Systems and Methods for SPIT Detection in VoIP: Survey and Future Directions

Muhammad Ajmal Azad
School of Computing
University of Newcastle
UK
Email: *majmalazad@gmail.com*

Ricardo Morla
Faculty of Engineering
University of Porto
Portugal
Email: *ricardo.morla@fe.up.pt*

Khaled Salah
ECE Department
Khalifa University of Science and Technology
UAE
Email: *khaled.salah@kustar.ac.ae*

Abstract

In recent years, VoIP (Voice over IP) telephony has shown a tremendous increase in the number of subscribers due to today's affordable telephony rates and flexible use of Internet technology for voice communication. At the same time, a proportional increase was exhibited in VoIP spamming and SPam over Internet Telephony (SPIT), which are forms of abuse and frauds that can have severe consequences and financial losses for both the service providers and subscribers. This paper surveys, reviews, and discusses the state-of-the-art detection and mitigation techniques and systems for VoIP spamming and SPIT. The paper highlights reasons and motivation behind such abuse and fraud, and it discusses the primary challenges in devising an effective and efficient anti-SPIT detection solutions. Moreover, the paper outlines shortcomings and limitations of existing solutions, and it identifies future research directions to aid in further improving and enhancing effectively and efficiently the detection and mitigation of SPIT and spamming.

KEYWORDS: Voice over IP (VoIP), VoIP Spam, SPam over Internet Telephony (SPIT), SPIT Detection

1. Introduction

Voice over IP (VoIP) is an Internet Protocol (IP)-based voice communication system that is increasingly used by a large number of people along with the traditional public switched telephone network (PSTN) for business and personal communications. In recent years, VoIP has seen an enormous growth in the number of subscribers due to its attractive affordability calling rates globally. Moreover, VoIP provides affordable value services and flexibility of using IP networks for the voice communication. According to Cisco [1], VoIP market is expected to reach more than 1200 million subscribers worldwide by 2018 with expected revenue of more than $77 billion [2]. Figure 1 depicts the growth of residential VoIP subscribers from 2011 to 2016. The number of business subscribers is also increasing at a rate of 7.58%, and is expected to reach 244 million business subscribers by 2018 [1]. The affordable calling rates of VoIP, its easy integration with the IP networks, and value added services have also created a lucrative opportunity for spammers and telemarketers to initiate unwanted, bulk un-solicited calls via VoIP. In VoIP terminology, these calls are referred as SPam over Internet Telephony (SPIT)

and mainly used for advertising products, harassing subscribers, convincing subscribers to dial premium numbers, or making Vishing (voice equivalent of web Phishing) attacks to get private information of call recipients etc. Spammers can also make unwanted calls to steal user's information [3], make calls to check unsecure gateways within the service provider for the termination of bulk un-billed calls [4], and cause disruption in the network services through flooding and denial of service attacks [5, 6, 80].

Unwanted phone calls and instant text messaging can be initiated at any hour of the day. These unwanted calls and instant messages require immediate response from the recipient—and thereby disturbing and annoying call recipients while at work, home, or in the middle of the night. Recent statistics on telephony spam have revealed that answering a spam call would result in an estimated loss of 20 million man hours for a small business enterprise in the United States with the estimated loss of about $475 million annually [7]. Furthermore, USFTC (US Federal Trade Communication) has estimated that the annual loss attributed to both scamming and spamming activities had reached $8.6 billions in the U.S., with the vast majority of these calls initiated from regular telephones.

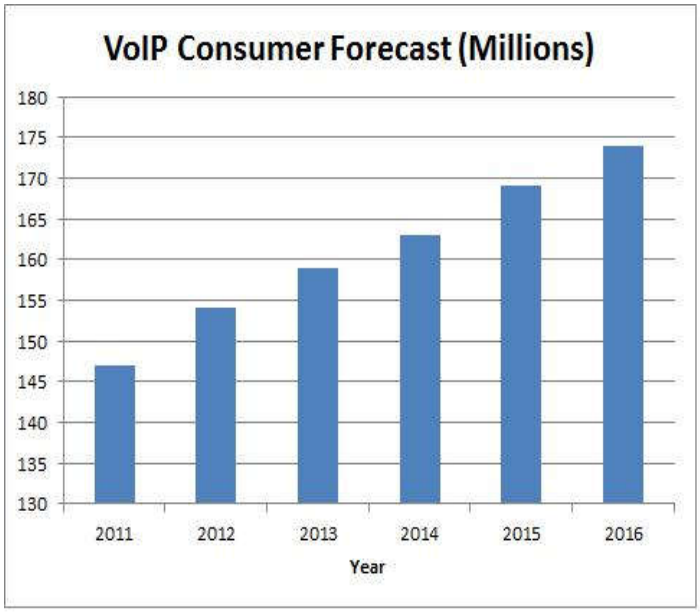

Fig. 1 VoIP Subscribers 2011-2016 [2]

Every year, service providers, regulators, and law enforcement agencies receive thousands of complaints from consumers for unsolicited, unauthorized, and fraudulent callers. In 2012, the USFTC has received four times more complaints against unwanted calls than number of complaints received in 2010 [8]. Additionally, the number of unidentified spam callers has risen to 162% from January 2013 to January 2014 [9].

This survey article has three primary objectives. First, it aims to present the taxonomy for SPIT detection systems; secondly, it clearly explains the idea used for spam detection in the telephone and VoIP network; thirdly, it provides possible future directions for SPIT detection systems. To date, and to the best of our knowledge, the literature lacks comprehensive and concise spam detection survey that discusses the state-of-the-art detection and mitigation techniques and systems for VoIP spamming and SPIT. Also the literature lacks sound review and discussion on the primary challenges and shortcoming in devising an effective and efficient anti-SPIT detection solutions.

Most of the survey articles that have been published in the literature focused on SPIT as threat to the VoIP network. IETF RFC 3039 lists different solutions that can be used to secure the VoIP network from the spamming attacks. In [12], authors survey the techniques designed for detecting and prevented unwanted calls in the telecommunication network. They also presented the evaluation criteria for assessing these approaches. In [91] and [6], the authors surveyed and analyzed published work related to securing the VoIP network from various security threats. The two survey papers summarized many previous related works and discussed the VoIP spamming as one of the categories in the social engineering attack. In [92], the authors presented a comprehensive survey that is also focused on discussing approaches designed for securing VoIP network from different security threats and SPIT attacks. In [93], the authors reviewed and discussed the solution that has been particular designed for SPIT detection in the VoIP network. In [94], the authors presented a survey discussing the attacks and solutions in VoIP, with VoIP Spam and Phishing being one of the attacks. Furthermore, the authors also proposed a solution to secure the VoIP network from number of different threats. In [95], the authors surveyed existing anti-SPIT system that categorized them in to three category Prevent, Detect, and Handle. In [96], the authors presented a survey reviewing various existing methods for preventing spam in IP telephony. The paper also presented a discussion on the implementation costs incurred for deploying these solutions. In [87], the authors presented a short survey of SPIT detection system, and proposed an approach that integrates and combines different features from call detailed records and SIP signaling messages. However, the survey does not provide enough details about each of the surveyed technique. In [88], the authors discussed number of approaches that have been designed for spam detection in the telephony network. However, they did not analyzed the systems in terms of their drawbacks and deployment challenges.

The main contributions of the paper can be summarized as follows:

- We present sufficient background and recent issues related to SPIT and spamming in a VoIP and telecommunication network. We describe reasons and motivations behind VoIP spamming, and discuss primary challenges in devising the effective SPIT detection techniques.

- We review, study, and provide taxonomy for the today's SPIT and VoIP spamming detection and filtering techniques that are designed for detecting spammers in a VoIP network. The major advantages as well as limitations of each of the detection schemes are also highlighted and discussed.
- We outline and discuss the latest work related to detecting VoIP and Email spamming in a collaborative manner across multiple service provides
- We provide comprehensive discussion, as well as a layout for future research directions for researchers to help them further enhancing the area of efficient and effective detection and filtering of unwanted calls and SPIT in a VoIP and telecommunication networks.

The remainder of this paper is organized as follows. Section 3 presents necessary background about spam over VoIP networks, and the scale of the problem in terms of implications, damage, and financial cost. Also, the section shed lights on the motivation of VoIP spammers, and the challenges in detecting SPIT spammers. Section 3 classifies and discusses anti-SPIT detection approaches and systems. The section presents a comprehensive taxonomy for identity-based and content-based detection schemes that are reported to date in the literature. Section 4 discusses a more advanced spam detection approaches that are collaborative in nature, and can be found for both VoIP and email networks. Section 5 provides concluding remarks, analysis and discussion. Section 6 outlines future research directions for enhancing and improving anti-SPIT detection and mitigation techniques.

2. Spam Over Internet Telephony

Voice spam or SPIT (Spam over Internet Telephony) can be defined as the unwanted, unsolicited, pre-recorded advertisement phone calls made by the spam sender to a large number of recipients that has no prior social relationship with the caller. VoIP spammers are similar to email spammers as both have the same intent of delivering unsolicited information to the recipients. This information typically contains advertisements of legal or illegal products. Some of these spammers commit fraud by attempting to access user private information by installing viruses that slow down system performance or steal user credential and sensitive information.

The spam calls and messages can also be sent to and from the mobile telephony system and the traditional PSTN (Public Switched Telephone Network) telephone network. The following are the additional forms of spam introduced due to advance forms of telephony:

Instant Message Spam: Bulk unsolicited instant messages (similar to email spam messages) but sent instantly and simultaneously to a massive number of users of messaging system like Skype [10], WhatsApp, Viber, etc.

Presence Spam: Presence spam is a bulk unsolicited set of presence requests messages sent to the subscribers with the aim of obtaining the subscriber's buddy or white list for sending IM.

Virus Spam: Sending viruses inside bulk SMS or IM messages that affect the functionality and operations of the mobile phone OS (operating system) and exploiting loop holes and vulnerabilities by spammers who would be able to remotely connect to spammers at a later stage.

2.1. Consequences of Spam over Internet Telephony

SPIT is one of the interactive forms of today's Internet and network abuses, whereby call recipient is required to respond immediately to the incoming call. Unlike email spammer, VoIP spammers are not only annoying and disturbing to callers, but also can be very costly causing significant loss because of answering a spam call while roaming and use of other value added services when answering the spam call. In telephony, spammers can be a threat to the subscribers of technology for any of the following reasons [11]:

i) Callee Account Credit. Telephony subscribers pay extra amount for the value added services for call forwarding, roaming, automatic call back, etc. The receiving of unsolicited calls and messages while roaming would charge callees for nothing. Similarly, automatic call back service might result in a call back to some premium numbers unintentionally by the users.

ii) Missing Important Calls. Legitimate users typically do not want to miss important calls and forward them to the voice mailbox during their periods of unavailability. However, a typical voice mailbox has very limited storage capacity which can be considerably exhausted by the spam call in the form of recorded message. This would result in resource unavailability for the calls from legitimate callers being forwarded to the voice mailbox. It would also be irritating and a waste of time for the callee to go through each spam message recorded in a voice mailbox. More importantly, the callee would also likely to miss some important recorded message if he dismissed them as spam on fly or if the mailbox was already full with the spam messages.

iii) Vishing. Vishing is equivalent to Phishing in the Web. In a Vishing attack, a spammer attempts to steal private and personally identifiable information by impersonating as a legitimate entity over the telephone network and later use this information for the financial frauds. Vishing is a type of social engineering fraud and is one of the emerging fraud across the world. In the years from 2014 to 2016 it brings the financial loss of around $1 billion [81]. In email network, Phishing might have a small impact on the users because users in the email network have some time to consult their friends and family before responding to the messages. Whereas in the telephony network, callees need to decide immediately whether to receive or

disconnect the call thus is prone to disclosing their personal information to the malicious and fraud callers that might bring some financial loss to them in a short time period [81].

iv) Financial Loss to Service Provider. Spam traffic consumes bandwidth and network resources, making resources unavailable to the legitimate users. Additionally, spammers or fraudsters try to identify some open and non-secure service providers for un-billed call termination. This can be very detrimental and damaging financially to service provider if a spammer remains undetected for long time periods. In the perspective of service provider's Quality of Service, users get annoyed if they do not receive sufficient network resources at the time of their calls. They also become annoyed if they receive large number unwanted advertisement calls without their prior agreement and consent. These scenarios would increase distrust of users to their service provider and can be damaging to the reputation of the service provider.

2.2. Motivation behind VoIP Spammers

The major motivation for all spammers is to gain financial benefits by practicing frauds on the recipients of calls and messages. Spammers undertake spamming attempts for advertising legal and illegal products, disseminating religious and political campaigns, convincing callees to buy product, call back to the premium numbers or disclose their private information for prize etc. The main goal of the spammers is to deliver their message to the public for greater financial benefits with small investment. Figure 2 depicts the spamming model of the spam caller in a VoIP network. A SPIT caller can be categorized into two groups: 1) auto dialer – where an automated machine or a computer generates a large number of spam calls to a large number of callees, and 2) a human SPIT caller – where the spammers hire inexpensive labor for making unsolicited calls to a large number of callees. The FTC received over 5.5 million complaints about unwanted calls, over 3.4 million of which involved a robocall and the human callers generate remaining 2.1 million calls [99].

In terms of financial benefits, spammers gain benefits in three ways. Firstly, they convince the callee to call them back on the premium numbers. For this purpose, spammers scam people by using social engineering attacks and exploit the needs of people in the particular societies such as offering expensive gifts and attractive tourism packages to historical and religious places. Secondly, they convince the callee to disclose his or her private information by impersonating as a legitimate entity. Thirdly, they make the callee listen to the complete advertisement and convenience them to buy products.

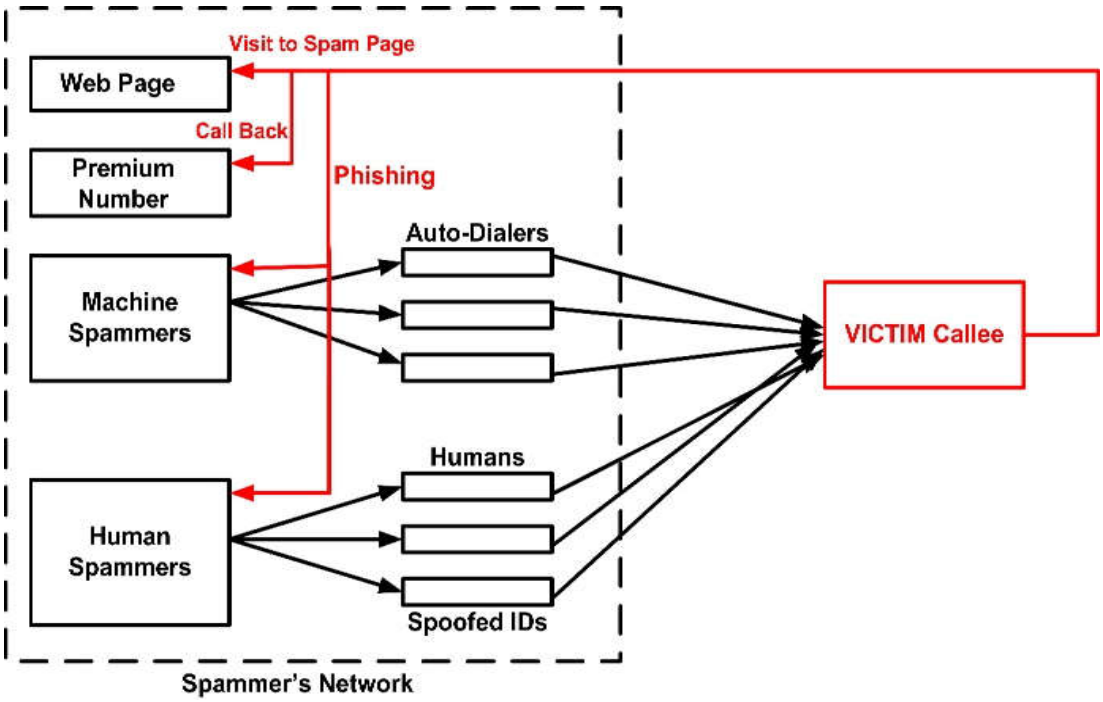

Fig. 2 Spammer Network Model

Spammers can also make spam calls or send spam messages for scanning end-user handset for the system vulnerabilities. The ability of making anonymous calls or missed calls to the specific callees from VoIP network also encourages spammers to use VoIP medium for threating or annoying callees with the denial of service attack to a specific callees. Beside financial benefits, spammers can also use telephony to spread a real time interactive religious hate or political messages to a large number of users. From the perspective of service provider or organization, spammers can make spam calls to learn the system vulnerabilities and identify some open gateways for a free of cost call terminations.

2.3. SPIT Differences from Email Spam

SPIT exhibits some similarity with the email spam. Both email spammers and SPIT callers use the Internet as a medium for sending their core message to callees but SPIT causes more serious discomfort to the victims because of the real time response for the call. In telephony, the callee has to decide immediately whether to accept or ignore the call. A callee has already been affected (waste of time) after realizing that the received call is a spam call. Besides similarity in motivations and IP medium, SPIT exhibits some differences from the email spams. E-mail spammer utilizes text messages, images or attachments for conveying their message to the victims, whereas SPIT caller uses digitized speech streams over the Internet for conveying his messages. In terms of deciding about sender and contents inside the message, the email service provider can hold e-mails for some time before finally delivering them to the callee's inbox, which is not noticeable to recipient. The VoIP or Voice service provider cannot hold speech streams and signaling messages without incurring noticeable delay

in a signaling and flow of speech streams between users. From the perspective of content processing, online processing of speech content is more challenging and resource intensive than offline processing of text messages and images.

From the user's perspective, a single e-mail spam can remain in the inbox unattended for as much time as the user wishes, but in the case of SPIT or voice call user has to respond back interactively, thereby making it more annoying and disturbing. From the perspective of user's resources, a single spam email typically consumes small number of bytes, but a voice message in a voice mailbox requires much greater space thus making voice mailbox space unavailable for the legitimate calls.

In terms of human efforts, the deletion of a SPIT call is more annoying and intrusive than the deletion of spam emails. In email network, firstly, the service provider assists end-user in classifying senders, and secondly end-user decides about email on a first look by reading the subject. On the other hand, in telephony, a user is required to listen the recorded call before thrashing it away as spam. The detection and deletion of spam speech content from the voice mailbox is more time consuming as it requires at least 6 steps to completely remove the speech content from the voice mailbox [12]. Furthermore, in telephony, user might also delete some important calls if he is making decision on a fly. From the perspective of protocol architecture, an E-mail message is composed of two parts: a header and a body. The email header part can also provide some information about sender's nature. Telephony calls also consist of two fundamental parts: (1) a call setup phase, and (2) a speech streaming phase; but the messages exchange in call setup phase though are available in a plain text but are not providing any information about the sender's nature and the speech stream is only available after the call setup.

2.4. Why is it hard to Detect SPIT Caller?

The affordable calling rates and use of telephony services over the Internet has convinced many subscribers and organizations to adopt VoIP as media for the business and non-business communications. Ironically, these features of VoIP have also attracted spammers to make use of this media for unsolicited activities on a large number of victims interactively. To make VoIP usable, to improve productive and trustworthiness of services, it is critical for the service provider to identifies and blocks spammers proactively and in a timely way. Unlike other forms of spams (for example email spam, blog spam, and web spam) VoIP spam or SIPT is much more difficult to detect. This is because spam can be distributed with the use of speech streams, which becomes available only after the call setup phase. Moreover, the service provider typically does not allocate sophisticated network resources for the processing of large number of speech samples in real-time. Though a human user can distinguish spam speech from non-speech, but this is always late as spammer has already annoyed callees with the unwanted content. The design of an effective SPIT detection system for a VoIP service provider is a challenging task. This can be attributed to the following reasons:

i) Unavailability of Speech Contents prior to Call-Setup. A typical VoIP call consist of two phases: 1) a call setup phase where a call request messages are exchange between caller, callee and the service provider, and 2) a real-time media exchange phase where a speech stream is exchange after a successful completion of call setup process. The messages that get exchanged between the caller, callee and proxy server during the call setup phase are in plain text and present call handling properties of involved parties. Though the signaling messages contain some information about caller and the callee, but they cannot provide enough information to be used for identifying SPIT caller. Additionally, it is very easy for the spammer to change his signaling messages and make them similar to the signaling messages of the legitimate users, thus leaving signaling content impractical for the spammer to detect. The speech contents could provide information about whether caller is spammer or not but speech content is only available after the call setup hence very late as spammer has already annoyed the callee with spam content. Beside this limitation, speech content- based spam detection systems have other limitations. Firstly, processing real-time speech content requires sophisticated and costly system resources for a real-time processing of speech content. Secondly, speech processing on active stream flows would add unnecessary delays to the conversations. Thirdly, processing and analyzing speech content is prohibited and restricted by law in many countries across the world because of privacy concerns. Further, content-based anti-SPIT system can easily evaded by the spammers by slightly manipulating the content and adding a random noise in the speech streams.

ii) Intrusiveness. Callers and the callees become annoyed and irritated if they are repeatedly get asked for solving a certain challenge or get asked to provide feedback about the callers at the end of every call transaction. Existing SPIT detection approaches involve callee in two ways: 1) after the call termination, and 2) before the call establishment. In the first, callees are asked to provide positive and negative feedback [13 – 15] about the caller's transaction as soon as call ends. This feedback is subsequently used for computing global reputation score of the caller within the network. In the second, the service provider provides credentials of the caller to the callee who can then decides whether to accept or reject the call [16]. Both of these approaches are intrusive and require changes to the VoIP infrastructure. For example, an email like spam button in the VoIP or telephony handset needs is to be added. In terms of caller, the existing approaches probe the caller to prove authentication in two ways. First, the caller is asked to solve a certain challenge in the form of CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apart) or Turing test, and second, the caller is asked to authenticate himself through the exchange of valid private and public keys. The first approach is intrusive while the second approach though is non-intrusive but require private-public infrastructure for the authentication and authorization. Moreover, challenge-response based approaches require computational overheads and system resources for handling a large number of concurrent callers, and would significantly add call setup delay.

iii) Calling Behavior and Automatic Classification. Telephony has become a popular medium for the communication. Through telephony, users communicate and talk with their family and friends, and conduct businesses. A clear understanding of the communication behavior of spammers and non-spammers can help for the effective design of a SPIT detection system for the VoIP network. The calling behavior or profile of users may characterize through the features like call-rates, the number of unique callees of the user, the time duration of user's interactions with their peer callees, and the number and duration of incoming calls to the user. The use of single feature for modeling behavior of user would not provide an effective prevention and protection against spammers, as spammers can easily evade them by manipulating the single feature. For example, SPIT detection approaches proposed in [10, 16] use average call duration for computing global reputation of the user for classifying user as a spammer and non-spammer. However, these systems can be evaded by spammers through the creation of a Sybil network among their own identities. Moreover, the use of an average call duration feature would assign high reputation scores to the caller having small duration calls with majority of his callees despite having large number of callees and non-repetitive calls. The challenge in an effective behavioral-based spam detection is three-fold: firstly, investigating which set of features are difficult to be manipulated by spammers, secondly, investigating which features could be used collectively for the computation of global reputation, and thirdly, having the system that automatically decides about nature of the user without any user intervention.

iv) Privacy-Aware Collaboration. Existing standalone SPIT detection systems consider locally recorded call logs for modeling the behavior of the users within the service provider network. The standalone detection systems lack global view of user profile and behavior in other service providers or home service provider. These systems can stretch the detection of spammers that initiate low rate spam calls to the callees of particular service provider, but distribute calls to the recipients of many service providers. Collaboration among service providers would naturally improve the detection time and detection accuracy, but this depends on the amount of information exchanged during the collaboration process. Service providers are reluctant in participating and collaborating with other service providers because they are business competitors and are concerned about privacy of their customers and their internal network configuration. The challenge in the design of an effective collaborative SPIT detection system is three-fold. Firstly, determining what filtered information should be exchanged among collaborators so that service provider remains confident and take parts in the collaboration. Secondly, understanding with whom this information should be exchanged such that privacy of collaborating service provider is not compromised. Thirdly, determining what information should be returned back to the collaborating service providers.

3. Standalone Detection Systems

In the past few years, social networks and telephony (mobile, landline and VoIP) have become the most popular form of communication for instant messaging and interactive live communications. Spam over the Internet has long been the problem since the begging of web technologies in the form of email spam that results in an overall loss of tens of billions of Dollars annually [84]. For example the so-called "Nigerian scam", resulted in a loss of around US$12.7 billion in 2013 [82], and the “CEO Email Scams” caused financial losses to organizations to more than US$2.3 billion [83]. However, recently email spamming has dropped drastically [17], [85] as spammers are finding new ways and means by which to target users of other technologies such as telephony and social networks with the unsolicited communications. Spamming in an interactive media through VoIP, instant messaging and traditional circuit switched network is more annoying, damaging, and disturbing than email spamming as callees are required to respond the incoming call request immediately. These unwanted calls not only effect productivity of subscribers but also causes a financial loss to the target victim. Besides gaining financial benefits, spammers also try to distribute malware to infect recipient’s mobile and VoIP handsets or to find system vulnerabilities.

Several approaches have been proposed for combating spammers in a VoIP network. These approaches can be mainly categorized as: (1) content-based SPIT detection systems, and (2) the identity-based SPIT detection systems. The content-based detection systems process the speech streams that are being exchanged between sender and the callee for the identification of a famous spam phrase or word. The identity-based SPIT detection systems use identity of the subscriber (calling identity or IP-address) to monitor the behavior of the subscriber within the service provider network.

The stand-alone spam detection system monitors behavior of users within the jurisdiction of a single telecom service provider. This section outlines and surveys some of the standalone SPIT detection approaches that have been proposed in the literature for mitigating spammers in a standalone VoIP network. The taxonomy of standalone SPIT detection systems is given in Figure 3.

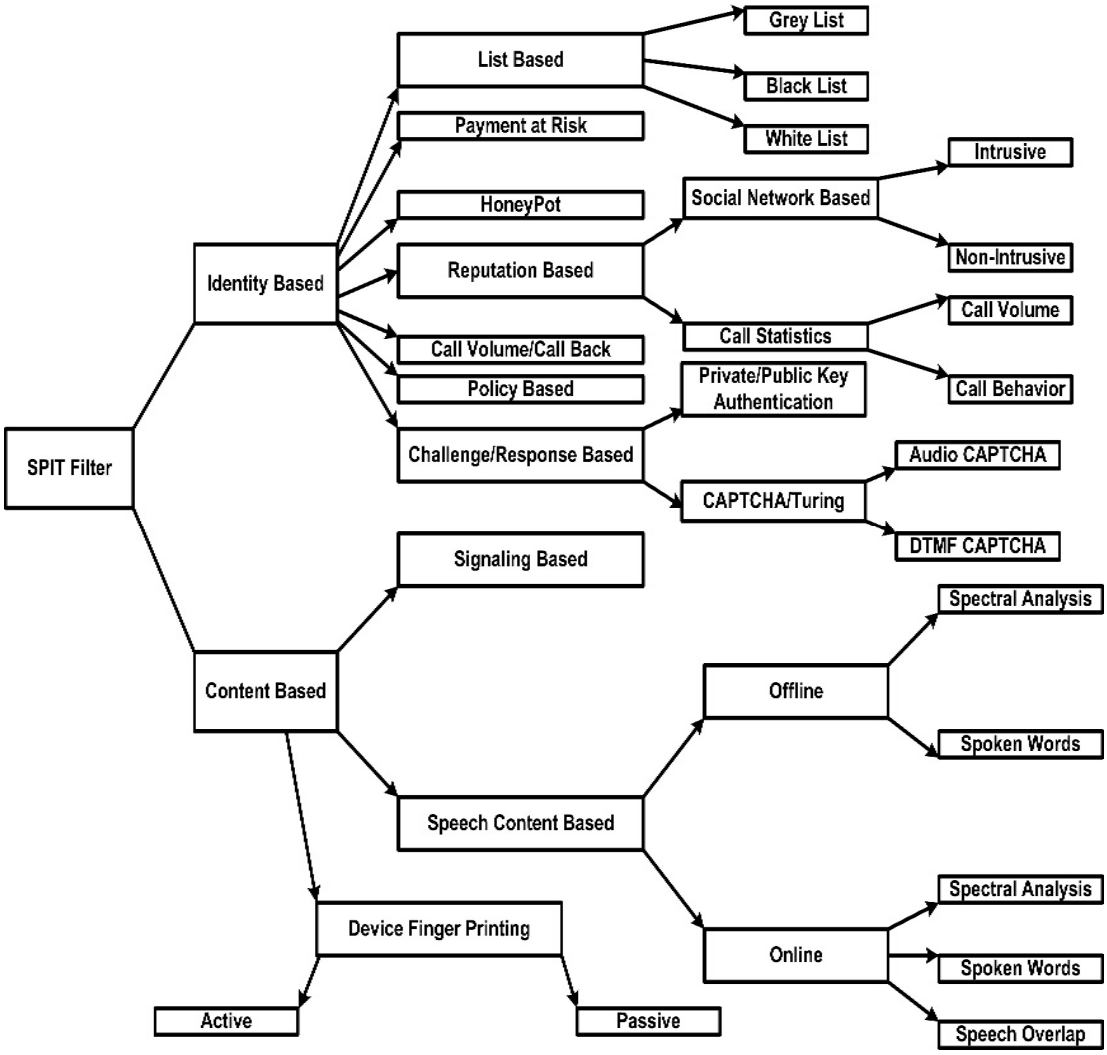

Fig. 3 A Taxonomy of Standalone Detection Systems for SPIT

3.1. Content-based Detection Systems

Content processing has been widely used for mitigation of email spams [18, 19], detection of spammers and spam content in the blogs [20], identification of malicious and spam web pages [21], filtering of spam messages in the online social networks [22], and filtering of SMS spam in a mobile network. The content-based approaches apply sophisticated machine learning mechanisms [18, 19] to the labelled (spam and non-spam) contents and classify new content as a spam and a non-spam. The content-based approaches have also been used in the VoIP networks for processing speech content in order to find famous spamming phrases and words.

In a VoIP network, content-based approaches can be classified into two categories: online speech processing and offline speech processing. In online speech processing, the detection system processes the speech content in real-time between sender and the recipient of the call, whereas the offline approaches process the speech

content stored on the voice mailbox and the media servers. A content independent offline speech processing method has been proposed in [23] that estimates similarity between speech samples left on the media server for unattended calls by the caller. The similarity between speech samples can also be estimated by estimating the distance between online speech samples and fingerprints [24 - 26] of speech samples left on media servers. The content-based approaches can also be implemented in the form of multistage systems which all are all integrated with other approaches in order to improve the detection rate and minimize false detection. According to authors in [27], a two stage Multi-layered Fusion-based method has been presented considered information from signaling messages and speech contents while making decisions about the behavior of a subscriber as a spammer or a non-spammer.

Content-based approaches have shown great resistance against spammers in emails or social networks. In these networks, contents are generally available in the form of text and images that do not require sophisticated system resources for processing and matching. Moreover, users of these networks are typically not worried about the delay incorporated between their conversations when spam detection systems have been deployed for analyzing the contents. However, in a VoIP network, applying content-based systems for spam detection have several limitations because of processing of voice streams in real-time. Firstly, the deployment of speech processing engine would introduce some noticeable delays between conversations of subscribers and this delay greatly annoy the subscribers as they do not want to have long delays in their conversation. Secondly, the service provider requires sophisticated signal processing system and network resources for the processing of online speech streams in near to real-time. Thirdly, content-based approaches examine and make a decision if the caller at a second stage and spammers have already annoyed call callees with the spam signaling (telephony ringing) and listening to some seconds of call. In the perspective of spammers, content-based approaches can easily be mitigated by slight modification of speech contents and addition of noisy speech streams.

The call setup messages are exchanged between subscribers which are usually available in the form of plain text. The call setup messages can also be processed for the identification of spam content and spammers. However, the structure and messages exchanged during the call setup phase and termination phase does not have valued information to be used for spam detection. In addition, the call setup messages of spam and non-spam calls are same and do not provide any information to be used for classifying subscriber as a spammer or a non-spammer. It is very easy for spammers to devise call setup messages so that it would appear that the call was originated from a legitimate caller. In some scenarios, content-based approaches would not have a global implication as people from different regions have different calling and greeting behavior. Lastly and most importantly, in most countries listening, recording and processing speech streams of subscribers is prohibited by law, thus the privacy of the subscribers is not protected at all.

3.2. Challenge/Response-based Detection Systems

A C/R (Challenge/Response) system is a type of spam filtering system that initiates an automatic challenge to the subscriber and subscriber needs to reply correctly to the given challenge. The C/R-based system allows subscriber to call if the subscriber correctly responds to the challenge and subsequently blocks subscriber if he fails to solve the challenge. Spam call can be generated by the human or an automated machine. Humans (legitimate and non-legitimate) can easily solve the challenge initiated by the proxy server, whereas machines would not be able to solve initiated challenges in a timely manner. The C/R-based anti-SPIT approaches can be implemented in two ways: 1) through authentication and authorization carried out in a non-intrusive way without involving the subscriber for the explicit response and instead using encryption or handshake mechanism with the subscriber at the time of registration and call request [28] in a seamless way, and 2) a method whereby the proxy server initiates a CAPTCHA challenge to be solved by the subscribers [29 - 31] in a timely manner. The CAPTCHA can be extremely annoying to subscribers as it takes time for them to respond to, especially if the subscribes need to make a call immediately or urgently.

In [28], subscribers are asked to successfully pass through two authentication phases before placing a call to the callee. In the first stage, authentication of subscriber is carried out through digest access authentication – exchange of credentials on which the VoIP proxy server and the subscriber agreed upon such as username or password, and in the second stage, a proxy server authenticates the subscriber through transport layer security and DNS service records. Legitimate subscribers typically have pauses in their conversations, whereas spammers do not have pauses especially at the start of the conversation that results in a speech overlap between spammers and their callees. The authors of [31] proposed a hidden Turing test mechanism that considers talking behavior of subscribers and monitors the overlaps in the speech streams flowing between subscriber and his call recipients. In [29], service provider asks subscriber to solve a Human Interactive Proof (HIPs) test by pressing phone keys against initiated challenge. Text and image-based CAPTCHA are not feasible to deploy in a VoIP or voice network, therefore speech-based CAPTCHA is the only available option for initiating the challenge to the subscriber. The subscriber will then respond by pressing combination of certain phone keys in the form of DTMF (Dual Tone - Multi Frequency) signals. In [32], audio CAPTCHA is generated to the subscriber and subscriber responds to the challenge that asks for pressing a specific key on the phone keypad. In [33], the authors combine audio CAPTCHA and game theoretic model for authenticating the subscribers. Besides, solving the challenge and providing the credentials, subscribers are also asked to call back the service provider [34] for proving their authentication. The approach presented in [34] uses anonymous verifying authority and mediator for initiating call back request and block subscriber if call back request is not fulfilled within a specific time period. In [98] author proposed a “voice CAPTCHA” that tells apart the calls made by the humans from those made by the automated machines in a VoIP network.

A C/R-based system can also be combined with other anti-SPIT systems so that challenge is initiated only to few subscribers or to only new subscribers. The multi-stage C/R-based system would minimize the false detection rate and improve the detection accuracy. C/R-based approaches are well suited for blocking machine or auto-dialer spammers but the deployment of such prevention mechanism in a real-network has a number of problems. For one thing, authenticating subscriber via CAPTCHA or public-private key would introduce notable call setup delay which would definitely introduce an annoying factor to the subscriber for each call made. However, public-private key-based C/R systems are non-intrusive but they require public-private key infrastructure for assigning keys. CAPTCHA test is intrusive to the subscribers and additionally has some other limitation. A CAPTCHA system usually requires solving some challenges that might be difficult to solve by the people with a certain degree of disability. From the perspective of spammers, CAPTCHA-based approaches can be circumvented by hiring cheap human spammers. For example, spammers can set-up a network of cheap workers for solving the audio CAPTCHA challenge and then relay voice streams from the media server upon successful authorization of the call. From the perspective of service provider, C/R-based system especially CAPTCHA system requires additional network and system resources for initiating and processing of challenges and their responses for a large number of subscribers in parallel.

3.3. Access-Control-List-based Detection Systems

The Access-Control-List (ACL) -based approaches are the simplest identity-based anti-SPIT filters. The proxy server checks the database list during the call setup phase. Whitelist maintains the database of identities that are allowed for using the network and blacklist maintains the database of identities barred from calling. A list database can be either global in scope (applied to all subscribers) or local in scope (applied to particular subscribers). In addition to black and white lists, a grey list [35, 36] can also be used for maintaining the list of subscriber to be observed for further time. A grey list maintains the database of identities that are being challenge for the authorization and monitored for the some extended times. In a grey list, if a subscriber exhibits spamming behavior over some pre-defined extended time then subscriber is permanently moved to a black-list, otherwise subscriber is moved to the whitelist. The list-based approaches need to be implemented along with other approaches that actually decides which list the user should be placed in [13, 14, 37].

A common problem with blacklist-based SPIT filtering is the inclusion of legitimate subscribers who were mistakenly reported or classified as the spammers because of personal dislikes. Additionally, it becomes difficult for subscribers to get their identities out of the blacklist. Spammers can easily evade the blacklist database by spoofing identities of legitimate subscribers not included in the blacklist. A whitelist is useful for controlling which subscribers are allow for calling globally and locally. However, it would not allow new subscriber or a legitimate subscriber wishing to call someone for the first time. From the perspective of service provider, maintaining a large global and local list database is problematic and requires continuous update.

Moreover, list-based approaches need to be implemented with other SPIT detection approaches that actually decides whether to include subscriber in the respective (black, white and grey) list or not.

In many countries, telecommunication regulators encourage consumers to put their telephone number in the do-not call register mode, if they do not want to receive any advertising or marketing calls. However, this mode is not effectively resistant against spammers, as some legitimate and the illegitimate companies do not pay attention to these lists or integrate them into their preventive strategies and approaches. Statistics show that a large number of users still receive a large number of spam calls despite putting their numbers in the national do-not call registry [90].

3.4. Cost-based Detection Systems

VoIP offers affordable calling rates that has not only attracted legitimate subscribers but has also attracted spammer for making unsolicited advertisement calls to a large number of callees. One way to block spammers is to impose some extra cost on subscribers classified as spammers. In this scenario, a service provider first deducts some money from the subscriber credit and returns back if subscriber is classified as legitimate later at end of call. If a subscriber classifies as a spammer later on, then the service provider will withhold the deducted money and distributes this money among spammer's victim callees. The cost-based approach needs to be implemented with other approaches [38] that actually decides whether to return back the money or not.

A major limitation of cost-based system is that it requires comprehensive micro-payment system for the computation which involves the deduction, recording, and tracking of charges and fees. Moreover, cost-based systems are dependent on other SPIT classification methods for the final decision. It might also be possible that cost-based systems would charge legitimate subscriber and would also mistakenly block legitimate subscriber if subscriber has insufficient balance required for the deduction but greater than the amount required for the call.

3.5. Policy-based Detection Systems

In a policy-based anti-SPIT systems, service provider defines policies that instructs anti-SPIT system to monitor the behavior of incoming call request according to subscriber's network-wide policies [39]. In [40], the authors proposed policy-based anti-SPIT system that maintains subscriber preferences and policies using call processing policy framework [41] for blocking the spammer. In [42], the authors used adaptive policy system based on the definition of a set of rules along with actions and controls for mitigating the SPIT attack. The authors in [43] presented a system that uses Security Assertion Markup Language (SAML) for authenticating the subscriber in the network. The policy-based systems are well suited for the small operators or local enterprise organization because maintaining policies for thousands of users in a big operator are difficult and unmanageable task. In [89] author proposed an adaptive policy-based SPIT management framework that detects SPIT callers using an adaptive anti-SPIT policy-based framework (ASPF).It incorporates a set of adaptive rules

that are based on the behavior spammers and non-spammers in a VoIP network along with set of actions and controls to block the SPIT caller.

The major limitation of policy-based SPIT detection systems is the maintenance and updating of subscriber public and private policies. Moreover, policy-based systems would add noticeable delay during call setup phase because of processing of policies of subscriber and his behavior towards these policies. From the perspective of spammer's attack, policy-based system can be circumvented by spammer through spoofing the policies of the legitimate callers.

3.6. Legislation-based Detection Systems

The primary goal of the legislation is to create a legislative framework that would make spamming illegal and impose punishment on those involved in spamming activities. Governments including European Union, USA and Canada have already made some good progress and reasonable efforts in terms of legislation against initiators of spamming [44 - 46]. These legislations prohibit unsolicited communication to reach the callee unless prior consent of the callee is obtained. The major limitation of legislation-based anti-SPIT system is the difficulty of tracing back the initiators of spam communication for the law enforcement agencies. Furthermore, if the regulator or law enforcement agencies trace-back the initiator of unsolicited communication even then there is no such global law exists that will apply to spammers across the world. Moreover, spammers make spams from anywhere around the world thus make anti-spam law of one country inapplicable to the spammer spamming from places where no such law exists.

3.7. Call Statistics-based Detection Systems

A VoIP call consists of two parts: 1) a signaling phase which consists of a series of call setup message exchanged between subscribers and the proxy server at the time of call request, and 2) a speech streaming phase which includes the exchange of actual speech payload between caller and the callee after the call establishment phase. The statistics-based SPIT detection systems monitor different call statistics of the subscriber that are: the call-rate, the call duration, inter-arrival time between call requests made by a subscriber, number of speech packets exchanged between subscribers, which subscribers disconnected the call etc. Several machine learning approaches have also been applied to call statistics and calling behavior of subscribers to separate spammers from the non-spammers [50]. The information from call statistics (call rate, call duration etc.) and calling behavior (number of friends, number of incoming calls etc.) of subscribers can also be used for computing reputation of the subscribers which is then used to block subscriber if reputation of subscriber is less than certain threshold. In [87], authors combined the set of features from the call detailed records and SIP signaling messages for detecting SPIT caller in a VoIP network. They evaluated their approach using synthetic data sets.

The authors have not provided any details as to why these features have been adopted and which feature shows optimal performance and under which conditions.

The major limitation of statistics-based approaches is learning the behavior of legitimate and non-legitimate subscribers. It would be difficult in a statistical system to differentiate the spammer from the non-spammer with small false positives. From the perspective of spammers, statistics-based systems could be easily circumvented by spammers by controlling the similar call statistics and by spoofing identity of legitimate subscribers. From the perspective of deployment, learning a dynamic threshold for each subscriber and for each aggregation cycle is also challenging and problematic.

3.8. Device Fingerprinting-based Detection Systems

A device fingerprint is the information that has been recorded on a proxy server or a remote computing device for the purpose of identifying devices and software used by the subscriber. The fingerprinting-based anti-SPIT systems provide fingerprints to a set of devices used for making calls in a VoIP network. The fingerprinting-based approaches assume that spammers and non-spammers use different telephony devices and communication protocol stakes for making and receiving calls. These approaches are categorized in two types: active fingerprinting and the passive fingerprinting. In active fingerprinting, remote device asks subscriber for transmitting packets to remote system for device analysis. However, in the passive fingerprinting, remote device actively monitor the fingerprints of device originating the calls. In [51], the authors analyze fingerprints of several commercial hard and soft phones for different call response messages using active and passive fingerprinting. The use of device fingerprinting in real deployment is not practical and scalable as it requires careful management of fingerprints large number of commercial and non-commercial VoIP devices. Additionally, spammer can bypass fingerprinting-based systems by adopting fingerprints and protocol stack similar to the devices used by the legitimate subscribers.

3.9. Honeypot-based Detection Systems

Spammers typically crawl the web or telephone directory to harness and collect target identities without knowing whether target identities are real or fake. Honey phones are virtual phones not assigned to human users but are used for analyzing the behavior of subscribers making calls to them, with the purpose of learning more about the origin of the spammers and the techniques they are employing. As honey-phone is not assigned to any physical person thus naturally would not make any call to users or receive any call from the legitimate subscriber. Subscribers calling honey-phones could be spammers but confirmation requires analysis of subscriber behavior towards honey-phones [3, 52, 53] and other network users. In [52], the authors deploy a large scale cloud-based honeynet system for analyzing the social behavior of callers making calls to these honey-phones. In [54], the authors proposed a MobiPot, which is a honeypot mobile system for collecting the

fraudulent calls and SMS. These calls and SMSs are then analyzed for identifying the mechanism used by spammers for collecting the identities of target victims and spamming attack patterns. Honeypot-based solution can learn about new spamming techniques and identify spammers who target the enterprise network users.

3.10. Reputation-based Detection Systems

Collaboration among subscribers can assist subscribers who wish to make decision about whether to receive or reject the call from subscribers not known to them. The reputation-based systems operate in two ways: (1) Distributed in which each subscriber directly collaborates with other subscribers, and (2) Centralized: in which the subscriber directly collaborates with the centralized service provider or system. Several reputation-based anti-SPIT systems have been proposed for filtering spam in a VoIP network. These systems consist of two steps: computing direct trust between subscribers engaged in communications and then aggregation of direct trust scores of subscribers for their global behavior. The direct trust represents strength of direct relationship between subscriber and his interacted subscriber, in which the global reputation represents aggregate behavior of subscriber towards all his interacted subscribers. If the score comprising the trust and reputation of the subscriber is higher than some learned or fixed threshold then subscriber is considered reputed; otherwise, subscriber is considered non-reputed.

The direct trust between the subscriber and his interacted subscriber can be computed in two ways: 1) Intrusive way which implicitly requires interaction with the call recipient of the subscriber [13 - 15] for the feedback about subscriber, and 2) Non-intrusive way which explicitly utilizes information from call logs recorded for the billing purposes [10, 16, 55]. The global reputation score of the subscriber can be computed by either applying Eigen Trust [16] to the direct trust scores or by applying machine learning approaches such as Bayesian networks and clustering to the direct trust scores of subscriber [13, 14],

Spammers typically exhibit different calling behaviors from the legitimate subscriber with the following properties: they target large number of callees, receive calls from only few callees and many of their received or made calls are of small duration. This calling behavior typically results in a spammer's disconnected social network with the large number of their target victims. On the other hand legitimate subscribers normally have small number of callees, have repetitive calling behavior with many of their called callees and also receive good number of calls from their called callees. This calling behavior of legitimate subscribers results in a subscriber's strong relationship network with many of his callees. In [16], the authors proposed a Call-Rank system which is a reputation-based that uses average call duration for computing direct trust between subscriber and his called callees. The global reputation of subscriber is then computed by applying Eigen trust algorithm to the subscriber's direct trust scores. Finally, Call-Rank asks the callee for the final decision (to either accept or reject the call) by sending reputation scores and social network credentials of caller to the callee. In [56], the

authors considered subscriber interaction with his callee as reputed if the call duration is greater than 20 seconds. In [57], the authors used call duration along with seven degree separation as a social network feature for the computing of reputation of the subscriber across the network. In [58], the authors proposed three reputation-based solutions for filtering SPIT subscriber. In the first approach, the authors in [58] used concept of strong and weak social ties among subscribers, in the second approach, they enhanced Progressive Multi Grey-Leveling list to the Enhanced Progressive Multi Grey-Leveling by using call density and reciprocity-index features, and in the third approach, they adopted Page-Rank algorithm for computing global reputation of the subscriber.

The direct trust between the subscriber and his callees can also be computed by collecting the feedback (positive or negative) from the callee for the subscriber's call transactions that just ended. In [15], the authors computed reputation of the subscriber by aggregating the callee's feedback about the subscriber's calls. The reputation scores and call statistics of the subscriber is then used along with MPCK-mean -- a semi supervised clustering algorithm, to cluster the subscriber into a SPIT and a non-SPIT clusters. The proposed system performs well only when the callee provides honest and accurate information about the caller. In [13, 14], the authors proposed a multistage system for the identification of spammers. The system consists of three stages that collaborate with each other. 1) The direct-trust stage computes the direct-trust of the subscriber by aggregating the feedback from the callees for his past calls; 2) the global reputation stage that computes reputation of the subscriber by applying the Bayesian network algorithm to the direct scores of the subscriber assigned by all his callees, and (3) the list database that maintains list of black and white listed subscribers. In [59], the authors proposed an approach that aggregates callee's feedback along with G-mail spam filtering method for blocking the spammers. In [60], the authors proposed two approaches based on the content processing and feedback aggregation about behavior of the subscriber from his callees. In [61], the authors computed reputation of the subscriber through a web of trust model between subscribers in a network. In [62], the authors proposed a multilayer system that incorporates behavioral characteristics of subscriber and his call signaling messages. One of the major limitations of intrusive reputation approaches is their intrusiveness and requirements to change in a VoIP handset and call signaling messages. Moreover, spammers can easily circumvent intrusive approaches by creating network among his identities and providing fake responses for their identities.

Reputation-based anti-SPIT systems have provided an effective protection against spammers in email and VoIP network but their effectiveness depend on the set of features used for the computation of global reputation. In some cases, a spammer could get high reputation scores by creating a Sybil network between his acquired identities and spoofed identities of the legitimate subscribers. The reputation systems could minimize the effect of Sybil attack but its underlying performance depends on features used for the computation of global

reputation scores. The spammer typically targets large number of callees without repeating his callees, thus typically resulting in small duration calls to a large number of callees and good duration calls to only few callees. In non-repetitive calls, the average call duration of caller with the callee is same as of his aggregate call duration. This behavior might result in a high reputation scores for the subscribers having good duration calls to a large number of their called callees. Moreover, spammers can also collude among their several identities with good duration calls that would also improve the score of their global reputation. In some CDR-based approaches, the global reputation score and social network credential of the caller are also sent to a callee for the final decision [16] which is not only intrusive to the callee but also poses threat to the privacy of the caller.

3.11. Multi-Stage Detection Systems

From the analysis of standalone systems, it is clear that no single detection approach provides the high true positive rate and the small false positive rates. Therefore, it is necessary to have SPIT detection system that combines different standalone system to improve the true positive rates and minimizes the false positive rates. The Multi-Stage anti-SPIT systems entail internal collaboration among many independent standalone systems which are integral parts of the service provider telecom network to improve the detection accuracy and detection time. The multistage systems collectively utilize information from many standalone components or systems in order to make decisions about behavior of the subscriber.

In [27] authors present two-stages system that processes speech streams and signaling messages of the subscriber in order to block the SPIT caller. Similarly, the authors in [47] presented a two-stage system that is based on the CAPTCHA test and a speech-processing phase. In [14], authors presented a content independent multistage system consisting of three major stages. 1) a Rate limiting stage -- which monitors call statistics such as call rate in a certain time period, 2) a blacklist module -- which places spammers in a blacklist based on its behavior, and 3) a multivariable Bayesian network -- which determines the reputation of a subscriber based on subscribers score assigned by its callees. In [30], the authors used different thresholds for computing reputation at each stage of multistage detection system. In a first stage, the system computes reputation score of the subscriber and compares this reputation score against two thresholds – low and high thresholds for forwarding call to the next stage. In a second stage, the system invokes Turing test for authorization and in a third stage the system asks call recipients for the feedback about the initiator of the call. In [48], the authors presented a four stage system for blocking spammers in a transit VoIP network using internal collaboration between many stages. In [36], the authors used subscriber's short and long time call patterns and blocks spammers using collaboration among various list databases – white, black and grey list. In [49], authors proposed multistage system that incorporates feedback from multiple stages while making decision about subscribers as a spammer and a non-spammers. The multistage stage system entails the collaboration among well-known detection

systems such as blacklists, whitelists and call statistical analysis (call duration and call rate) into a multistage SPIT detection system.

The multistage systems typically combine sequentially the output from standalone SPIT detection. The systems can be standalone as an individual modules, or components where each component makes decisions based on the nature and behavior of the callers once the call is passed through all the combined components. This is suitable for combining the techniques that use information from multiple components. For example, caller information is first processed using SIP signaling message, and subsequently, and the caller is presented with the CAPTCHA test to prove that he is a human or a machine. However, these systems would increase the call setup delay – which can annoy the end-users. Furthermore, some multistage phased systems require feedback from the caller and is intrusive to the callers. The system can also be implemented in a way where standalone system analyse the call in parallel and assign weights to the call, which then gets combined to produce the final score by applying weighted aggregation methods. The final classification is carried out based on final aggregated scores. This multistage system is suitable for combining the feedback in parallel from various standalone systems. However, the primary challenge for deploying these systems is determining accurately the value of the weights to be assigned for each standalone systems.

It is obvious that multistage systems would provide better detection accuracy and detection time but their performance depends on the types of standalone systems used together. The content-based multi-stage systems have same limitations as of approaches based on the speech content processing. Similarly, the CAPTCHA and reputation-based systems require interaction with the caller and the callee, thus are intrusive to the end-users. Another major issue with multi-stage systems is that they would introduce considerable high delay during call setup phase and incurs high deployment cost.

4. Collaborative SPIT Detection Systems

Spammers always try to find ways for evading the spam detection systems. They can manipulate content by adding noisy messages, acquire large number of identities for colluding and Sybil attack, and controlling number of spam calls to a single service provider but target many service providers in parallel. Furthermore, spammer can make a large number of spam calls in aggregate from a given calling identity but distribute calls among many service providers. Existing anti-SPIT system decides about behavior of the subscriber-based on his calling behavior observed at a single service provider. These approaches prolong detection time and block spammers only if spammers make significantly large number of spam calls to a single service provider. The unavailability of calling behavior of spammers across the service providers limits standalone system to react effectively against spammers.

The calling behavior of a subscriber becomes more meaningful when subscribers are observed across many service providers. Naturally, collaboration among service providers would improve the detection time and detection accuracy because of collective use of information about behavior of subscriber from many autonomous collaborating service providers. Service provider or domain collaboration has been applied in various networking domains for identifying malicious or unauthorized intruders and spammers in a network. In this section, we provide detailed discussion on the collaborative systems for SPIT.

In order to evade the standalone systems, spammer targets large number of callees that are dispersed across many service providers but his calling behavior remains effectively the same across all telecom service providers. Collaboration among service provider is a natural way for early detection of spammer before they spam a large number of callees.

A number of collaborative systems have been proposed for filtering spammers in an email network and detection of intruders in the IP-based networks. From the perspective of email networks, the collaborative systems typically require collaboration in the form of exchanging spam message contents, spam HTML (Hyper Text Markup Language) tags and exchanging feedback about behavior of particular sender. Typically spammers send the same spam message to a large number of callees and collaboration with the message content would greatly improve the spam detection time but it poses threat to the privacy of email users. In [68], the authors presented a privacy-aware collaborative system called ALPACAS that invokes collaboration among service provider with the exchange of encrypted fingerprints of message contents. The similarity between fingerprints of user messages and messages stored in the database of ham and spam is computed to classify the new message into a spam or a ham. HTML tags are available within the email headers and can be used to distinguish spam content from the non-spam. In [69], the authors presented a COSDES system that collaborates with the exchange of HTML tags and computes distance between spam and ham tages using near duplicate approach. In [70], the authors presented a collaborative framework that entails distributed collaboration for making decisions about users. The collaborating email domain directly exchanges information to each other and imposes some restrictions on domains not taking part in a collaboration process.

The behavior of spammer remains the same in all target service provider or domains, and when analyzed collectively would decrease the detection time. In [71], the authors proposed a system called SpamTrackers that requires collaboration among domains with the communication behavioral patterns of email senders within collaborating domains. In [72], the authors presented a three layered P2P architecture-based on communication patterns of spammers and non-spammers. The architecture requires collaboration among end-users, mail service and the super peers. The super peer handles the exchange of messages among themselves for tagging and subsequently classifying incoming mail digest as a spam or a non-spam. In [73], the authors presented RepuScore that requires collaboration among email domains with the exchange of local reputation score of

email sender within the domain. In RepuScore, a centralized system computes global reputation of user by aggregating local reputation scores. In [74, 75], the authors proposed social filter, a centralized system that aggregates feedback from individual spam detection systems for early detection of Phishers and spammers.

Collaboration can also be carried out directly among end-users. In [76], the authors presented a collaborative spam filter that uses collaboration and social network information of users for blocking spammers. In [77], the authors proposed two spam detection systems: a simple Mail-Rank and a personalized Mail-Rank that computes global reputation of email user by aggregating direct trust score through power iteration algorithm. Spammers are moving to different online social networks for increasing their footprint. The collaboration among the different social network platforms would considerably improve the detection accuracy and detection time. In [78], the authors proposed a system that incorporates collaboration among different online social network platforms with the exchange of spam contents to be used for effective spam detection.

All of the above-proposed collaborative approaches classify incoming email into spam or non-spam by analyzing the contents of messages, contents of HTML tags and static rule for some features. These approaches cannot be applied directly for filtering spammers in a voice networks. In a voice network, contents are available in the form of speech signals and having collaboration with the exchange of speech signals is not feasible. Moreover, it requires sophisticated system and network resources for speech processing, storage and matching. In case of non-content-based collaborative approaches, the detection approaches make use of structure of the network while ignoring the weights on the links between users. In a voice network, few additional features such as call rate and call duration could provide information about relationship strength among users and should be used in a collaborative way.

A very few works have been reported that incorporate collaboration among service providers for rating the subscribers. In [11, 63], 3GPP a standardization body on next generation network formalized best practice standards for fighting spammers in a VoIP and IMS networks. Particularly, they encouraged service providers to have collaboration among themselves for early and effective IM (Instant Messaging) and voice spam detection. However, these technical standards do not provide any information about how collaboration is achieved. In [64], the authors exchange scorecards of subscriber among collaborating service providers and collaborating service providers react independently against the subscriber's scorecard whether to allow or block the subscriber. This framework requires predefined trust relationship between collaborators which is in reality and practice not feasible to implement and integrate in a telecommunication network. Additionally, no mechanism has been presented for the computation of subscriber score and trust assessment of collaborators. In [97], the authors present a decentralized system for SPIT detection that incorporates the privacy protected collaboration among the telecommunication service providers. The privacy of telecommunication operators and their customers is protected through the exchange of encrypted reputation scores among telecommunication

operators. The experimental results revealed that proposed approach outperforms standalone systems in terms of true positive rate and false positive rate.

In [65], the authors proposed a collaborative system where the home service provider collaborates with the visiting service provider with the exchange of information about their SPIT detection system in the form of call tags. The visiting service provider then would access the performance of SPIT detection system deployed in a home service provider of the subscriber calling recipients of his network. This approach has three obvious limitations: 1) It only evaluates performance of a SPIT detection system deployed in a service provider that provides the tag information, 2) It requires establishment of predefined trust between collaborating service providers, and 3) It requires changes in the call signaling messages exchanged between collaborating service providers in order to incorporate tag information. The authors in [66] proposed SPACEDIVE that detects intrusion in a VoIP network by correlating local and remote information from many collaborating service providers. In [67], the authors presented P2PAVS that computes reputation of subscribers through collaborative response from the subscriber from many service providers. However, it does not provide any mechanism for sharing and propagation of trust among subscribers. In VoIP typically collaboration is achieved in the form of multistage systems or collaboration among proxy servers within the service providers. In [33], the authors proposed an approach that uses collaboration among several local VoIP servers within the service provider. However, mechanism for collection and aggregation of feedback between subscribers and proxy servers has not provided. In [86], the authors proposed a game-theoretic based system to combat the spam calls in a VoIP service provider by having a competition among VoIP service providers. The approach is based on the fact that two different service providers adopt different strategies to use the shared resource to increase their profit.

5. Concluding Remarks

To date, anti-SPIT systems can be divided into two main approaches: 1) content-based filtering and 2) identity-based filtering. Both of these approaches have clear limitations as spammers are finding new mean by which to evade them. In Section 3, we outlined existing anti-SPIT systems and their limitations. Content-based system could block certain type of spammers but they can be circumvented by spammer through slight modification of speech content. Moreover, applying content-based approaches in VoIP for filtering spammer has some additional limitations. Firstly, content-based approaches make decision about the caller after the call has already been established and is already late as the spammer has already reached the callee with the ringing message and initial contents. Secondly, processing speech requires extensive computation and storage resources and incur additional delays in conversation among users. Thirdly, content-based approaches do not have global implication because of different conversation languages and talking behavior. Finally, caller and callee would not provide consent for monitoring and processing their conversation. Identity-based reputation systems are

viable for blocking spammers in a VoIP network. These approaches mainly consider call and social network statistics of the caller but can also be evaded by the spammers if approaches are not carefully designed as discussed in Section 3.

Existing reputation-based anti-SPIT system still exhibits clear limitations. Firstly, anti-SPIT system must involve interaction with the caller or callee at any stage of detection process. Secondly, there is a need to identity set of features to be used collectively so as it become difficult for the spammer to evade the detection system. Thirdly, there is a need to have automatic classification of caller as spammer or non-spammer with high true positive rate and small false positive rate. A number of reputation-based systems have been proposed [13 - 16] and all are non-intrusive. These approaches either involve caller for authorization or involve callee for the final decision and feedback about the caller. Call-Rank [16] though computes caller's reputation automatically but it relays on callee for final decision and has other limitations. Firstly, it discloses caller's reputation scores and social network credentials to the callee which could be threat to the privacy of caller. Secondly, it requires public and private key infrastructure for authentication and authorization. Thirdly, it asks caller to solve the CAPTCHA challenge to minimize the false positives and handle calls from new callers. However, we believe that anti-SPIT systems must be non-intrusive and should number of call and social network features collective. Moreover, classification threshold needs to be computed automatically without incurring high false positives and small true positives. The collective use of social network and call features in both directions (incoming and outgoing) would considerably improve the detection accuracy and minimize interaction with caller or the callee. It also becomes difficult for spammers to evade such system because they are not able to control multiple social and call features.

The standalone system whether based on reputation or content prolongs detection of spammers if spammers make small rate spamming to callees but targets callees of many service providers. Collaboration among service providers is the natural and practical approach for blocking these small rate spammers in a timely way. However, to-date, there is no such collaborative system exists, whereby telecommunication service providers collaborate with each other and exchanges the information about behavior of certain caller in their network. Moreover, Service providers are not willing to take part of collaboration with other because they business competitor and worried about privacy of their customers. The challenge in the design of collaborative system is to convince telecommunication service provider for taking part in the collaboration process. The simplest collaborative approach to convince service provider is to use the trusted centralized system so that service provider directly collaborate with the trusted centralized repository. However, service provider requires absolute protection of privacy of their customers and only interested in exchanging such information which does not contain any threat to privacy of their customers. In this article, collaboration is carried out in the following two ways: (1) It corporates use of trusted centralized repository for aggregation of information

provided by collaborators, and (2) It uses filtered non-sensitive information from the collaborating service provider. With this, service providers locally computes reputation of user using local call logs and exchange reputation score to the centralized repository for reputation aggregation and decision. The adversary or intruder at centralized repository is not able to infer the relationship network of end-user thus privacy protected at all.

Spammers can have a large number of identities because of VoIP cheap telephony rate. These identities either collude with each other for high reputation score so that to make spam calls to other users or use new identity once blocked by the service provider. However, a spammer having multiple identities has overlap in target identities among his different identities and exhibits similar calling behavior towards many target identities. The standalone reputation-based anti-SPIT systems react very slowly towards these spammers having large number of identities. It is of utmost importance to correlate and link the identities that belong to one physical person so as to collective using the reputation of caller while making decision. The identity linking is not only beneficial for timely identification of spammers but also helpful in characterizing the complete behavior of users having many identities. To date, there is no anti-SPIT system exists that computes reputation of the physical caller by connecting all of the caller's multiple identities. Moreover, to the best of our knowledge, there exists no such system that links identities that associate to one physical caller using callers behavioral and call features in telecommunication network.

The notable limitations of existing anti-SPIT systems call for designing a SPIT filtering system that blocks spammer through collectively use of call and social network features without interacting with the caller and callee. To accomplish the objectives of effective non-intrusive SPIT filtering system, this system proposes the following. First, a standalone detection system is presented that collectively uses social network and call feature of users while computing reputation of users. The standalone system relies on the rational that spammer and non-spammer exhibit different calling behavior in terms of call-rate, duration and social circle. To this end, the design system neither requires any interaction with user nor does it require any change in call setup messages and user hardware. Second, a collaborative anti-SPIT system is presented which incentivize service provider to take part in the collaboration without disclosing any sensitive information that characterizes the relationship network of end-user. To this extent, the system utilizes trusted centralized repository and exchange of filtered information which does not provide any information which can be a threat to privacy of the user and the service provider. Third, an identity linking approach is presented that connects identities that belong to single individual together. To this extent, we used social and call features of caller while computing similarity between the identities from two different time periods within the network. The identity link would help in characterizing the complete social behavior of a physical user having multiple identities.

6. Future Directions

As discussed in this paper, the existing proposed spam detection systems have clear limitations and entail future research work for enhancement and improvement in many aspects. Future research directions for anti-SPIT detection and mitigation can summarized as follows:

Identification of Other Call Features. In this paper, we have used call and social network features for the computation of global reputation of the caller within the telecommunication network. However, the presented work can be extended to leverage the use of call statistics information such as call release cause codes, inter-arrival time between call requests made by the caller, number of calls disconnected by the caller and callee of the call, number of failed calls, etc. Our aim is to integrate these call features along with the Caller-REP system in order to understand the impact of these features on the detection performance. We would also like to investigate the ways for personalized spam detection where the spam calls are only allowed to specific callees who wish to receive these spam and promotion calls.

Secure Multi-party Computation and Collaborative Systems: In this paper, we have relied on the use of centralized repository for reputation aggregation and decision about behavior of the user across all service providers. However, the centralized repository can be a single point of failure. Also, it would not be scalable, and service providers might not be willing to collaborate with the centralized repository. Alternatively, the service provider might exchange encrypted scores with the peer service providers. Hence, the idea of distributed collaboration can be a viable option, whereby each service provider directly collaborates with his peer service provider without worrying about privacy of his customers. This collaboration approach has two challenges: 1) privacy protection of collaborators data, and 2) convergence and communication overhead required for reputation aggregation. To achieve these objectives, the design of a Secure Multiparty Computation (SMC) scheme without overwhelming the network bandwidth can be an adequate solution.

Collaboration among different Mobile Applications. The proliferation of smart phones has attracted people to replace their legacy phones with smart phones and use newly developed VoIP applications for free voice conversation with friends and family. There is a number of free VoIP calling application. WhatsApp, Viber and Skype are some of the famous ones. The free calling rates and the ease of integration of these applications with the spamming tools have also attracted spammers and scammers to target users of these applications with the unwanted calls and messages. Mostly, these applications operate using mobile number of users except for Skype that uses email identity. A spammer can exploit the users of these mobile apps in parallel or one by one. However, with intra-application collaboration the spamming effect could be minimized and spammers are identified at earlier. This work can be extended towards the design of a spam detection application that collectively uses information from all VoIP application and aggregately characterize the behavior of the caller

targeting an application user. We have already developed an android mobile application as part of a master thesis [79] for detecting spammers on a standalone android mobile device using user's call logs and collaboration among end-users. The design of intra-application collaboration has a challenge of developing a common application programming interface for excessing the data from different applications.

Caller ID Spoofing. Caller ID spoofing occurs when someone changes the Caller ID to any number that he wants to display at the receiving end. Caller ID is meant for providing information to callee about the caller. There are several techniques and smart phone applications are available that can be used to spoof the identity. It can be used to identify the telemarketers and spammers; however, spammers, telemarketers and prank callers are spoofing identities of legitimate users to circumvent the detection system and fraud user by pretending to be a legitimate entity. It is important to design a system that can identify the users hiding their original identity. The challenge in this regard is two-fold: 1) Making decision during the call setup phase, and 2) Making decision about identity of the caller without involving caller and the callee for a certain response messages. The call setup messages can provide complete information about the caller location, calling identity, devices, etc. and can be used for blocking illegitimate user spoofing their identities.

Acknowledgement

We thank the associate editor and the anonymous reviewers for their assessment and valuable feedback which significantly enhanced the quality and the presentation of the paper.

References

[1] Cisco Inc., "CISCO VNI Service Adoption Forecast 2013 2018 White Paper," Cisco, Tech. Rep., 2013-2018. [Online]. Available: http://goo.gl/X3efVF

[2] Infonetics, "Research for VoIP Services Forecast," [Online]. Available: http://goo.gl/rh21kQ

[3] M. Nassar, S. Niccolini, R. State, and T. Ewald, "Holistic VoIP Intrusion Detection and Prevention System," in 1st IPTCOMM, 2007.

[4] R. Zhang, X. Wang, X. Yang, and X. Jiang, "Billing Attacks on SIP-based VoIP Systems," In the 1st USENIX workshop on Offensive Technologies, 2007. [Online]. Available: http://dl.acm.org/citation.cfm? id=1323276.1323280

[5] S. Ehlert, D. Geneiatakis, and T. Magedanz, "Survey Of Network Security Systems To Counter SIP-Based Denial-Of-Service Attacks," Computers & Security, vol. 29, no. 2, pp. 225–243, 2010.

[6] A. Keromytis, "A Comprehensive Survey of Voice over IP Security Research," IEEE Communications Surveys Tutorials, vol. PP, no. 99, pp. 1 –24, 2011.

[7] Spam Phone Calls Cost U.S. Small Businesses Half-Billion Dollars in Lost Productivity, Marchex Study Finds. [Online]. Available: http://goo.gl/jTrgp3

[8] C. K. JENNIFER, "Complaints about Automated Calls up Sharply (last retrieved August 2015)." [Online]. Available: http://goo.gl/H5HTBh

[9] The state of phone fraud 2014-2015 a global, cross-industry. [Online]. Available: https://www.pindrop.com/phone-fraud-report/

[10] Y. Rebahi, D. Sisalem, and T. Magedanz, “SIP Spam Detection,” in ICDT ’06, 2006.

[11] Study of Mechanisms for Protection against Unsolicited Communication for IMS (PUCI), 3GPP Std. Release 11.

[12] H. Tu, A. Doupe,´ Z. Zhao, and G. Ahn, “SoK: Everyone Hates Robocalls: A Survey of Techniques against Telephone Spam,” in 37th IEEE Symposium on Security and Privacy, 2016.

[13] P. Kolan and R. Dantu, “Socio-Technical Defense Against Voice Spam-ming,” ACM Transactions on Autonomous and Adaptive Systems, vol. 2, 2007.

[14] R. Dantu and P. Kolan, “Detecting Spam in VoIP Networks,” in The Steps to Reducing Unwanted Traffic on the Internet, 2005.

[15] Y. Wu, S. Bagchi, N. Singh, and R. Wita, “Spam Detection in Voice-Over-IP Calls through Semi-Supervised Clustering,” in 39th Annual IEEE/IFIP DSN, 2009, pp. 307–316.

[16] V. Balasubramaniyan, M. Ahamad, and H. Park, “CallRank: Combating SPIT Using Call Duration, Social Networks and Global Reputation,” in 4th CEAS, 2007.

[17] (2015) Symantec Intelligence Report (Reterived September 2015). [Online]. Available: https://goo.gl/4C7xu5

[18] M. Uemura and T. Tabata, “Design and Evaluation of a Bayesian-filter-based Image Spam Filtering Method,” in ISA 2008, 2008.

[19] J. Suykens and J. Vandewalle, “Least Squares Support Vector Machine Classifiers,” Neural Processing Letters, vol. 9, pp. 293–300, 1999. [Online]. Available: http://dx.doi.org/10.1023/A%3A1018628609742

[20] P. Kolari, A. Java, T. Finin, T. Oates, and A. Joshi, “Detecting Spam Blogs: A Machine Learning Approach,” in 21st National Conference on Artificial Intelligence - Volume 2, 2006, pp. 1351–1356. [Online]. Available: http://dl.acm.org/citation.cfm?id=1597348.1597403

[21] A. Ntoulas, M. Najork, M. Manasse, and D. Fetterly, “Detecting Spam Web Pages Through Content Analysis,” in 15th International WWW, 2006, pp. 83–92. [Online]. Available: http://doi.acm.org/10. 1145/1135777.1135794

[22] I. Santos, I. Marcos, C. Laorden, P. Garc´ıa, A. Ibirika, and P. Bringas, “Twitter Content-Based Spam Filtering,” in International Joint Conference SOCO13-CISIS13-ICEUTE13. Springer, 2014, pp. 449– 458. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-01854-6 46

[23] S. Iranmanesh, H. Sengar, and H. Wang, “A Voice Spam Filter to Clean Subscribers Mailbox,” in Security and Privacy in Communication Networks. Springer, 2013, pp. 349–367. [Online]. Available: http: //dx.doi.org/10.1007/978-3-642-36883-7 21

[24] J. Strobl, B. Mainka, G. Grutzek, and H. Knospe, “An Efficient Search Method for the Content-based Identification of Telephone-SPAM,” in 2012 IEEE ICC, 2012, pp. 2623–2627.

[25] D. Lentzen, G. Grutzek, H. Knospe, and C. Porschmann, “Content-Based Detection and Prevention of Spam over IP Telephony - System Design, Prototype and First Results,” in 2011 IEEE ICC, 2011, pp. 1–5.

[26] C. Porschmann¨ and H. Knospe, “Analysis of Spectral Parameters of Audio Signals for the Identification of Spam Over IP Telephony,” in 5th Conference on Email and Anti-Spam CEAS, 2008.

[27] R. Schlegel, S. Niccolini, S. Tartarelli, and M. Brunner, “SPam over In-ternet Telephony (SPIT) Prevention Framework,” in IEEE GLOBECOM, 2006, pp. 1–6.

[28] K. Srivastava and H. Schulzrinne, “Preventing Spam for SIP-based Instant Messages and Sessions,” Columbia University Technical Report CUCS-042-04, Tech. Rep., October 2004.

[29] J. Lindqvist and M. Komu, “Cure for Spam Over Internet Telephony,” in 4th IEEE CCNC, 2007.

[30] J. Quittek, S. Niccolini, S. Tartarelli, and R. Schlegel, “On Spam over Internet Telephony (SPIT) Prevention,” in IEEE Communications Magazine, 2008, pp. 80–86.

[31] J. Quittek, S. Niccolini, S. Tartarelli, M. Stiemerling, M. Brunner, and T. Ewald, “Detecting SPIT Calls by Checking Human Communication Patterns,” in IEEE ICC, 2007.

[32] Y. Soupionis and D. Gritzalis, “Audio CAPTCHA: Existing Solutions Assessment and a New Implementation for VoIP Telephony,” Computer & Security, vol. 29, pp. 603–618, 2010. [Online]. Available: http://dx.doi.org/10.1016/j.cose.2009.12.003

[33] Y. Soupionis, R. Koutsiamanis, P. Efraimidis, and D. Gritzalis, “A Game-theoretic Analysis of Preventing Spam over Internet Telephony via Audio CAPTCHA-based Authentication,” Journal Computer Security, vol. 22, pp. 383–413, 2014. [Online]. Available: http://dl.acm.org/citation.cfm?id=2597910.2597912

[34] N. Croft and M. Olivie, “A Model for Spam Prevention in IP Telephony Networks Using Anonymous Verifying Authorities,” in ISSA 2005, 2005.

[35] M. Hansen, M. Hansen, J. Moller,¨ T. Rohwer, C. Tolkmit, and H. Waack, "Developing a Legally Compliant Reachability Management System as a Countermeasure against SPIT," in 3rd Annual VoIP Security Workshop, 2006.

[36] D. Shin, J. Ahn, and C. Shim, "Progressive Multi Gray-Leveling: a Voice Spam Protection Algorithm," in IEEE Network, 2006, pp. 18–24.

[37] K. Ono and H. Schulzrinne, "Have I met you before? Using Cross-Media Relations to Reduce SPIT," in 3rd IPTCOMM, 2009, pp. 1–7.

[38] J. Rosenberg and C. Jennings. (2008) The session initiation protocol (sip) and spam (rfc 5039). IETF.

[39] D. Gritzalis, P. Katsaros, S. Basagiannis, and Y. Soupionis, "Formal Analysis for Robust anti-SPIT Protection using Model Checking," International Journal of Information Security, vol. 11, pp. 121– 135, 2012. [Online]. Available: http://dblp.uni-trier.de/db/journals/ijisec/ ijisec11.html#GritzalisKBS12

[40] N. D'Heureuse, J. Seedorf, and S. Niccolini, "A Policy Framework for Personalized and Role-based SPIT Prevention," in 3rd IPTComm, 2009, pp. 1–11. [Online]. Available: http://doi.acm.org/10.1145/1595637. 1595653

[41] H. Schulzrinne, H. Tschofenig, and J. Morris. Common Policy: A Document Format for Expressing Privacy Preferences (RFC 4745).

[42] Y. Soupionis, S. Dritsas, and D. Gritzalis, "An Adaptive Policy-Based Approach to SPIT Management," in ESORICS Computer Security. Springer, 2008, pp. 446–460. [Online]. Available: http: //dx.doi.org/10.1007/978-3-540-88313-5 29

[43] H. Tschofenig, R. Falk, J. Peterson, J. Hodges, D. Sicker, and J. Polk, "Using SAML to Protect the Session Initiation Protocol (SIP)," Network Management of Global Internet Networking, vol. 20, pp. 14–17, 2006. [Online]. Available: http://dx.doi.org/10.1109/MNET.2006.1705878

[44] ITU Survey On Anti-spam Legislation Worldwide, ITU Std., July 2015.

[45] Criminal and J. U. England. Sending unsolicited ext messages. [Online]. Available: http://goo.gl/8vwZTJ

[46] Canada's anti-spam Legislation (CASL). [Online]. Available: http://goo.gl/8vwZTJ

[47] D. Gritzalis and Y. Mallios, "A sip-oriented SPIT management framework," Computers & Security, vol. 27, pp. 136 – 153, 2008. [Online]. Available: http://www.sciencedirect.com/science/article/ pii/S0167404808000333

[48] M. Azad and R. Morla, "Multistage SPIT Detection in Transit VoIP," in 19th IEEE SoftCOM, 2011, pp. 1–9.

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

[50] K. Toyoda and I. Sasase, "SPIT Callers Detection with Unsupervised Random Forests Classifier," in IEEE ICC, 2013, pp. 2068–2072.

[51] H. Hong, K. Sripanidkulchai, H. Zhang, Z. Shae, and D. Saha, "Incorpo-rating Active Fingerprinting into SPIT Prevention Systems," in VSW06, 2006.

[52] P. Gupta, B. Srinivasan, V. Balasubramaniyan, and M. Ahamad, "Phoneypot: Data-driven Understanding of Telephony Threats," in 20th NDSS, 2015.

[53] K. Lee, J. Caverlee, and S. Webb, "Uncovering Social Spammers: Social Honeypots + Machine Learning," in 33rd ACM SIGIR Conference on Research and Development in Information Retrieval, 2010, pp. 435–442. [Online]. Available: http://doi.acm.org/10.1145/1835449.1835522

[54] M. Balduzzi, P. Gupta, L. Gu, Gao.D, and M. Ahamad, "MobiPot: Understanding Mobile Telephony Threats with Honeycards," in In Proceedings of the 11th ACM ASIACCS, 2016.

[55] Y. Rebahi and D. Sisalem, "SIP Service Providers and the Spam Problem," in Voice over IP Secuity Workshop, 2005.

[56] R. Zhang and A. Gurtov, "Collaborative Reputation-based Voice Spam Filtering," in DEXA '09., 2009, pp. 33–37.

[57] N. Chaisamran, T. Okuda, G. Blanc, and S. Yamaguchi, "Trust-Based VoIP Spam Detection Based on Call Duration and Human Relation-ships," IEEE/IPSJ International Symposium on Applications and the Internet, vol. 0, pp. 451–456, 2011.

[58] H. Bokharaei, A. Sahraei, Y. Ganjali, R. Keralapura, and A. Nucci, "You can SPIT, but you can't hide: Spammer Identification in Telephony Networks," in IEEE INFOCOM, 2011, pp. 41–45.

[59] S. Phithakkitnukoon and R. Dantu, "Defense against SPIT using Com-munity Signals," in IEEE International Conference on Intelligence and Security Informatics, 2009, pp. 232–232.

[60] P. Patankar, N. Gunwoo, G. Kesidis, and C. Das, "Exploring Anti-Spam Models in Large Scale VoIP Systems," in 28th International Conference on Distributed Computing Systems, 2008, pp. 85–92.

[61] J. Seedorf, N. D'Heureuse, S. Niccolini, and M. Cornolti, "Detecting Trustworthy Real-Time Communications Using a Web-of-Trust," in IEEE GLOBECOM, 2009, pp. 1–8.

[62] H. Guang, W. Ying, and Z. Hong, "SPIT Detection and Prevention Method Based on Signal Analysis," in 3rd International Conference on Convergence and Hybrid Information Technology, 2008, pp. 631–638.

[63] Study of Mechanisms for Protection against Unsolicited Communication for IMS (PUCI), 3GPP Std.

[64] A. Schmidt, A. Leicher, Y. Shah, I. Cha, and L. Guccione, "Sender Scorecards for the Prevention of Unsolicited Communication," in 2nd IEEE Workshop onCollaborative Security Technologies, 2010, pp. 1–6.

[65] C. Sorge and J. Seedorf, "A Provider-Level Reputation System for Assessing the Quality of SPIT Mitigation Algorithms," in IEEE ICC, 2009, pp. 1–6.

[66] Y. 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, 2009. [Online]. Available: http://dx.doi.org/10.1007/s10207-008-0071-0

[67] F. Wang, Y. Mo, and B. Huang, "P2P-AVS: P2P Based Cooperative VoIP Spam Filtering," in IEEE WCNC, 2007, pp. 3547–3552.

[68] K. Li, Z. Zhong, and L. Ramaswamy, "Privacy-Aware Collaborative Spam Filtering," IEEE Transactions on Parallel and Distributed Systems, vol. 20, pp. 725–739, 2009.

[69] T. Yao, S. Pin-Chieh, and C. Ming-Syan, "Cosdes: A Collaborative Spam Detection System with a Novel E-Mail Abstraction Scheme," IEEE Transactions on Knowledge and Data Engineering, vol. 23, pp. 669– 682, 2011.

[70] N. Foukia, L. Zhou, and C. Neuman, "Multilateral Decisions for Collaborative Defense Against Unsolicited Bulk e-MailMul," in 4th International Conference on Trust Management, 2006, pp. 77–92. [Online]. Available: http://dx.doi.org/10.1007/11755593 7

[71] A. Ramachandran, N. Feamster, and S. Vempala, "Filtering Spam with Behavioral Blacklisting," in 14th ACM CCS, 2007, pp. 342–351. [Online]. Available: http://doi.acm.org/10.1145/1315245.1315288

[72] E. Damiani, S. Vimercati, S. Paraboschi, and P. Samarati, "P2P-based Collaborative Spam Detection and Filtering," in the 4th International Con-ference on Peer-to-Peer Computing,, 2004, pp. 176–183.

[73] G. Singaraju and B. ByungHoon-Kang, "RepuScore: Collaborative Reputation Management Framework for Email Infrastructure," in 21st Conference on Large Installation System Administration Conference, 2007, pp. 1–9. [Online]. Available: http://dl.acm.org/citation.cfm?id= 1349426.1349445

[74] M. Sirivianos, K. Kyungbaek, and X. Yang, "SocialFilter: Introducing Social Trust to Collaborative Spam Mitigation," in IEEE INFOCOM, 2011, pp. 2300–2308.

[75] CloudMark. (2016) CloudMark SpamNet. [Online]. Available: https: //www.cloudmark.com/en

[76] J. Kong, B. Rezaei, N. Sarshar, V. Roychowdhury, and P. Boykin, "Collaborative Spam Filtering Using E-Mail Networks," Computer, vol. 39, pp. 67–73, 2006.

[77] P.-A. Chirita, J. Diederich, and W. Nejdl, "MailRank: Using Ranking for Spam Detection," in 14th ACM CIKM, 2005, pp. 373–380. [Online]. Available: http://doi.acm.org/10.1145/1099554.1099671

[78] D. Wang, D. Irani, and C. Pu, "A Social-spam Detection Framework," in 8th CEAS, 2011, pp. 46–54. [Online]. Available: http://doi.acm.org/10.1145/2030376.2030382

[79] R. Cardoso, "Mobile Application for Blocking Spam Callers," Master's thesis, Faculty of Engineering, University of Porto, 2015

[80] M. A. Akbar and M. Farooq, "Securing SIP-based VoIP Infrastructure against Flooding Attacks and Spam over IP Telephony," Journal of Knowledge and information systems, Springer, Vol. 38, No. 2, 2014, pp. 491–510.

[81] M. Keyworth "Vishing and Smishing: The Rise of Social Engineering Fraud" http://www.bbc.com/news/business-35201188

[82] GeekTime.Com, "Millions of victims lost $12.7B last year falling for Nigerian scams", http://www.geektime.com/2014/07/21/millions-of-victims-lost-12-7b-last-year-falling-for-nigerian-scams/

[83] FBI, "$2.3 Billion Lost to CEO Email Scams", https://krebsonsecurity.com/2016/04/fbi-2-3-billion-lost-to-ceo-email-scams/

[84] Scam statistics, https://www.scamwatch.gov.au/about-scamwatch/scam-statistics

[85] David Glance "Phone scams cost billions. Why isn't technology being used to stop them?", https://phys.org/news/2017-06-scams-billions-isnt-technology.html

[86] A.B. Shahroudi, R.H. Khosravi, H.R. Mashhadi and M. Ghorbanian, “Full Survey on SPIT and Prediction of How VoIP Providers Compete in Presence of SPITTERS Using Game Theory”, in: Proc. of the 2011 IEEE International Conference on Computer Applications and Industrial Electronics (ICCAIE), Abu Dhabi, 2011, pp. 402–406.

[87] F.M. Hossein, S.M.Amanian and F. H. Khosravi,” A Survey of Different SPIT Mitigation Methods and a Presentation of a Comprehensive SPIT Detection Framework”, In International Journal of Machine Learning and Computing (IJMLC) 2014 Vol.4(2).

[88] S. Kamas and M.A Aydan, “SPIT Detection and Prevention,” In IU-JEEE Vol. 17(1), (2017).

[89] Soupionis, Y., Gritzalis, D., “ASPF: an Adaptive anti-SPIT Policybased Framework,” In: Pernul G., et al. (ed.) Proceedings of the 6th International Conference on Availability, Reliability and Security (ARES-2011), pp. 153–160, Austria (2011)

[90] Consumers Union, “Robocalls Keep Coming!”, http://consumersunion.org/end-robocalls/problems/

[91] A. D. Keromytis, "A Survey of Voice over IP Security Research", *Information Systems Security. Springer*, pp. 1-17, 2009.

[92] S. Phithakkitnukoon, R. Dantu, and E.-A. Baatarjav, “VoIP Security Attacks and Solutions,” Information Security Journal: A Global Perspective, vol. 17, no. 3, pp. 114–123, 2008.

[93] S. Dritsas, Y. Soupionis, M. Theoharidou, Y. Mallios, and D. Gritzalis, “SPIT Identification Criteria Implementation: Effectiveness and Lessons Learned,” in Proceedings of 23rd International Information Security Conference. Springer, 2008, pp. 381–395.

[94] R. Dantu, S. Fahmy, H. Schulzrinne, and J. Cangussu, “Issues and Challenges in Securing VoIP,” Computers & Security, vol. 28, no. 8, pp. 743–753, 2009.

[95] G. F. Marias, S. Dritsas, M. Theoharidou, J. Mallios, and D. Gritzalis, “SIP Vulnerabilities and anti-SPIT Mechanisms Assessment,” in Computer Communications and Networks, 2007. ICCCN 2007.

[96] S. F. Khan, M. Portmann, and N. W. Bergmann, “A Review of Methods for Preventing Spam in IP Telephony,” Modern Applied Science, vol. 7, no. 7, p. p48, 2013.

[97] M.A. Azad and S. Bag, “Decentralized Privacy-aware Collaborative Filtering of Smart Spammers in a Telecommunication Network”, *Proceedings of the Symposium on Applied Computing* (SAC '17), 2017.

[98] A. Markkola, J. Lindqvist, "Accessible Voice CAPTCHAs for Internet Telephony", *Proceedings of the Symposium on Accessible Privacy and Security (SOAPS)*, 2008.

[99] US Federal Trade Commission (FTC), “Blocking Unwanted Calls”, https://www.consumer.ftc.gov/articles/0548-blocking-unwanted-calls