Identification of Spam over Internet Telephony through Acoustic Fingerprint Technique

Nitikesh S. Thakare
ME(CSE)2nd semester G.H Raisoni College of Engineering, Amravati.

Avinash P.Wadhe
M-Tech(CSE) G.H Raisoni College of Engineering, Amravati.

Abstract
Spam has posed a serious problem for users of email since its infancy. Today, automated strategies are required to deal with the massive amount of spam traffic. In an email setting, we have a significant chance to quarantine incoming emails to check if they are spam or a virus. In fact, as Simple Mail Transfer Protocol (SMTP) allows a repeated negotiation step between the sender and receiver, there are plenty of chances to tap into the negotiation steps and quarantine emails. Research in this area includes various approaches — such as Bayesian-based, content-based, DNS based or signature-based and sometimes combines methods for a hybrid approach. As opposed to email protocols represented by SMTP, VoIP does not tolerate any negotiation in the signaling step or screening of content because of its slim delay limitation. Thus, it is harder to implement any Spam over Internet Telephony (SPIT) protection or control algorithms in a VoIP setting. SPIT detection and mitigation can also be based on the caller's audio data and that's why in order to identify such technique is required. Thus we approaches, uses of the audio identification techniques (similar to music identification) to detect calls with identical audio data including certain degradations (e.g., noise and different audio codecs) called as robust Acoustic fingerprint is derived from spectral parameters of the audio data and replayed calls are identified by a comparison of fingerprints for dealing with spam and its voice means precisely.

Index Terms
SIP, VoIP, Spam, SPIT.

I. INTRODUCTION

NOWADAYS, spammers use forged messages, stolen identities, bogus cancellation addresses, and relay hijacking to hide their identities when sending their advertisements or bogus messages[2]. This activity uses a lot of Internet bandwidth. We can categorize spam issues from both the client and server perspectives in this process further. VoIP spam is also known as SPIT, or Spam over Internet Telephony. As the popularity of VoIP increases for consumers and businesses, so do the same un-wanted messages we see every day over traditional phone and email networks.
As a service provider, especially in the mobile arena, the expectation is that you will do all you can to protect your valued subscribers from SPIT wherever possible. VoIP spam or SPIT (Spam over Internet Telephony) are bulk unsolicited, automatically dialed, pre-recorded phone calls using the Voice over Internet Protocol (VoIP). Telephone spam is comparable to E-mail spam, but due to its synchronous character, different mitigation methods are needed. Voice over IP systems, like e-mail and other Internet applications, are susceptible to abuse by malicious parties who initiate unsolicited and unwanted communications. Telemarketers, prank callers, and other telephone system abusers are likely to target VoIP systems increasingly, particularly if VoIP supplants conventional telephony. The VoIP technology provides convenient tools (e.g. Asterisk and SIPp) and low-priced possibilities to place a large number of Spam calls. The underlying technology driving this threat is Session Initiation Protocol (SIP)[10]. This technology has received significant support from most major telecommunication vendors, and is showing signs of becoming the industry standard for voice, video and other interactive forms of communication such as instant messaging and gaming[16]. To avoid VoIP users to manually label nuisance calls, our voice spam filter automatically marks all VoIP calls with short call durations as unsolicited. The preliminary simulation results show that our approach is effective to counter voice spam. Like human fingerprints, Audio fingerprints allow identifying an audio file among a set of candidates but does not allow retrieving any other characteristics of the files.

II. Related Work

When SMTP came into existence, no one thought about it needing a security mechanism. Thus, SMTP is a simple, text-based protocol that supports
only a basic mechanism to avoid spam — that is, the receivers can check whether or not the sender's system meets the RFC standards. So, spammers and criminals have misused the SMTP system by flooding it with lots of spam. There is little information available on implementations of SPIT mitigation measures by Telephone companies. SPIT is generally not yet considered to be problem with similar relevance as E-

Manuscript received May 5, 2014
Manuscript revised May 20, 2014

mail spam. An automated analysis of the call signaling flow can help to discover SPIT[5].Relevant parameters and indications of SPIT are, for example, a high call attempt frequency, concurrent calls, low call completion and low call duration average. An Audio fingerprint is a small digest of an audio file computed from its main perceptual propertiesIn the spam protection aspect, voice is more difficult to protect than email, due to its synchronicity. Since email is delivered asynchronously, users and administrators have a lot of chances to check and quarantine them. One well-known email spam protection algorithm, Bayesian Spam Filtering[1] , is able to analyze the content of emails and detect spam mails at a certain checkpoint. Because email is delivered asynchronously, it does not really matter if email delivery is delayed for a short period of time. However, in the case of SPIT, we have to decide if the incoming call is spam or not within the connection time. Once the connection is set up, it is too late to take action because the voice spam is already disrupting the call server and recording voice mail.

III. Proposed system

We propose in this paper a new fingerprint extraction methodology which combines a segmentation method with a new fingerprint construction scheme. The proposed method is robust against compression and time shifting alterations of the audio files. The focus of this paper is set on the topic of so called SPAM over Internet Telephony although SPIT contains the phrase "SPAM" and has some parallels with email spam, it also has major differences. The similarity is that in both cases senders (or callers) use the Internet to target recipients (or callees) or a group of users, in order to place bulk unsolicited calls [15]. The main difference is that an email arrives at the server before it is accessed by the user. This means that structure and content of an email can be analysed at the server before it arrives at the recipient and so SPAM can be detected before it disturbs the recipient[9]. As in VoIP scenarios delays of call establishment are not wished, session establishment messages are forwarded immediately to the recipients. Besides this fact the content of a VoIP call is exchanged not until the session is already established[9]. In other words if the phone rings it is too late for SPIT prevention and the phone rings immediately after session initiation, while an email can be delayed and even, if it is not delayed, the recipient can decide if he wants to read the email immediately or not. In addition to these aspects another main difference between spam and SPIT is the fact, that the single email itself contains information that can be used for spam detection. The header fields contain information about sender, subject and content of the message. A single SPIT call in contradiction is technically indistinguishable from a call in general. A SPIT call is initiated and answered with the same set of SIP messages as any other call.

IV. Methodology and Approaches

A. Method 1

In content-based methods, the information contained in the mail body or header (such as the
subject) is compared to different rules to automatically classify it as spam or legitimate. This
approach uses different classifiers to label a message as spam, including rule- and Bayesian based classifiers. This is the main reason that, in some classifications, Bayesian-based approaches are also listed under this category. Bayesian-based is the most popular technique because it's easy to implement[6]. The algorithm's native version is the naive Bayes probabilistic, statistical classifier. In the learning stage, messages are classified as spam according to the probability of the frequency of a particular term used in them. Later, this data is used as a criterion to detect spam.

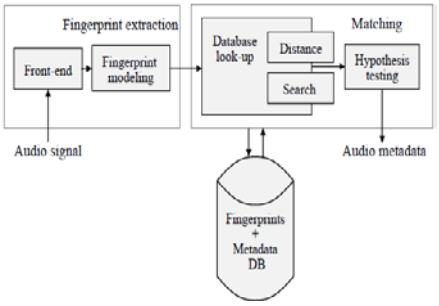

Figure 1. RGB Color Values [12]

B. Method 2

Perceptual characteristics often exploited by audio fingerprints include average zero crossing rate, estimated tempo, average spectrum, spectral flatness, prominent tones across a set of bands, and bandwidth. Most audio compression techniques (AAC, MP3, WMA, Vorbis) will make radical changes to the binary encoding of an audio file, without radically affecting the way it is perceived by the human ear[3]. A robust acoustic fingerprint will allow a recording to be identified after it has gone through such compression, even if the audio quality has been reduced significantly. For use in radio broadcast monitoring, acoustic fingerprints should also be insensitive to analog transmission artifacts. On the other hand, a good acoustic fingerprint algorithm must be able to identify a particular master recording among all the productions of an artist or group. For use as evidence

in a court of law, an acoustic fingerprint method must be forensic in its accuracy. Fingerprint hashing, hash functions allow comparison of two large objects, X and Y, by just comparing their respective hash values H(X) and H(Y)[4]. For a properly designed fingerprint function F, there should be a threshold T such that, If X and Y are similar, then ||F(X) – F(Y)|| < T with very high probability, And ||F(X) – F(Y)|| > T if X and Y are dissimilar.

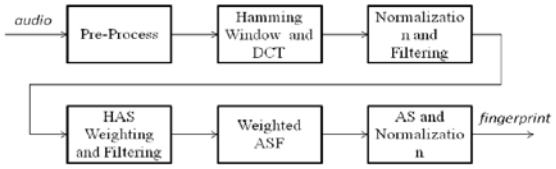

Figure2.the framework of the feature extraction [3]

This step contains pre-process, framing, time frequency transformation and data filtering. A stereo waveform should be converted into a mono waveform in the pre-process phase because this proposed algorithm is aimed to the mono waveforms. In order to extract robust feature from the dynamic audio data, a framing window should be applied to the audio waveform to obtain relatively static audio clips. After the front-process, we begin to use the weighted ASF descriptor to compute the audio feature. To obtain a robust feature, we should get the most sensitive part of the frequency spectrum[3].

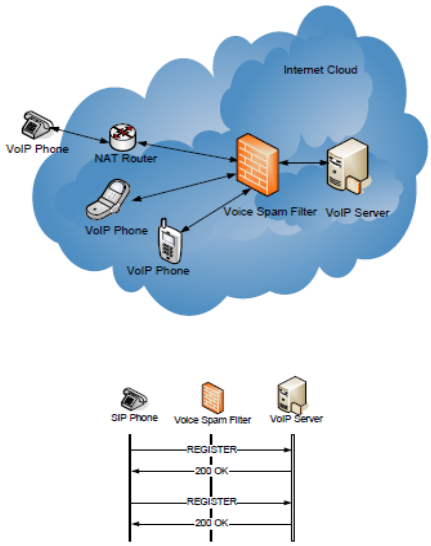

Figure.3. Transmitting SIP REGISTER Messages Through the Voice Spam Filter [14]

In our experiment, we select the frequency range in 250 Hz-2000 Hz to extract the audio feature. Then the frequency spectrum of each frame is partitioned into bands in a logarithmic spacing and these bands are not overlapped.

V. Hypothetical experimental study

Let us denote by Ti the set of intervals used to compute the subfngerprints of a signal si. Using the method of Kalker and Haitsma [12], this set is equal to the number of sliding windows used by this algorithm. Using our method this set is equal to the number of If intervals defined by the segmentation step. Some additional quantities may be usefully defined to measure the performances of our method: Given an audio file si, let us denote by SPi _ Ti the set of intervals located at a same position within si and a degraded version of si. We consider, that two intervals have a same position if the distance between the center of the two intervals is less than 0.25ms. When the degraded version of si is shifted, the position of the interval within the shifted version is translated by the shift before computing the distance. Moveover, let us additionally consider the set SVi-SPi of intervals having a same location and a same subfingerprint value within si and its degraded version. Given a specific degradation, several quantities may be defined in order to measure the performances of our algorithm:

1. Segmentation rate: This quantity represents the mean value of If intervals located at a same position within the original signal and its degraded version. This quantity is thus a measure of the performances of our segmentation algorithm. It is formally defined by :

SR = 1 N N Xi=1 |SPi| |Ti| (1)

where |.| denote the cardinal of the set and N the number of audio files of the database.

2. Recognition rate: The robustness of our method used to compute subfingerprints is measured for each si by the ratio between the cardinals of SVi and SPi. We measure thus the ratio of intervals whose subfingerprint value remains unchanged by a degradation. The mean value of this ratio over the whole database defines the recognition rate and is formally defined by:

RR = 1 N N Xi=1 |SVi| |SPi| (2)

3.Total recognition rate : The recognition rate defined above, measures the robutness of our subfingerprints independently of the segmentation step. A global measure of both steps may be achieved by computing for each signal si the ratio between SVi and Ti. This measure may be understood, as the product, for each si, of |SPi| |Ti| and |SVi| |SPi| respectively used to define the segmentation and the recognition rates. The mean value of this ratio over the database defines the total recognition rate formally defined as :

TRR = 1 N N Xi=1 |SVi| |Ti| (3)

We also measured the performance of Kalker and Haitsma [5] algorithm. The segmentation rate and the recognition rate are meaningless for this method since it does not perform a segmentation step. We thus only use

the total recognition rate to measure the performances of this algorithm. Further experiments using the bit rate error between files may be found in [?] Note that the quantities defined in this section measure the robustness of our fingerprints against degradation. These quantities don't readily allow to measure the efficiency of an indexing scheme which does not constitute the core of this paper (section 5). Indeed, the sets SPi and SVi are usually not known when a request on the fingerprint database is performed using the fingerprint of an unknown signal[12].

VI. conclusion

The reference model described earlier can be applied in a wide range of scenarios from SPIT prevention at a high-throughput peering point between VoIP operators to SPIT prevention at a VoIP terminal. However, concrete instances of the system for different scenarios may vary significantly. Also, the selection of applied prevention methods strongly depends on application requirements and constraints. An acoustic fingerprint is a condensed digital summary, deterministically generated from an audio signal, that can be used to identify an audio sample or quickly locate similar items in an audio database[2]. Practical uses of acoustic fingerprinting include identifying songs, melodies, tunes, or advertisements; sound effect library management; and video file identification.. This identification has been used in copyright compliance, licensing, and other monetization schemes. A robust acoustic fingerprint algorithm must take into account the perceptual characteristics of the audio. If two files sound alike to the human ear, their acoustic fingerprints should match, even if their binary representations are quite different. Acoustic fingerprints are not bitwise fingerprints, which must be sensitive to any small changes in the data. Acoustic fingerprints are more analogous to human fingerprints where small variations that are insignificant to the features the fingerprint uses are tolerated[6]. One can imagine the case of a smeared human fingerprint impression which can accurately be matched to another fingerprint sample in a reference database; acoustic fingerprints should work in a similar way to deal with SPIT.

References

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

[2] H. Rafiee,M. Löwis,C. Meinel." IPv6 Deployment and Spam Challenges ".Internet Computing, IEEE Volume: 16 Issue: 6 Page(s): 22 – 29, November 2012.

[3] J. Chen, Tiejun Hang ."A Robust Feature Extraction Algorithm for Audio Fingerprinting ". IEEE Transaction on speech and processings, 2002.

[4] J.Burges, J.Platt , S. Jana. "Distortion Discriminant Analysis for Audio Fingerprinting". IEEE Transaction on speech and processings, Vol.11, No.3, May, 2003.

[5] A. Schmidt, N. Kuntze1, R. Khayari." Spam Over Internet Telephony and How to Deal with it" Security & Privacy, IEEE Volume: 8 , , Page(s): 45 – 50,2010 .

[6] J. Yagnik, D. Strelow, D. Ross, R. Lin, "The Power of Comparative Reasoning," in Proceedings of ICCV, 2011.

[7] J. Haitsma, "A Highly Robust Audio Fingerprinting System," in Proceedings of ISMIR, 2002.

[8] C. Sorge, S. Niccolini, J. Seedorf. "The Legal Ramifications of Call-Filtering Solutions ".Security & Privacy,IEEE,Volume: 8 , Page(s): 45 – 50,2010 .

[9] Keromytis, A.D. "Voice-over-IP Security: Research and Practice ". Security & Privacy, IEEE , 2010 .

[10] B. Mathieu,S. Niccolini,D. Sisalem. "SDRS: A Voice-over-IP Spam Detection and Reaction System ". Security & Privacy, IEEE .Volume: 6 , Issue: 6 , Page(s): 52 – 59,2008 .

[11] J. Quittek, S.Niccolini, S. Tartarelli, R. Schlegel." On Spam over Internet Telephony (SPIT) Prevention ". Communications Magazine, IEEE Volume: 46 , Issue: 8 , Page(s): 80 – 86,2008.

[12] P. Cano , E. Batlle, T.Kalker, J. Haitsma." A Review of Algorithms for Audio Fingerprinting".

[13] K. Li , Z.Zhong , L. Ramaswamy." Privacy-Aware Collaborative Spam Filtering ".Parallel and Distributed Systems, IEEE Transactions on Volume: 20 , Issue: 5 , Page(s): 725 – 739,2009 .

[14] R. Zhang, A. Gurtov." Collaborative Reputation-based Voice Spam Filtering". IEEE Transaction on speech and processings, 2004.

[15] Wikipedia,http://en.wikipedia.org/wiki/Ascouticfingerprint,June 2011.

[16] Wikipedia,http://en.wikipedia.org/wiki/VOIPSpam, April 2011.

Mr. Nitikesh S. Thakare Received the B.E and from SGBAU Amravati university and he is currently pursuing ME(CSE) from G.H Raisoni College of Engineering, Amravati. His research interest include Data mining system .He had presented research paper at conference in AIMS , Bangalore in 2013.

Avinash P. Wadhe Received the B.E and from SGBAU Amravati university and M-Tech (CSE) From G.H Raisoni College of Engineering, Nagpur (an Autonomous Institute). He is Currently an Assistant Professor with the G.H Raisoni College of Engineering and Management, Amravati SGBAU Amravati university. His research interest include Network Security , Data mining and Fuzzy system .He has contributed to more than 20 research paper. He had awarded with young investigator award in international conference.