PREVENTING SPIT WITH NAIVE BAYES IN VOIP COMMUNICATION

Intesab Hussain Sadhayo*, Fareed Ahmed Jokhio**, Umair Ali Khan*, Pardeep Kumar*, Nisar Ahmed Memon*

ABSTRACT

Spams over Internet telephony is a serious threat to consumers and enterprises. In comparison to e-mail spams, voice spams have drastic effects. It consumes time and resources of the victim for example telemarketing. To counter this problem, we propose a two-step solution with analysis and detection to such threats. In first step, we extract the useful data from VoIP calls to organize and to get the input for our detection phase using bloom filter. In second phase, we examine the data obtained during the first step using Naive Bayes (NB). The efficiency of NB is analyzed by simulations. The results show that our proposed technique for voice spam delivers a high level of accuracy.

Keywords: *VoIP, SIP Network, security, SPIT, DoS*

1. INTRODUCTION

Voice over Internet Protocol (VoIP) is a rapidly growing Internet service. Due to its remarkable flexibility, feasible implementation, and low-cost connectivity to international telephones, it has become extremely popular among end-users and communication engineers [1-4]. However, in spite of its immense benefit, this technology contains certain drawbacks also [5] [6]. Session Initiation Protocol (SIP) [7], which is one of the widely used protocol for VoIP services, is vulnerable to many types of attacks. Apart from this, there are also different types of attacks targeting VoIP protocols [8] [9], such as DoS attacks, call hijacking, toll fraud, SPam over Internet Telephony (SPIT) and phishing. SPIT is one of the most serious threats to IP telephony.

Internet telephony is vulnerable to spam calls due to its low cost. Voice over IP systems, like e-mail and other Internet applications, are susceptible to abuse by malicious parties to initiate unsolicited and unwanted communications for telemarketing. Apparently, attackers generate a number of machine automated calls to launch a DoS attack. SPIT is similar to the email spaming problem which many Internet users face quite often. Although email spam will still be a big challenge in the future. The numerous solutions [10-15] proposed over the last few years have helped to mitigate the problem significantly.

Many of the well-known techniques which are used for email spam detection fail completely in the context of VoIP due to many reasons. First, an email usually arrives at a server before it is finally downloaded by the user. Such a mail server can apply many filtering strategies. For instance, it can check whether the text body of the email mentions certain products. In contrast, recognizing human voices and to determine whether the message is spam or not is still a very difficult task. Additionally, the recipient of a call only learns about the subject of the message when someone is actually listening to it. Also from a user's perspective, SPIT is quite different from e-mail spam. Although a spam email is nuisance, it is typically easy to delete such an email. But it can be very un-productive if a regular email from a friend is considered spam and not delivered to a user's inbox. Having said that, it may be tolerable if an email spam filter yields a large ratio of false negatives, but the filter should avoid false positives completely [16].

The work presented in this paper is related to the analysis and detection of threats encountered in Internet telephony. We record a large number of VoIP calls and extract certain input features for spam analysis using bloom filter. The results obtained from the bloom filter are further used as inputs to a Naïve Bayes (NB) classifier for spam detection.

The rest of the paper is organized as follows. Section 2 discusses the related work in the domain of spam detection. Section 3 describes our proposed framework for SPIT detection. Section 4 gives a brief overview of the data analysis through bloom filter. Section 5 describes the spam detection using Naïve Bayes classifier. Section 6 discusses the performance evaluation of the proposed technique. Section 7 concludes the paper.

* Quaid-e-Awam University of Engineering, Science & Technology, Nawabshah, Pakistan.

** Åbo Akademi University, Finland

***email*: {**intesab, fajokhio, umair.khan, pardeep.kumar, drnisar}@quest.edu.pk

2. RELATED WORK

The relevant literature about spam detection demonstrates various techniques. In [17], SPIT detection is done through human telephony communication patterns. Turing test is used to differentiate between human and computer botnets. The two human communication patterns are discussed which are double-talk and call-start pattern. In double-talk, the fraction of time in which both caller and callee are in talking mode is checked. In start pattern, the proposed scheme checks the starting pattern by asking questions having very short answers. If the callee replies with a long answer for the question which has short answer expected, the Turing test will declare it a SPIT. The limitation of this technique is that it can detect spams in only machine generated calls.

The authors in [18] propose a local centric approach based on signaling protocol analysis to counter SPIT. For detection purpose, they consider few facts, e.g., unidirectional spammers, validity of signaling routing data, termination of calls by the same conversation parties, and spammer does not call the same destination for some fraction of time. This solution detects external spammers on a recipient side. The counters are maintained for call setup and call termination in different time stamps. The simultaneous deviation of these counters indicate the spammers' activity with respect to certain probabilities. However, false positives may lead this method to consider legitimate callers as SPIT.

In [19], authors propose an AntiSPIT method based on a blacklist. This module takes input from Call Detail Record (CDR) and decides whether to put the caller in a black list or not. Although this method is simple, it could block legitimate callers too.

The technique proposed in [20] uses a decoy to catch the spammers without their knowledge. In fact, this technique puts a decoy system in front of a server. The spammers treat the decoy as a server and their status are stored in the decoy. This solution is implemented in SIP Express Router (SER). The method is suitable for the IP Multimedia System (IMS) with strong authentication to avoid spoofing. However, this scheme performs well in the situation where two or more decoys are hit by a spammer.

The authors in [21] compare different types of Naive Bayes to get the best choice of Naive Bayes for e-mail spams. Many machine learning algorithms are used for e-mail spam filtering, e.g., support vector machines, boosting and Naïve Bayes classifier. The authors present a comparison between different Naive Bayes classifiers including multi-variate Bernoulli Naive Bayes, multinomial Naive Bayes, TF attribute, boolean attribute, Multi-variate Gauss Naive Bayes, and flexible Bayes. According to this work, the best results are obtained by the flexible Bayes and multinomial Naive Bayes with boolean attributes.

The initial platform to use Naive Bayes in intrusion detection systems for VoIP threats is proposed in [22]. The detection process in this technique keeps track of request intensity, error response, number of destinations, SIP methods and response distribution of VoIP. However, this work can be further extended to detect SPIT.

In [23], the consequences of SPIT attacks are highlighted. The first consequence described is *space consumption* due to voice messages. Second, employees' disturbance in companies and the last one is a user's service complaint. In [24-27], different solutions are proposed for some specific scenarios.

3. PROPOSED SOLUTION FOR SPIT DETECTION

We consider the properties of various VoIP calls to detect SPIT. The high-level overview of our scheme is depicted in Figure 1.

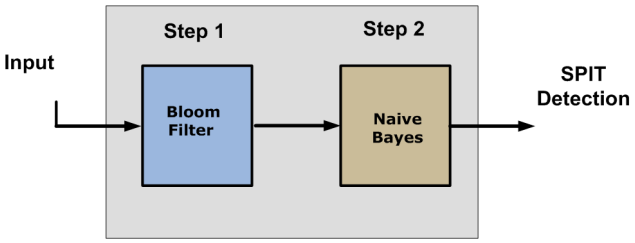

Fig.1: Two-step scheme for SPIT detection

We calculate a call-completion-ratio by comparing the number of call initiation, successful acknowledgements and termination. The length of a call is measured by analyzing the call duration. Furthermore, nature of the calls is checked to confirm whether they are machine generated or (genuine) human calls. Our approach also detects the abnormally large number of calls (machine recorded or human) which exhaust the server resources and result in a Denial of Service (DoS) at victim's side.

4. ANALYSIS THROUGH BLOOM FILTERS

We use bloom filter in the first step of our proposed technique as described in Figure 2 and Figure 3. In this step, we need to gather the information regarding the callers. This step serves as an input to step 2 in which we use Naive Bayes to predict the probability of calls being SPIT. The bloom filter is a space-efficient technique used for probabilistic dataset for testing whether the element is a member of set or not. An empty bloom filter is a bit array of m bits, all set to 0. There must also be k different hash functions defined, each of which maps or hashes some set element to one of the m array positions with a uniform distribution.

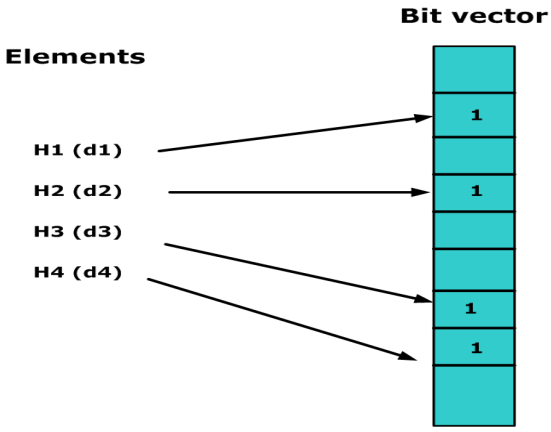

Fig. 2: Generic structure of Bloom filter

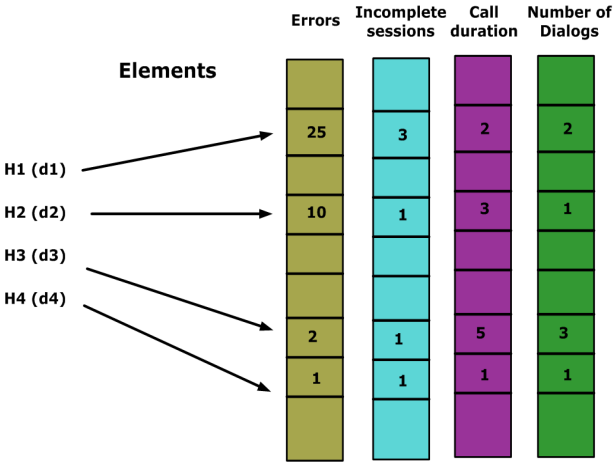

Fig. 3: Bloom filter modified for our proposed technique

To add an element in the bloom filter's array, we feed it to each of the *k* hash functions to get *k* array positions. We set the bits at all these positions to 1. Bloom filters have a strong space advantage over other data structures for representing sets, such as self-balancing binary search trees, hash tables, simple arrays or linked lists.

Table 1 shows the statistics of the calls analysis with respect to their behavior. We count the number of requests and number of responses through bloom filter. This helps us to check the incomplete sessions of calls. We also count the probability of waiting calls, ongoing traffic sessions and number of maximum dialogues. These parameters help us to specify the respective probabilities to detect the SPIT calls. The bloom filter is used for this data collection.

We calculate the probabilities of different frequently used SIP methods. These probabilities are shown in Table 2. This table shows the methods which are necessary for call completion, call cancellation, acknowledgement and registration. These probabilities help us to find the SPIT calls used as input in the second step of our proposed technique. Table 3 shows the probabilities of error response type. In SIP responses, the error could be client error, server error, global error, redirection, successful or provisional response. This step is based on the traffic analysis, and later we use this data to the next step of our proposed scheme.

TABLE 1: SIP CALL DATA

	Number of Requests	Number of Responses	Calls Waiting	RTP Open ports	Maximum no: of Dialogues
0 -10	1	---	---	0.8	1
> 10	0	---	---	0.2	0
0 - 4	---	0.2	---	---	---
> 4	---	0.8	---	---	---
0 - 7	---	---	0	---	---
> 7	---	---	1	---	---

TABLE 2 : SIP METHODS FOR SPIT

INVITE	REGISTER	ACK	CANCEL	BYE
0.40	0.00	0.40	0.00	0.20

TABLE 3 : SIP ERROR RESPONSES

1xx	2xx	3xx	4xx	5xx	6xx
0.30	0.00	0.05	0.20	0.20	0.05

5. NAIVE BAYES CLASSIFIER

A Naive Bayes classifier is a simple probabilistic classifier based on Bayes theorem [28] with strong independence assumptions. Naive Bayes classifier assumes that the presence of a particular feature of a class is unrelated to the presence of any other feature, given the class variable. The success of Naive Bayes relies on the dependencies. Even if these features depend on each other or on the existence of other features, a Naive Bayes

classifier considers all of these properties to independently contribute to the probability of an item. Equation 1 shows the probabilistic model of Naïve Bayes classifier.

$$p(H \mid e) = \frac{p(e \mid H)p(H)}{P(e)} \quad (1)$$

In equation 1, $p(H \mid e)$ represents the probability of instance e being in class H (in our case this represents the probability of a call being legitimate or malicious), $p(e \mid H)$ is the probability of generating instance e given class H, $p(H)$ is the probability of occurrence of class H, and $p(e)$ is the probability of occurrence of instance e.

Depending on the precise nature of the probability model, a Naive Bayes classifier can be trained very efficiently by a supervised learning algorithm. In many practical applications, parameter estimation for Naive Bayes models uses the method of maximum likelihood. In spite of its simple design and apparently over-simplified assumptions, Naive Bayes classifier performs quite well in many complex real-world situations. Some analysis of the Bayesian classification [28] shows that there are some theoretical reasons for the apparently unreasonable efficacy of Naive Bayes classifiers. However, a comprehensive comparison with other classification methods [29] shows that Bayes classification is outperformed by more current approaches, such as boosted trees or random forests. However, the main advantage of the Naive Bayes classifier is that it only requires a small amount of training data (means and variances of the variables) to estimate the parameters necessary for classification. Since independent variables are assumed, only the variances of the variables for each class needs to be determined and not the entire co-variance matrix.The training is quite easy to implement and just requires the (prior) conditional properties calculated from the offline data analysis.

6. PERFORMANCE EVALUATION

We use Weka [30], a suit of machine learning software written in Java, for the performance evaluation of our proposed technique. It is a data mining tool used for data pre-processing, classification, regression, clustering, association rules, and visualization.

We generate random data to check the records of VoIP calls to analyze SPIT. The gathered data includes total number of requests, responses, waiting calls, etc. We further use bloom filter to classify the SIP call data. We generate margin curves which show the difference between the probability calculated for the actual class and the probabilities calculated for the other classes. Margin curves also show the threshold curves of YES/NO (legitimate/malicious) class. The points illustrating prediction tradeoffs can be obtained by varying the threshold values between classes such as cost curves and cost-benefit analysis of YES and NO. This is particularly useful for the analysis of predictive analytic outcomes. The Naive Bayes classifier is applied to the data retrieved from the bloom filter. Figure 4 depicts margin curve of Naive Bayes. Figure 5 and Figure 6 show threshold curve of YES and threshold curve of NO, respectively. After comparing training data and actual data, Naive Bayes calculates the probability of the calls being SPIT based on the given information and predicts the future probabilities. Thus Naive Bayes classifier is able to detect the SPIT calls through the calculated probability based on the given SIP data. Figure 7 and Figures 8 depict cost-benefit analysis of YES and NO. Figures 9 shows cost analysis of YES and Figure 10 shows cost analysis of NO.

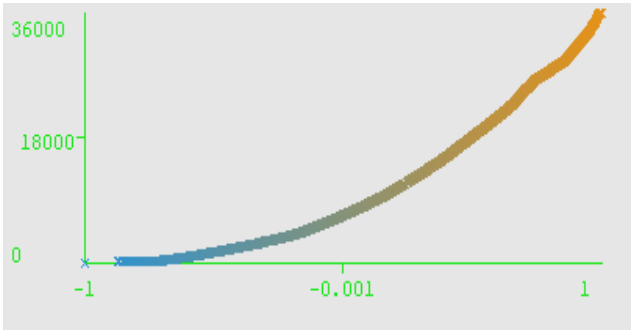

Fig. 4 : Naive Bayes Margin curves

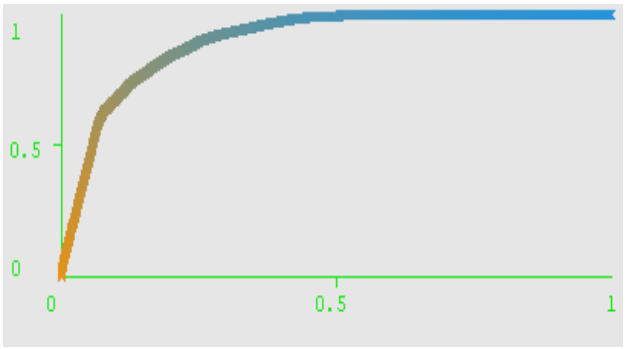

Fig.5: Naive Bayes threshold curve of YES

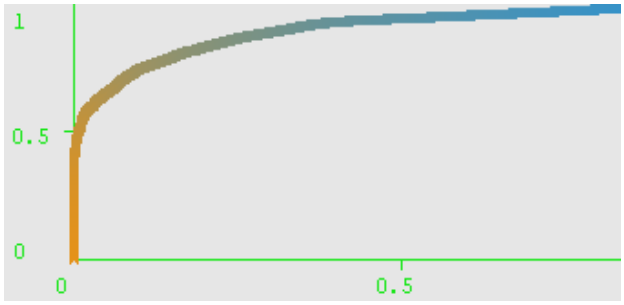

Fig. 6: Naive Bayes threshold curve of NO

Fig. 7: Naive Bayes cost-benefit analysis of YES

Fig. 8: Naive Bayes cost-benefit analysis of NO

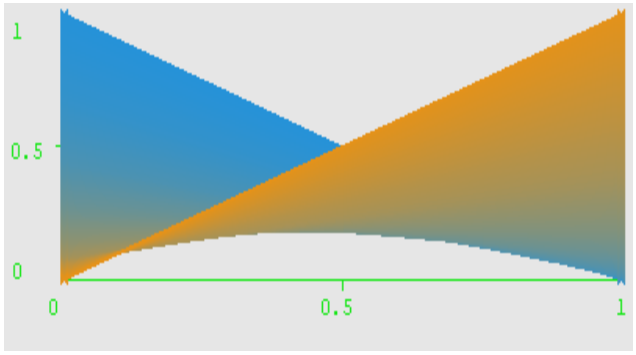

Fig. 9 : Naive Bayes cost analysis of Yes

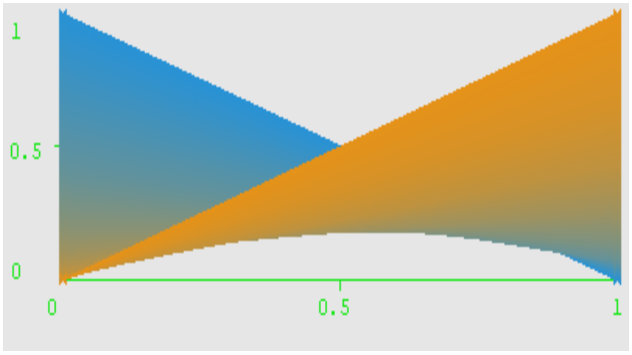

Fig. 10: Naive Bayes cost analysis of No

The proposed technique is easy to configure and does not require a special architecture or an additional hardware on the servers. It is cheaper than other proposed techniques in the sense that most of the existing solutions require Intruder Detection Systems (IDS), firewalls and detection devices at both server and client side. The proposed solution is memory-efficient due to the usage of Bloom filter. Future prediction through Naive Bayes works well to block SPIT. Few solutions have been proposed [31-35] for this problem, but they are less efficient than our proposed scheme as shown in our performance evaluation.

8. CONCLUSION

SPIT is a growing threat to SIP based VoIP systems keeping in view that the email spam filtering is out of context in VoIP [35-39]. SPIT are harmful for the organizations and for the individuals too. In this paper, we have proposed a two-step solution for spam detection. The first step arranges the SIP data through a bloom filter, while the second step classifies the data to be legitimate or spam using a Naive Bayes classifier. We calculate the probability of the SIP calls being SPIT through a Naive Bayes classifier. At the same time, Naive Bayes helps to predict the future probability of a call being SPIT. The margin curve, cost curve and cost-benefit analysis verify the efficiency of Naive Bayes in our proposed technique. In future, we aim to compare different types of Naïve Bayes classifiers to further improve the accuracy of our proposed technique.

REFERENCES

[1] http://www.iptel.org/ser

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

[3] http://www.kamailio.org/w/

[4] http://www.trainsignal.com/blog/voip-signaling-protocols

[5] D. Kuhn, T. Walsh, S. Fries, “Security considerations of voice over IP Systems”, National Institute of Standards and Technology (NIST), Gaithersburs, MD, USA, Computer Security Division, Special Publication 800-58, January 2005.

[6] Ge Zhang, "Unwanted Traffic and Information Disclosure in VoIP Networks Threats and Count-ermeasures", DISSERTATION, Karlstad University Faculty of Economic Sciences, Communication and IT Computer Science 2012.

[7] H. Schulzrinne et al., "RTP: A Transport Protocol for Real-Time Applications", *RFC 1889*, January 1996.

[8] M. Handley et al.," SDP: Session Description Protocol", *RFC 2327* , April 1998.

[9] J. Rosenburg et al., "Sip: Session Initiation Protocol", *RFC 3261*, June 2002

[10] T. Dierks et al., "The Transport Layer Security (TLS) Protocol", *R*FC 5246, August 2008.

[11] D. Geneiatakis N. Vrakas, C. Lambrinoudakis, "Utilizing bloom filters for detecting flooding attacks against SIP based services", Computers Security,Vol. 28, No. 7, pp. 578-591, 2009.

[12] T. Dierks, E. Rescorla, "The Transport Layer Security (TLS) protocol version 1.1", *R*FC 4346, April 2006.

[13] J. Franks et al., “HTTP Authentication: Basic and Digest Access Authentication”, Internet Engineering Task Force, RFC 2617, June 1999.

[14] S. Salsano et al., “SIP security issues: the SIP authentication procedure and its processing load”, IEEE Trans. on Network, Vol. 16, No. 6, pp. 38-44, 2002.

[15] B. Ramsdell, “Secure/Multipurpose Internet Mail Extensions (S/MIME) Version 3.1 Message Specification”, IETF RFC 3851, July 2004.

[16] S. Ehlert, D. Geneiatakis, T. Magendaz, "Survey of Network Security Systems to Counter SIP-based Denial-of-Service Attacks ", Trans. on Computer and Security, Vol. 29, No. 2, pp. 225-243, 2010.

[17] D. Xianglin, M. Shore "Advanced flooding attack on a SIP server", In Proc. of IEEE International Conference on Availability, Reliability and Security (ARES), pp. 647-651, 2009.

[18] “http:www.math.uiuc.edu/r-ash/BPT.html", Computer Communications, Volume 31 , Issue 10, June 2008.

[19] M. Bsihop, " Introduction to computer security", Pearson Education India, 2005.

[20] J. Loughney and G. Camarillo, “Authentication, Authorization, and Accounting Requirements for the Session Initiation Protocol (SIP)”, RFC 3702, Feb. 2004.

[21] Keromytis, Angelos D. "A Comprehensive Survey of Voice over IP Security Research.", IEEE Trans. on Communications Surveys and Tutorials, Vol. 14, No. 2, pp. 514-537, 2011.

[22] S. Ehlert, D. Geneiatakis, T. Magendaz, "Survey of Network Security Systems to Counter SIP-based Denial-of-Service Attacks ", Trans. on Computer and Security, Vol. 29, No. 2, pp. 225-243, 2010.

[23] D. Geneiatakis, A. Dagiouklas, G. Kambourakis, C. Lambrinoudakis, S. Gritzalis, S. Ehlert, D. Sisalem, "Survey of Security Vulnerabilities in Session Initiation Protocol", IEEE Trans. on Communications Surveys and Tutorials, Vol. 8, No. 3, pp. 68-81, 2006.

[24] M.A. Akbar, Z. Tariq, M. Farooq, "A comparative study of Detection Algorithms for Detection of SIP Flooding in IMS", In Proc. of IEEEConference on Internet Multimedia Services Architecture and Applications(IMSAA), pp.1-6, 2008.

[25] D. Geneiatakis T. Daiuklas, G. Kambourakis et.al, "Survey of Security vulnerabilities in Session Initiation Protocol", IEEE Trans. on Communications surveys and tutorials,Vol. 8, No. 1-4, pp. 68-81, 2006.

[26] D. Sisalem, J. Kuthan, "Denial of service attacks and SIP infrastructure", IEEE Trans. on Network, Vol. 20, No. 5, pp. 26-31, 2006.

[27] I. Husaain, F. Nait-Abdesselam, "Strategy based proxy to secure user agent from flooding attack in SIP", In Proc. of 7th International Wireless Communications and Mobile Computing Conference (IWCMC), pp. 430-435, 2011.

[28] Albrecht, Keno, Nicolas Burri, and Roger Wattenhofer. "Spamato-an extendable spam filter system.", In Proc. of 2nd Conference on Email and Anti-Spam (CEAS), California, USA. 2005.

[29] MacIntosh R, Vinokurov D., “Detection and mitigation of SPAM in IP telephony networks using signalling protocol analysis”, In Proc. of IEEE symposium on advances in wired and wireless communication, pp. 49-52, 2005.

[30] www.Weka.com

[31] J. Quittek, S. Niccolini, S. Tartarelli, M. Stiemerling, M. Brunner and T. Ewald, “Detecting SPIT calls by checking human communication patterns”, In Proc. of IEEE International Conference on Communications (ICC07), pp. 1979-1984, 2007.

[32] Salehin, SM Akramus, and Neco Ventura. "Blocking unsolicited voice calls using decoys for the IMS." In Proc. of IEEE International Conference on Communications, ICC’07, 2007.

[33] V. Metsis, I. Androutsopoulos, G. Paliouras, “Spam filtering with Naïve Bayes - which Naives Bayes?”, In Proc. of the 3rd Conference on Email and Anti-Spam, Mountain View, pp. 27-28, 2006.

[34] Nassar et al., “Intrusion detection mechanisms for VoIP applications”, 2006.

[35] "http://sipsak.org/"

[36] Nucci, Antonio et al., "SIP-based VoIP traffic behavior profiling", United States Patent 7441429, 2008.

[37] H. Sinnreich, A. Johnston., “Internet communications using SIP: delivering VoIP and multimedia services with session initiation protocol”, Vol. 27, Wiley, ISBN: 978-0-471-41399-8, 2001.

[38] Wang, Haiyan, Runsheng Zhou, and Yi Wang, "An anti-spam filtering system based on the Naive Bayesian Classifier and Distributed Checksum Clearinghouse", In Proc. of IEEE Third International Symposium on Intelligent Information Technology Application (IITA), pp.128-131, 2009

[39] M. Voznak and F.Rezc, “VoIP SPAM and a defence against this type of threat”, In Proc. Of 14th International Conference on Communications, pp. 172-177, 2010.