, June 2008.

[100] J. Tang, Y. Cheng, and C. Zhou, "Sketch-based SIP Flooding Detection Using Hellinger Distance," in *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM)*, pp. 1–6, November/December 2009.

[101] G. Zhang, S. Ehlert, T. Magedanz, and D. Sisalem, "Denial of Service Attack and Prevention on SIP VoIP Infrastructures Using DNS Flooding," in *Proceedings of the 1^{st} International Conference on Principles, Systems and Applications of IP Telecommunications (IPTCOMM)*, pp. 57–66, July 2007.

[102] J. Fiedler, T. Kupka, S. Ehlert, T. Magedanz, and D. Sisalem, "VoIP Defender: Highly Scalable SIP-based Security Architecture," in *Proceedings of the 1^{st} International Conference on Principles, Systems and Applications of IP Telecommunications (IPTComm)*, pp. 11–17, July 2007.

[103] W. Conner and K. Nahrstedt, "Protecting SIP Proxy Servers from Ringing-based Denial-of-Service Attacks," in *Proceedings of the 10^{th} IEEE International Symposium on Multimedia (ISM)*, pp. 340–347, December 2008.

[104] M. Luo, T. Peng, and C. Leckie, "CPU-based DoS Attacks Against SIP Servers," in *Proceedings of the IEEE Network Operations and Management Symposium (NOMS)*, pp. 41–48, April 2008.

[105] C. Fuchs, N. Aschenbruck, F. Leder, and P. Martini, "Detecting VoIP-based DoS Attacks at the Public Safety Answering Point," in *Proceedings of the ACM Aymposium on Information, Computer and Communications Security (ASIACCS)*, pp. 148–155, March 2008.

[106] N. Aschenbruck, M. Frank, P. Martini, J. Tolle, R. Legat, and H.-D. Richmann, "Present and Future Challenges Concerning DoS-attacks against PSAPs in VoIP Networks," in *Proceedings of the 4^{th} IEEE International Workshop on Information Assurance (IWIA)*, pp. 103–108, April 2006.

[107] C. Hyun-Soo, R. Jea-Tek, R. Byeong-hee, K. Jeong-Wook, and J. Hyun-Cheol, "Detection of SIP De-Registration and Call-Disruption Attacks Using a Retransmission Mechanism and a Countermeasure Scheme," in *Proceedings of the IEEE International Conference on Signal Image Technology and Internet Based Systems (SITIS)*, pp. 650–656, November/December 2008.

[108] D. Geneiatakis and C. Lambrinoudakis, "A Lightweight Protection Mechanism against Signaling Attacks in a SIP-based VoIP Environment," *Telecommunication Systems*, vol. 36, pp. 153–159, December 2007.

[109] D. Geneiatakis and C. Lambrinoudakis, "A Cost-Effective Mechanism for Protecting SIP Based Internet Telephony Services Against Signaling Attacks," in *Proceedings of the IMS and Mobile Multimedia Workshop*, July 2008.

[110] G. Ormazabal, S. Nagpal, E. Yardeni, and H. Schulzrinne, "Secure SIP: A Scalable Prevention Mechanism for DoS Attacks on SIP Based VoIP Systems," in *Proceedings of the 2^{nd} International Conference on Principles, Systems and Applications of IP Telecommunications (IPTComm)*, pp. 107–132, July 2008.

[111] S. Ehlert, G. Zhang, D. Geneiatakis, G. Kambourakis, T. Dagiuklas, J. Markl, and D. Sisalem, "Two Layer Denial of Service Prevention on SIP VoIP Infrastructures," *Computer Communications*, vol. 31, pp. 2443–2456, June 2008.

[112] S. Ehlert, Y. Rebahi, and T. Magedanz, "Intrusion Detection System for Denial-of-Service Flooding Attacks in SIP Communication Networks," *International Journal of Security and Networks*, vol. 4, pp. 189–200, July 2009.

[113] S. Ehlert, C. Wang, T. Magedanz, and D. Sisalem, "Specification-based Denial-of-Service Detection for SIP Voice-over-IP Networks," in *Pro-*

ceedings of the 3^{rd} International Conference on Internet Monitoring and Protection, pp. 59–66, July 2008.

[114] D. Geneiatakis, N. Vrakas, and C. Lambrinoudakis, "Utilizing Bloom Filters for Detecting Flooding Attacks against SIP Based Services," *Computers and Security*, vol. 28, pp. 578–591, October 2009.

[115] D. Geneiatakis, N. Vrakas, and C. Lambrinoudakis, "Performance Evaluation of a Flooding Detection Mechanism for VoIP Networks," in *Proceedings of the 16^{th} International Workshop on Systems Signals and Image Processing*, pp. 1–5, June 2009.

[116] A. Awais, M. Farooq, and M. Y. Javed, "Attack Analysis & Bio-inspired Security Framework for IP Multimedia Subsystem," in *Proceedings of the GECCO Conference Companion on Genetic and Evolutionary Computation*, pp. 2093–2098, July 2008.

[117] Y. Rebahi, M. Sher, and T. Magedanz, "Detecting Flooding Attacks Against IP Multimedia Subsystem (IMS) Networks," in *Proceedings of the IEEE/ACS International Conference on Computer Systems and Applications*, pp. 848–851, March/April 2008.

[118] M. A. Akbar and M. Farooq, "Application of Evolutionary Algorithms in Detection of SIP based Flooding Attacks," in *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO)*, July 2009.

[119] M. A. Akbar and M. Farooq, "RTP-Miner: A Real-time Security Framework for RTP Fuzzing Attacks," in *Proceedings of the 20^{th} International Workshop on Network and Operating Systems Support for Digital Audio and Video (NOSSDAV)*, June 2010.

[120] M. Nassar, R. State, and O. Festor, "Monitoring SIP Traffic Using Support Vector Machines," in *Proceedings of the Symposium on Recent Advances in Intrusion Detection (RAID)*, pp. 311–330, September 2008.

[121] M. Z. Rafique, M. A. Akbar, and M. Farooq, "Evaluating DoS Attacks Against SIP-Based VoIP Systems," in *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM)*, November/December 2009.

[122] M. A. Akbar, Z. Tariq, and M. Farooq, "A Comparative Study of Anomaly Detection Algorithms for Detection of SIP Flooding in IMS," in *Proceedings of the International Conference on Internet Multimedia Services Architecture and Applications (IMSAA)*, December 2008.

[123] P. Battistello, "Work in Progress: Inter-Domain and DoS-Reistant Call Establishment Protocol (IDDR-CEP)," in *Proceedings of the 4^{th} Annual ACM Conference on Principles, Systems and Applications of IP Telecommunications (IPTCOMM)*, pp. 25–34, August 2010.

[124] P. Truong, D. Nieh, and M. Moh, "Specification-based Intrusion Detection for H.323-based Voice over IP," in *Proceedings of the IEEE International Symposium on Signal Processing and Information Technology*, December 2005.

[125] W. Mazurczyk and Z. Kotulski, "New Security and Control Protocol for VoIP Based on Steganography and Digital Watermarking," technical report, Institute of Fundamental Technological Research, Polish Academy of Sciences, June 2005.

[126] Z. Kotulski and W. Mazurczyk, "Covert Channel for Improving VoIP Security," in *Proceedings of the 13^{th} International Multi-Conference on Advanced Computer Systems (ACS)*, pp. 311–320, October 2006.

[127] W. Mazurczyk and Z. Kotulski, "New VoIP Traffic Security Scheme with Digital Watermarking," in *Proceedings of International Conference on Computer Safety, Reliability, and Security (SafeComp)*, pp. 170–181, September 2006.

[128] R. Zhang, X. Wang, X. Yang, and X. Jiang, "Billing Attacks on SIP-based VoIP Systems," in *Proceedings of the 1^{st} USENIX Workshop On Offensive Technologies (WOOT)*, pp. 1–8, August 2007.

[129] H. Abdelnur, T. Avanesov, M. Rusinowitch, and R. State, "Abusing SIP Authentication," in *Proceedings of the 4^{th} International Conference on Information Assurance and Security (ISIAS)*, pp. 237–242, September 2008.

[130] R. State, O. Festor, H. Abdelanur, V. Pascual, J. Kuthan, R. Coeffic, J. Janak, and J. Floroiu, "SIP digest authentication relay attack." draft-state-sip-relay-attack-00, March 2009.

[131] D. Geneiatakis, G. Kambourakis, and C. Lambrinoudakis, "A Mechanism for Ensuring the Validity and Accuracy of the Billing Services in IP Telephony," in *Proceedings of the 5^{th} International Conference on Trust, Privacy & Security in Digital Business (TrustBus)*, pp. 59–68, September 2008.

[132] R. Ackermann, M. Schumacher, U. Roedig, and R. Steinmetz, "Vulnerabilities and Security Limitations of current IP Telephony Systems," in *Proceedings of the Conference on Communications and Multimedia Security (CMS)*, pp. 53–66, May 2001.

[133] P. Hunter, "VOIP the Latest Security Concern: DoS Attack the Greatest Threat," *Network Security*, vol. 2002, pp. 5–7, November 2002.

[134] A. Batchvarov, "Security Issues and Solutions for Voice over IP Compared to Circuit Switched Networks," tech. rep., INFOTECH Seminar Advanced Communication Services (ACS), 2004.

[135] D. Bradbury, "The Security Challenges Inherent in VoIP ," *Computers & Security*, vol. 26, pp. 485–487, December 2007.

[136] J. Chau, "Security Issues Around the Deployment of VoIP and Multimedia Protocols in Wireless and Firewalled Environments," *Computer Fraud & Security*, vol. 2006, pp. 14–16, August 2006.

[137] D. Sicker and T. Lookabaugh, "VoIP Security: Not an Afterthought," *ACM Queue Magazine*, vol. 2, pp. 56–64, September 2004.

[138] S. Vuong and Y. Bai, "A Survey of VoIP Intrusions and Intrusion Detection Systems," in *Proceedings of the 6^{th} International Conference on Advanced Communication Technology (ICACT)*, pp. 317–322, February 2004.

[139] D. Geneiatakis, G. Kambourakis, C. Lambrinoudakis, T. Dagiuklas, and S. Gritzalis, "SIP Message Tampering: THE SQL code INJECTION attack," in *Proceedings of 13^{th} IEEE International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*, September 2005.

[140] G. S. Tucker, "Voice Over Internet Protocol (VoIP) and Security," white paper, SANS Institute, 2005.

[141] J. Posegga and J. Seedorf, "Voice Over IP: Unsafe at any Bandwidth?," in *Proceedings of the Eurescom Summit: Ubiquitous Services and Applications Exploiting the Potential*, Apri 2005.

[142] E. Edelson, "Voice over IP: Security Pitfalls," *Network Security*, vol. 2005, pp. 4–7, February 2005.

[143] J. Albers, B. Hahn, S. McGann, S. Park, and R. Zhu, "An Analysis of Security Threats and Tools in SIP-Based VoIP Systems," M.Sc. Capstone Paper, Univesity of Colorado, Boulder, 2005.

[144] S. McGann and D. Sicker, "An Analysis of Security Threats and Tools in SIP-Based VoIP Systems," in *Proceedings of the 2^{nd} VoIP Security Workshop*, June 2005.

[145] D. Geneiatakis, G. Kambourakis, T. Dagiuklas, C. Lambrinoudakis, and S. Gritzalis, "SIP Security Mechanisms: A state-of-the-art review," in *Proceedings of the 5^{th} International Network Conference (INC)*, July 2005.

[146] F. Cao and S. Malik, "Security Analysis and Solutions for Deploying IP Telephony in the Critical Infrastructure," in *Proceedings of the workshop of the 1^{st} International Conference on Security and Privacy for Emerging Areas in Communication Networks*, pp. 171–180, September 2005.

[147] F. Cao and S. Malik, "Vulnerability Analysis and Best Practices for Adopting IP Telephony in Critical Infrastructure Sectors," *IEEE Communications Magazine*, vol. 44, pp. 138–145, April 2006.

[148] B. Allain, "VoIP Security Challenges and Approaches," in *Proceedings of the 2^{nd} Workshop on Securing Voice over IP*, June 2005.

[149] A. Adelsbach, A. Alkassar, K.-H. Garbe, M. Luzaic, M. Manulis, E. Scherer, J. Schwenk, and E. Siemens, "Voice over IP: Sichere Umstellung der Sprachkommunikation auf IP-Technologie." Bundesanzeiger Verlag, 2005.

[150] D. R. Kuhn, T. J. Walsh, and S. Fries, "Security Considerations for Voice Over IP Systems." US National Institute of Standards and Technology (NIST) Special Publication SP 800-58, January 2005.

[151] T. J. Walsh and D. R. Kuhn, "Challenges in Securing Voice over IP," *IEEE Security & Privacy Magazine*, vol. 3, pp. 44–49, May/June 2005.

[152] Z. Anwar, W. Yurcik, R. E. Johnson, M. Hafiz, and R. H. Campbell, "Multiple Design Patterns for Voice over IP (VoIP) Security," in *Proceedings of the IEEE Workshop on Information Assurance (WIA), held in conjunction with the 25^{th} IEEE International Performance Computing and Communications Conference, (IPCCC)*, April 2006.

[153] D. Geneiatakis, T. Dagiuklas, G. Kambourakis, C. Lambrinoudakis, S. Gritzalis, K. S. Ehlert, and D. Sisalem, "Survey of Security Vulnerabilities in Session Initiation Protocol," *IEEE Communications Surveys & Tutorials*, vol. 8, pp. 68–81, 3^{rd} Quarter 2006.

[154] D. Geneiatakis, C. Lambrinoudakis, and G. Kambourakis, "An Ontology Based Policy for Deploying Secure SIP-based VoIP Services," *Computers and Security*, vol. 27, pp. 285–297, October 2008.

[155] G. Me and D. Verdone, "An Overview of Some Techniques to Exploit VoIP over WLAN," in *Proceedings of the International Conference on Digital Telecommunications (ICDT)*, pp. 67–73, August 2006.

[156] R. Singhai and A. Sahoo, "VoIP Security," technical report, Indian Institute of Technology, Mumbai — School of Information Technology, 2006.

[157] W. J. Rippon, "Threat Assessment of IP Based Voice Systems," in *Proceedings of the 1^{st} IEEE Workshop on VoIP Management and Security (VoIP MaSe)*, pp. 19–28, April 2006.

[158] P. C. K. Hung and M. V. Martin, "Security Issues in VoIP Applications," in *Proceedings of the Canadian Conference on Electrical and Computer Engineering (CCECE)*, pp. 2361–2364, May 2006.

[159] P. C. K. Hung and M. V. Martin, "Through the looking glass: Security issues in VoIP applications," in *Proceedings of the IADIS International Conference on Applied Computing*, February 2006.

[160] M. Zandi, M. V. Martin, and P. C. K. Hung, "Overview of Security Issues of VOIP," in *Proceedings of the IASTED European Conference on Internet and Multimedia Systems and Applications (IMSA)*, pp. 254–259, March 2007.

[161] J. Xin, "Security Issues and Countermeasure for VoIP ," white paper, SANS Institute, 2007.

[162] D. Persky, "VoIP Security Vulnerabilities," white paper, SANS Institute, 2007.

[163] V. M. Quinten, R. van de Meent, and A. Pras, "Analysis of Techniques for Protection Against Spam over Internet Telephony," in *Proceedings of the* 13^{th} *Open European Summer School and IFIP TC6.6 Workshop (EUNICE)*, pp. 70–77, July 2007.

[164] P. Hansen and A. Woodward, "Network Security—Is IP Telephony Helping The Cause?," in *Proceedings of the* 5^{th} *Australian Information Security Management Conference*, pp. 73–79, December 2007.

[165] P. James and A. Woodward, "Securing VoIP: A Framework to Mitigate or Manage Risks," in *Proceedings of the* 5^{th} *Australian Information Security Management Conference*, pp. 103–116, December 2007.

[166] 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*, vol. 37, pp. 1152–1162, November 2007.

[167] P. Thermos and A. Takanen, *Securing VoIP Networks*. Pearson Education, 2008.

[168] A. Kurmus and J.-F. Garet, "Studying and Experimenting with Threats Against Voice over IP Systems," Tech. Rep. Masters Thesis, EURECOM, 2009.

[169] D. Sisalem, J. Floroiu, J. Kuthan, U. Abend, and H. Schulzrinne, *SIP Security*. Wiley, 2009.

[170] V. K. Gurbani and V. Kolesnikov, "A Survey and Analysis of Media Keying Techniques in the Session Initiation Protocol (SIP)," *IEEE Communications Surveys and Tutorials (to appear)*, 2011.

[171] Y. Rebahi, S. Ehlert, M. Theoharidou, J. Mallios, S. Dritsas, G. F. Marias, L. Mitrou, T. Dagiuklas, M. Avgoustianakis, D. Gritzalis, B. Pannier, O. Capsada, and J. Markl, "SPIT Threat Analysis," deliverable wp2/d2.1, SPIDER COOP-32720, January 2007.

[172] G. F. Marias, S. Dritsas, M. Theoharidou, J. Mallios, L. Mitrou, D. Gritzalis, T. Dagiuklas, Y. Rebahi, S. Ehlert, B. Pannier, O. Capsada, and J. F. Juell, "SPIT Detection and Handling Strategies for VoIP Infrastructures," Tech. Rep. Deliverable WP2/D2.2, SPIDER COOP-32720, March 2007.

[173] S. Dritsas, J. Mallios, M. Theoharidou, G. F. Marias, and D. Gritzalis, "Threat Analysis of the Session Initiation Protocol Regarding Spam," in *Proceedings of the* 26^{th} *IEEE International Performance Computing and Communications Conference (IPCCC)*, pp. 426–433, April 2007.

[174] G. F. Marias, S. Dritsas, M. Theoharidou, J. Mallios, and D. Gritzalis, "SIP Vulnerabilities and Anti-SPIT Mechanisms Assessment," in *Proceedings of the* 16^{th} *International Conference on Computer Communications and Networks (ICCCN)*, pp. 597–604, August 2007.

[175] S. Dritsas, Y. Soupionis, M. Theoharidou, Y. Mallios, and D. Gritzalis, "SPIT Identification Criteria Implementation: Effectiveness and Lessons Learned," in *Proceedings of the* 23^{rd} *IFIP TC11 International Information Security Conference (SEC)*, pp. 381–395, September 2008.

[176] N. d'Heureuse, J. Seedorf, S. Niccolini, and T. Ewald, "Protecting SIP-Based Networks and Services from Unwanted Communications," in *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM)*, pp. 1–5, November/December 2008.

[177] D. Sisalem, J. Kuthan, and S. Ehlert, "Denial of Service Attacks Targeting a SIP VoIP Infrastructure: Attack Scenarios and Prevention Mechanisms," *IEEE Network*, vol. 20, pp. 26–31, September/October 2006.

[178] J. Seedorf, "Security challenges for peer-to-peer SIP," *IEEE Network*, vol. 20, pp. 38–45, September/October 2006.

[179] C. Wieser, M. Laakso, and H. Schulzrinne, "Security Testing of SIP Implementations," Tech. Rep. CUCS-024-03, Columbia University, Department of Computer Science, 2003.

[180] R. Kaksonen, M. Laakso, and A. Takanen, "Software Security Assessment through Specification Mutations and Fault Injection."

[181] "CVE-2003-1109." http://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2003-1109, 2003.

[182] T. Berson, "Skype Security Evaluation," October 2005.

[183] S. A. Baset and H. Schulzrinne, "An Analysis of the Skype Peer-to-Peer Telephony Protocol," in *Proceedings of IEEE INFOCOM*, April 2006.

[184] P. Biondi and F. Desclaux, "Silver Needle in the Skype," in *BlackHat Europe Conference*, March 2006. www.blackhat.com/presentations/bh-europe-06/bh-eu-06-biondi/bh-eu-06-biondi-up.pdf, **(reverse engineer)**.

[185] P. Thermos and G. Hadsall, "Vulnerabilities in SOHO VoIP Gateways," in *Proceedings of the* 2^{nd} *VoIP Security Workshop*, June 2005.

[186] H. Scholz, "Attacking VoIP Networks," in *Proceedings of the* 3^{rd} *Workshop on Securing Voice over IP*, June 2006.

[187] H. Abdelnur, R. State, and O. Festor, "Fuzzing for Vulnerabilities in the VoIP Space," in *Proceedings of the* 17^{th} *Annual Conference of the European Institute for Computer Anti-Virus Research (EICAR)*, May 2008.

[188] H. Abdelnur, R. State, and O. Festor, "KiF: A stateful SIP Fuzzer," in *Proceedings of the* 1^{st} *International Conference on Principles, Systems and Applications of IP Telecommunications*, pp. 47–56, July 2007.

[189] H. Abdelnur, V. Cridlig, R. State, O. Festor, and J. Bourdellon, "VoIP Security Assessment: Methods and Tools," in *Proceedings of the* 1^{st} *IEEE Workshop on VoIP Management and Security (VoIP MASe)*, pp. 29–34, April 2006.

[190] H. Abdelnur, R. State, I. Chrisment, and C. Popi, "Assessing the security of VoIP Services," in *Proceedings of the* 10^{th} *IFIP/IEEE Symposium on Integrated Management (IM)*, pp. 373–382, May 2007.

[191] P. Gupta and V. Shmatikov, "Security Analysis of Voice-over-IP Protocols," in *Proceedings of the* 20^{th} *IEEE Computer Security Foundations Symposium (CSFW)*, pp. 49–63, July 2007.

[192] J. Floroiu and D. Sisalem, "A Comparative Analysis of the Security Aspects of the Multimedia Key Exchange Protocols," in *Proceedings of the* 3^{rd} *International Conference on Principles, Systems and Applications of IP Telecommunications (IPTComm)*, pp. 2:1–2:10, July 2009.

[193] J. Reason and D. Messerschmitt, "The Impact of Confidentiality on Quality of Service in Heterogeneous Voice over IP Networks," in *Proceedings of the IEEE Conference on Management of Multimedia Networks and Services*, pp. 175–192, November 2001.

[194] A. D. Elbayoumy and S. Shepherd, "Stream or Block Cipher for Securing VoIP?," *International Journal of Network Security*, vol. 5, pp. 128–133, September 2007.

[195] A. D. Elbayoumy and S. Shepherd, "A Comprehensive Secure VoIP Solution," *International Journal of Network Security*, vol. 5, pp. 233–240, September 2007.

[196] S. Salsano, L. Veltri, and D. Papalilo, "SIP Security Issues: The SIP Authentication Procedure and its Processing Load," *IEEE Network*, vol. 16, pp. 38–44, November/December 2002.

[197] N. Modadugu and E. Rescorla, "The Design and Implementation of Datagram TLS," in *Proceedings of the ISOC Symposium on Network and Distributed Systems Security (NDSS)*, February 2004.

[198] R. Barbieri, D. Bruschi, and E. Rosti, "Voice over IPsec: Analysis and Solutions," in *Proceedings of the* 18^{th} *Annual Computer Security Applications Conference (ACSAC)*, pp. 261–270, December 2002.

[199] M. K. Ranganathan and L. Kilmartin, "Performance Analysis of Secure Session Initiation Protocol Based VoIP Networks," *Computer Communications*, vol. 26, pp. 552–565, April 2003.

[200] D. Harkins and D. Carrel, "The Internet Key Exchange (IKE)." RFC 2409 (Proposed Standard), Nov. 1998. Obsoleted by RFC 4306, updated by RFC 4109.

[201] C. Kaufman, "Internet Key Exchange (IKEv2) Protocol." RFC 4306 (Proposed Standard), Dec. 2005. Updated by RFC 5282.

[202] J. Bilien, "Key Agreement for Secure Voice over IP," Master of Science Thesis IMIT/LCN 2003-14, Royal Institute of Technology, Sweden, December 2003.

[203] J. Bilien, E. Eliasson, and J.-O. Vatn, "Call Establishment Delay for Secure VoIP," in *Proceedings of the Workshop on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt)*, March 2004.

[204] J. Bilien, E. Eliasson, J. Orrblad, and J.-O. Vatn, "Secure VoIP: Call Establishment and Media Protection," in *Proceedings of the* 2^{nd} *Workshop on Securing Voice over IP*, June 2005.

[205] H. Xiao and P. Zarrella, "Quality Effects of Wireless VoIP Using Security Solutions," in *Proceedings of the IEEE Military Communications Conference (MILCOM)*, vol. 3, pp. 1352–1357, October/November 2004.

[206] E. T. Lakay and J. I. Agbinya, "Security Issues in SIP Signaling in Wireless Networks and Services," in *Proceedings of the International Conference on Mobile Business*, pp. 639–642, July 2005.

[207] C. Eun-Chul, C. Hyoung-Kee, and C. Sung-Jae, "Evaluation of Security Protocols for the Session Initiation Protocol," in *Proceedings of the 16^{th} International Conference on Computer Communications and Networks (ICCCN)*, pp. 611–616, August 2007.

[208] C. Shen, E. Nahum, H. Schulzrinne, and C. P. Wright, "The Impact of TLS on SIP Server Performance," in *Proceedings of the 4^{th} Annual ACM Conference on Principles, Systems and Applications of IP Telecommunications (IPTCOMM)*, pp. 63–74, August 2010.

[209] Y. Rebahi, J. J. Pallares, G. Kovacs, N. T. Minh, S. Ehlert, and D. Sisalem, "Performance Analysis of Identity Management in the Session Initiation Protocol (SIP)," in *Proceedings of the IEEE/ACS International Conference on Computer Systems and Applications (AICCSA)*, pp. 711–717, March/April 2008.

[210] A. Buschel, "Authentication in VoIP," in *Proceedings of the 2^{nd} Workshop on Securing Voice over IP*, June 2005.

[211] C. C. Yang, R. C. Wang, and W. T. Liu, "Secure Authentication Scheme for Session Initiation Protocol," *Computers and Security*, vol. 24, pp. 381–386, August 2005.

[212] A. Durlanik and I. Sogukpinar, "SIP Authentication Scheme Using ECDH," *World Academy Science, Engineering and Technology (WASET)*, vol. 8, pp. 350–353, September 2005.

[213] E.-J. Yoon and K.-Y. Yoo, "A New Authentication Scheme for Session Initiation Protocol," in *Proceedings of the International Conference on Complex, Intelligent and Software Intensive Systems*, pp. 549–554, March 2009.

[214] J. L. Tsai, "Efficient Nonce-based Authentication Scheme for Session Initiation Protocol," *International Journal of Network Security (IJNS)*, vol. 9, pp. 12–16, July 2009.

[215] R. Srinivasan, V. Vaidehi, K. Harish, K. L. Narasimhan, S. L. Babu, and V. Srikanth, "Authentication of Signaling in VoIP Applications," in *Proceedings of the 11^{th} Asia-Pacific Conference on Communications (APCC)*, pp. 530–533, October 2005.

[216] A. Mohammadi-nodooshan, Y. Darmani, R. Jalili, and M. Nourani, "A Robust and Efficient SIP Authentication Scheme," in *Proceedings of the 13^{th} International CSI Computer Conference (CSICC)*, pp. 551–558, March 2008.

[217] H.-F. Huang and W.-C. Wei, "A New Efficient Authentication Scheme for Session Initiation Protocol," in *Proceedings of the Joint Conference on Information Sciences (JCIS), 9^{th} International Conference on Computer Science and Informatics*, October 2006.

[218] C.-C. Chang, Y.-F. Lu, A.-C. Pang, and T.-W. Kuo, "Design and Implementation of SIP Security," in *Proceedings of the International Conference On Information Networking (ICION)*, pp. 669–678, February 2005.

[219] C. C. Lee, "On Security of An Efficient Nonce-based Authentication Scheme for SIP," *International Journal of Network Security*, vol. 9, pp. 201–203, November 2009.

[220] F. Cao and C. Jennings, "Providing Response Identity and Authentication in IP Telephony," in *Proceedings of the 1^{st} International Conference on Availability, Reliability and Security (ARES)*, April 2006.

[221] K. Insu and K. Keecheon, "Secure Session Management Mechanism in VoIP Service," in *Proceedings of the Workshop on Ubiquitous Processing for Wireless Networks (UPWN), held in conjunction with the 5^{th} International Symposium on Parallel and Distributed Processing and Applications (ISPA)*, pp. 96–104, August 2007.

[222] H. Schmidt, C.-T. Dang, and F. J. Hauck, "Proxy-based Security for the Session Initiation Protocol(SIP)," in *Proceedings of the 2^{nd} International Conference on Systems and Networks Communications (ICSNC)*, pp. 42–47, August 2007.

[223] F. Wang and Y. Zhang, "A New Provably Secure Authentication and Key Agreement for SIP Using Certificateless Public-Key Cryptography," *Computer Communications*, vol. 31, pp. 2142–2149, June 2008.

[224] S. Al-Riyami and K. Paterson, "Certificateless Public Key Cryptography," in *Proceedings of AsiaCrypt*, pp. 452–473, November/December 2003.

[225] J. Ring, K.-K. R. Choo, E. Foo, and M. Looi, "A New authentication Mechanism and Key Agreement Protocol for SIP Using Identity-based Cryptography," in *Proceedings of AusCERT, R&D Stream*, pp. 61–72, May 2006.

[226] K. Singh and S. Vuong, "Blaze: A Mobile Agent Paradigm for VoIP Intrusion Detection Systems," in *Proceedings of the 1^{st} International Conference on E-Business and Telecommunication Networks (ICETE)*, August 2004.

[227] V. Casola, R. Chianese, A. Mazzeo, N. Mazzocca, and M. Rak, "A Policy-based Design Methodology and Performance Evaluation Framework for a Secure VoIP Infrastructure," in *Proceedings of the International Conference on E-business and TElecommunication Networks (ICETE)*, August 2004.

[228] V. Casola, M. Rak, A. Mazzeo, and N. Mazzoccca, "Security Design and Evaluation in a VoIP Secure Infrastracture: A Policy Based Approach," in *Proceedings of the International Conference on Information Technology: Coding and Computing*, pp. 727–732, April 2005.

[229] Y. Wu, S. Bagchi, S. Garg, and N. Singh, "SCIDIVE: A Stateful and Cross Protocol Intrusion Detection Architecture for Voice-over-IP Environments," in *Proceedings of the Conference on Dependable Systems and Networks (DSN*, pp. 433–442, June/July 2004.

[230] V. Apte, Y.-S. Wu, S. Bagchi, S. Garg, and N. Singh, "SPACEDIVE: A Distributed Intrusion Detection System for Voice-over-IP Environments (Fast Abstract)," in *Proceedings of the International Conference on Dependable Systems and Networks (DSN)*, pp. 25–28, June 2006.

[231] Y.-S. Wu, V. Apte, S. Bagchi, S. Garg, and N. Singh, "Intrusion Detection in Voice over IP Environments," *International Journal of Information Security*, vol. 8, pp. 153–172, June 2009.

[232] M. V. Martin and P. C. K. Hung, "Towards a Security Policy for VoIP Applications," in *Proceedings of the Canadian Conference on Electrical and Computer Engineering (CCECE)*, pp. 65–68, May 2005.

[233] D. Sisalem, S. Ehlert, D. Geneiatakis, G. Kambourakis, T. Dagiuklas, J. Markl, M. Rokos, O. Botron, J. Rodriguez, and J. Liu, "Towards a Secure and Reliable VoIP Infrastructure," Tech. Rep. Deliverable D2.1, SNOCER COOP-005892, May 2005.

[234] T. Dagiuklas, D. Geneiatakis, G. Kambourakis, D. Sisalem, S. Ehlert, J. Fiedler, J. Markl, M. Rokis, O. Botron, J. Rodriguez, and J. Liu, "General Reliability and Security Framework for VoIP Infrastructures," Tech. Rep. Deliverable D2.2, SNOCER COOP-005892, September 2005.

[235] D. Geneiatakis and C. Lambrinoudakis, "An Ontology Description for SIP Security Flaws," *Computer Communications*, vol. 30, pp. 1367–1374, April 2007.

[236] S. Niccolini, R. G. Garroppo, S. Giordano, G. Risi, and S. Ventura, "SIP Intrusion Detection and Prevention: Recommendations and Prototype Implementation," in *Proceedings of the 1^{st} IEEE Workshop on VoIP Management and Security (VoIP MaSe)*, pp. 47–52, April 2006.

[237] W. Marshall, A. F. Faryar, K. Kealy, G. de los Reyes, I. Rosencrantz, R. Rosencrantz, and C. Spielman, "Carrier VoIP Security Architecture," in *Proceedings of the 12^{th} International Telecommunications Network Strategy and Planning Symposium*, pp. 1–6, November 2006.

[238] M. Sher and T. Magedanz, "Protecting IP Multimedia Subsystem (IMS) Service Delivery Platform from Time Independent Attacks," in *Proceedings of the 3^{rd} International Symposium on Information Assurance and Security (IAS)*, pp. 171–176, August 2007.

[239] Y. Ding and G. Su, "Intrusion Detection System for Signal-based SIP Attacks Through Timed HCPN," in *Proceedings of the 2^{nd} International Conference on Availability, Reliability and Security (ARES)*, pp. 190–197, April 2007.

[240] M. Nassar, R. State, and O. Festor, "VoIP Honeypot Architecture," in *Proceedings of the 10^{th} IFIP/IEEE International Symposium on Integrated Network Management*, pp. 109–118, May 2007.

[241] M. Nassar, S. Niccolini, R. State, and T. Ewald, "Holistic VoIP Intrusion Detection and Prevention System," in *Proceedings of the 1^{st} International Conference on Principles, Systems and Applications of IP Telecommunications (IPTCOMM)*, pp. 1–9, July 2007.

[242] B. I. A. Barry and H. A. Anthony, "On the Performance of a Hybrid Intrusion Detection Architecture for Voice over IP Systems," in *Proceedings of the 4^{th} International Conference on Security and Privacy in Communication Networks (SecureComm)*, pp. 1–10, September 2008.

[243] K. Rieck, S. Wahl, P. Laskov, P. Domschitz, and K.-R. Müller, "A Self-learning System for Detection of Anomalous SIP Messages," in *Proceedings of the 2^{nd} Internation Conference on Principles, Systems and Applications of IP Telecommunications. Services and Security for Next Generation Networks: Second International Conference, (IPTComm)*, pp. 90–106, July 2008.

[244] G. F. Cretu, A. Stavrou, M. E. Locasto, S. J. Stolfo, and A. D. Keromytis, "Casting out Demons: Sanitizing Training Data for Anomaly Sensors," in *Proceedings of the IEEE Security and Privacy Symposium*, pp. 81–95, May 2008.

[245] R. Dantu, S. Fahmy, H. Schulzrinne, and J. Cangussu, "Issues and Challenges in Securing VoIP," *Computers & Security (to appear)*, 2010.

[246] B. Reynolds and D. Ghosal, "STEM: Secure Telephony Enabled Middlebox," *IEEE Communications Magazine*, vol. 40, pp. 52–58, October 2002.

[247] U. Roedig, R. Ackermann, and R. Steinmetz, "Evaluating and Improving Firewalls for IP-Telephony Environments," in *Proceedings of the* 1^{st} *IP Telephony Workshop*, April 2000.
[248] J. Kuthan, "Internet Telephony Traversal Across Decomposed Firewalls and NATs," in *Proceedings of the* 2^{nd} *IP Telephony Workshop*, April 2001.
[249] P. Sijben, W. van Willigenburg, M. de Boer, and S. van der Gaast, "Middleboxes: Controllable Media Firewalls," *Bell Labs Technical Journal*, vol. 7, pp. 141–157, August 2002.
[250] D. T. Stott, "SAFENeT: Server-based Architecture For Enterprise NAT and Firewall Traversal," in *Proceedings of the* 2^{nd} *VoIP Security Workshop*, June 2005.
[251] T. Bessis, A. Rana, and V. K. Gurbani, "Session Initiation Protocol (SIP) Firewall for Internet Multimedia Subsystem (IMS) Core," *Bell Labs Technical Journal*, Fall 2010.
[252] V. K. Gurbani, D. Willis, and F. Audet, "Cryptographically Transparent Session Initiation Protocol (SIP) Proxies," in *Proceedings of the IEEE International Conference on Communications (ICC)*, pp. 1185–1190, June 2007.
[253] H. Sengar, D. Wijesekera, S. Jajodia, and R. Dantu, "Securing VoIP from Signaling Network Vulnerabilities," in *Proceedings of the* 2^{nd} *Workshop on Securing Voice over IP*, June 2005.
[254] H. Sengar, R. Dantu, and D. Wijesekera, "Securing VoIP and PSTN from Integrated Signaling Network Vulnerabilities," in *Proceedings of the* 1^{st} *IEEE Workshop on VoIP Management and Security*, pp. 1–7, April 2005.
[255] H. Sengar, D. Wijesekera, H. Wang, and S. Jajodia, "VoIP Intrusion Detection Through Interacting Protocol State Machines," in *Proceedings of the International Conference on Dependable Systems and Networks (DSN)*, pp. 393–402, June 2006.
[256] S. Ehlert, G. Zhang, and T. 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)*, pp. 1–6, September 2008.
[257] M. Mandjes, I. Saniee, and A. L. Stolyar, "Load characterization and Anomaly Detection for Voice over IP Traffic (Extended Abstract)," in *Proceedings of the ACM SIGMETRICS Conference*, June 2001.
[258] M. Mandjes, I. Saniee, and A. A. L. Stolyar, "Load Characterization and Anomaly Detection for Voice over IP Traffic," *IEEE Transactions on Neural Networks*, vol. 16, pp. 1019–1026, September 2005.
[259] D. Geneiatakis, G. Kambourakis, T. Dagiuklas, C. Lambrinoudakis, and S. Gritzalis, "A Framework for Detecting Malformed Messages in SIP Networks," in *Proceedings of* 14^{th} *IEEE Workshop on Local and Metropolitan Area Networks (LANMAN)*, September 2005.
[260] D. Geneiatakis, G. Kambourakis, T. Dagiuklas, C. Lambrinoudakis, and S. Gritzalis, "A Framework for Detecting Malformed Messages in SIP Networks," *Computer Networks: The International Journal of Computer and Telecommunications Networking*, vol. 51, pp. 2580–2593, July 2007.
[261] D. Geneiatakis, T. Dagiuklas, C. Lambrinoudakis, G. Kambourakis, and S. Gritzalis, "Novel Protecting Mechanism for SIP-based Infrastructure against Malformed Message Attacks: Performance Evaluation Study," in *Proceedings of the* 5^{th} *International Conference on Communication Systems, Networks and Digital Signal Processing (CSNDSP)*, pp. 261–266, July 2006.
[262] D. Geneiatakis, G. Kambourakis, C. Lambrinoudakis, A. Dagiouklas, and S. Gritzalis, "A Framework for Protecting SIP-based Infrastructure against Malformed Message Attacks," *Computer Networks*, vol. 51, pp. 2580–2593, July 2007.
[263] D. Geneiatakis and A. D. Keromytis, "Towards a Forensic Analysis for Multimedia Communication Services," in *Proceedings of the* 7^{th} *International Symposium on Frontiers in Networking with Applications (FINA)*, March 2011.
[264] N. Hantehzadeh, A. Mehta, V. K. Gurbani, L. Gupta, T. K. Ho, and G. Wilathgamuwa, "Statistical analysis of self-similar Session Initiation Protocol (SIP) messages for anomaly detection," in *Proceedings of the* 4^{th} *IFIP/IEEE International Conference on New Technologies, Mobility, and Security (NTMS)*, February 2011.
[265] A. Mehta, N. Hantehzadeh, V. K. Gurbani, T. K. Ho, J. Koshiko, and R. Vishwanathan, "On the inefficacy of Euclidean classifiers for detecting self-similar Session Initiation Protocol (SIP) messages," in *Proceedings of the* 12^{th} *IFIP/IEEE International Symposium on Integrated Network Management (IM)*, May 2011.
[266] D. Seo, H. Lee, and E. Nuwere, "Detecting More SIP Attacks on VoIP Services by Combining Rule Matching and State Transition Models," in *Proceedings of the* 23^{rd} *IFIP TC11 International Information Security Conference (SEC)*, pp. 397–411, September 2008.
[267] F. Cao, B. Ha, R. Padmanabhan, A. Yuan, and K. Tran, "Call Filtering and Tracking in IP Telephony," in *Proceedings of the* 7^{th} *IASTED International Conference on Internet and Multimedia Systems and Applications (IMSA)*, January 2003.
[268] M. Steinebach, F. Siebenhaar, C. Neubauer, R. Ackermann, U. Roedig, and J. Dittmann, "Intrusion Detection Systems for IP Telephony Networks," in *Proceedings of the Real Time Intrusion Detection Symposium*, pp. 1–9, May 2002.

Angelos D. Keromytis is an Associate Professor of Computer Science and Director of the Network Security Lab at Columbia University in New York. His research interests focus on network and systems security and cryptography. He received his B.Sc. from the University of Crete, and his M.Sc. and Ph.D. from the University of Pennsylvania.