SPIT DETECTION AND PREVENTION

Selin KAMAS[1], Muhammed Ali AYDIN [2]

[1]Netaş Telecommunication, Istanbul, Turkey
[2]Department of Computer Engineering Istanbul University, Istanbul, Turkey
skamas@netas.com.tr, aydinali@istanbul.edu.tr

Abstract: *In telecommunication technology VoIP protocol has become a very popular technology as it is cheap, efficient. Also it has easy deployment. While it has lots of advantages it brings lots of vulnerabilities. These are Man in the middle Attack, Replay Attack, Teardown Attacks, Flooding Attacks, Toll Fraud and SPIT (Spam over IP Telephony). Spam over IP Telephony (SPIT) is an known threat in the Voice over IP Networks (VoIP). Even though evolved from email spam, SPIT is more obstructive and intrusive in nature. SPIT attack is called important threat of reliability and availability of VoIP system and also it is difficult to make SPIT call in PSTN (Public Switched Telephone Network) system. In this work It is tried to say how SPIT attacks occur, how attackers do it and also it is mentioned that prevention mechanisms and compare them in terms of feasibility, advantages and disadvantages..*

Keywords: *VoIP, VoIP security, SPAM over IP Telephony ,SPIT,Captcha, Whitelist, Blacklist.*

1. Introduction

VoIP spam is unwanted and automatic calls that are consecutive records have been recorded previously. VoIP system has much vulnerability because of its IP Infrastructure. One of them is SPIT. In early 2004 found 50% of e-mail is determined to be spam call. E-mail and old phone system protocol's addressing system is similar with VoIP. Therefore VoIP system is vulnerable spam call, too. In VoIP systems, Spamming reveals more effective results than can be done in e-mails protocols. Because spam calls obstruct people to use phone. Additionally, VoIP systems are cheaper than PSTN system. Therefore this makes VoIP system an easier target for telemarketers [1]. Telemarketers are people who make unwanted phone calls to sell products or services.

VoIP protocols have a lot of tools (SIPp, Asterisk) that are used by attackers to make spam call[2]. The other reason VoIP system vulnerable to SPIT is SIP (Session Initiation Protocol)'s vulnerabilities. The Session Initiation Protocol (SIP) is a communications protocol for signaling and controlling multimedia communication sessions. SIP characterizes the messages that are sent between endpoints, which govern establishment, termination and other essential elements of a call. The protocol can be used for creating, modifying and terminating sessions consisting of one or several media streams. SIP has information about voice, codec, application type and status of call [3]. "Spammers" starts session and If they use SIP, used request message type is "INVITE" to start session. After called answer call, they send automated voice record (SIPp, Asterisk) to spam called. These calls are unwanted, irrelevant, unsolicited and unexpected and called SPIT. To make SPIT call is ordinary and its result is effectively dangerous. Therefore, to make secure VoIP system, providers should apply variety of methods to detect and prevent SPIT call. To mention SPIT attack's visibility and its effective dangerous result Softbank in Japanese reported they have seen three big SPIT attack in their VoIP system [15]. Voice over IP (VoIP) is a methodology and group of technologies for the distributing of voice communications and multimedia sessions over Internet Protocol (IP) networks, such as the Internet. To prevent SPIT call; it is not idea to change its IP infrastructure. Using IP protocol has common usage and not opens the change [5].

2. Compare SPIT and SPAM

SPAM e-mails do not disturb users or system until they open their e-mails. After they open e-mails they can understand this is spam mail and they can point this mail's sender as a spammer. E-mail protocol can block this sender to prevent send spam mail after this point. Also spam mail can detect before user open mail by checking content of mail. But SPIT call cannot prevent like that scenario. Because, SPIT calls disturb user when users open call. And until users open the call system cannot understand it is SPIT because content of

Received on: 10.06.2016
Accepted on: 08.02.2017

voice communication cannot be seen. And this prevents user access to service.

Also VoIP systems and e-mail protocols have different in terms of time. VoIP systems work in real time.

Because of e-mail service is content-based, spam mails can be detect by checking content but in VoIP system cannot check content of communication until conversation starts. Moreover, filtering cannot be made by looking content of conversation.

The following table shows the system in PSTN or VoIP systems is that when compared to the costs of spamming.

Table 1. Comparison of Costs of Spam Attacks on PSTN and VoIP systems [8]

Cost	SPAM (PSTN)	SPIT (VoIP)	Note
Software Cost.	A.	A.	**A** (change according to signaling protocol).
Hardware Cost.	10B-100B.	B.	**B** (does not change according to signaling protocol).
Cost of every spam.	About 1000C.	C.	**C** (does not change according to signaling protocol).

3. Spamming Over Internet Telephony (SPIT)

Attackers make phone unwanted and unexpected continuous call to inhibit users to access services or to advertise or discredit providers[6]. VoIP system vulnerable this attack the same reason with e mail services. This reason is every person can call every person really cheaply.

As just mentioned, Telemarketers also benefit from SIP addressing that like email addressing. Telemarketers use several web pages, e-mail lists or crawling technic to takeover SIP addresses. Also to seizure SIP addresses or usernames attackers make Brute Force or Dictionary Attack.

While attackers make SPIT , for example If there is and 30sn packet to send, attackers use RTP (Real Time Protocol) and it takes 30sn to deliver this packet and a system security administrator can think this feature can be used to prevent SPIT. But it is not idea because telemarketers use parallelism to handle this condition [5].

Voicemail services are services that facilitate the feasibility of SPIT attacks. Telemarketer send SPIT call even offline user thanks to previously recorded messages.

SIP provide user to be anonym.. SIP enables this capability through e-mail services, unlike VoIP protocol circuit-switched system, resulting in the vulnerability of a spamming attack. SPITTers create a botnet for themselves and hide their IP addresses. SPITTers are the same people with telemarketers. A botnet (also known as a zombie army) is a number of Internet computers that, although their owners are unaware of it, have been set up to forward transmissions (including spam or viruses) to other computers on the Internet.

In one experiment, without any SPIT attack network usage is 21kbps and for 30sn needed space is 75KB and in 30sn 100 spam e-mail can be send in experimental network. From this point 100.000 voice record and every one lasts 30sn. Therefore it shows that this SPIT calls needed 7.2GB uplink network usage capacity. From there SPIT calls can be detected [5].

List of requirement of architecture to prevent SPIT

- ✓ Do not block legal users
- ✓ Maximize possibility of detect attacker who make SPIT calls
- ✓ Stop communication with attacker and victim called
- ✓ Prevent the SPITTers to define themselves as legal
- ✓ Be used as appropriate for different language, infrastructure, environment (office, home) [8].

To prevent SPIT there are a lot of methods but none of them have all of requirement which are mentioned above. The methods should exclude called users while detect or prevent SPIT calls. Based on this assumption; get feedback from a caller will be way more intelligent solution. From this point there is an algorithm which defines black and grey list. While prevention mechanism make classification, it check caller from inter-domain. Secondly, mechanism have waited proof from caller about his*her honesty about call goal’s. This proof can be done with Computational Puzzles, sender checks, Turing test etc. But with this kind of mechanism problem is that: caller has to proof his/her honesty and this cause users wait long time until call establish. Moreover, computational puzzles and Turing test’s complexity is not effective to implement real time application even though their complexity is median [8].

3.1 Solution Methods

In e-mail services if users can manage their e-mail individually, defining black and white list approaches to prevent spam mail in level of proxy and client can be feasible. Because users should able to edit their mail according to type of mail lists (spam, social, advertisement, all etc.).However service providers should able to filter e-mail in their servers. Also defining black and white list merely is not enough to prevent spam mails because of ability of create botnet or IP spoofing attack.

Another method is CAPTCHA (Completely Automated Public Turing to Tell Computers and Humans Apart). In many places, end users expected from brute-force authentication mechanism used to verify user to prevent attacks on the proxy. This verification mechanism's random code instantaneous transmission can be produce in Proxy (on the fly)'s process or using sound recordings produced by the user [8]. But expecting user to strive in this protection mechanism is not very accurate. In addition, the reliability of the records carried out on-the-fly process should also be discussed.

Transmission of these records must be secured using some encryption methods. This requires a distinct performance.

Other method is non-reputation. From historical call details can be found caller and called information. But SPITTers have found ways to overcome this. They have agreed with peer and pretending as a legal user. After that they start to SPIT call and handle non-reputation mechanism. Non-reputation does not require any effort to SPITTers is a deterrent method.

To define White list is somehow limits SPITTers. But defining black list is not efficient method to limit or stop SPITTers to make SPIT call as it is mentioned before. Reason of this is ability to create botnet or dynamic IP addresses etc. [8].

Using CAPTCHA in web pages as a Picture or text is common way. Even the use of the Web page has security vulnerabilities. It should not be defined directly in the codes. Hash algorithms should be used during displaying of these numbers to users. This method cannot be used in e-mail services because e-mail services works asynchronous. It can be used in voice transmission but this prevention mechanism should be secure, too. Users should not have to expend extra effort for this mechanism [8].

Other prevention mechanism called Domain Based Authentication and Policy Enforced for SIP(DAPES) and it uses TLS(Transport Layer Security) and digest authentication mechanism[11]. Implementation of this method is infeasible and complicated. Because implementing this method require to change other modules.

In RFC5039; there are lots of methods to prevent SPIT. These are content filtering, black and white list, Consent-Based Communications, Reputation Systems, Address Obfuscation, Limited-Use Addresses, Turing Tests, Computational Puzzles, Payments at Risk, Legal Action, Circles of Trust. All of these have some disadvantages and because of this it is mentioned that every of methods has some vulnerabilities [12].

1. Content Filtering

Content filtering is a method which is used in e-mail services to prevent spam mails. However in VoIP system cannot be applied because voice communication is real time and no one check content until called answer to call. If content is saved as a voice record in voice mail, this method can be applied. In this case, to control this content, prevention mechanism should have sound/video recognition algorithm. But these algorithms can be broken by attackers and also these algorithms are complex and hard to implement. In addition in sound recognition system %40 of sound is noise [14].

2. Black List and White List

Black list is not a best practice for SPIT attack detection a prevention mechanism even though in e-mail services. Although SIP protocol makes inter-domain authentication, attackers can create limitless addresses and they do not care of being in black list.

White list and black list work oppositely. Attackers want to be white list. If SIP authentication mechanism work truly, SPIT calls will be detected. If there is a black list protection, users who are not in white list make call with effort. This is a restrictive method in terms of users comfort. Also expecting users to be in white in in their first call is not expected way. Moreover first SPIT calls will not be detected [17]. Also looking universal list and try to detect SPIT call is not pragmatic or feasible way to protecting from SPIT calls.

3. Grey List

Defining grey list; when users make their first call they will be in grey list and after a while system want to users to make call again. If user make call in this specific time line, he/she will be in white list else black list [17]. This is more feasible approach is based on previous black / white list identification.

D.Shin[17] define two grey list identification according to duration time of calls. There is a threshold for duration of call.

If duration of call longer then this threshold, call be marked as a SPIT call. Authentication mechanism of this method is weak. Therefore is not applicable [17].

4. Consent-Based Communications

This is hybrid solution of black/white list protection approaches. Users can accept calls directly or request for authentication. At first glance it seems applicable but its authentication mechanism is not sufficient for detect/prevent SPIT calls in VoIP systems [12]

5. Reputation Systems

This method is also hybrid soltion of black/white lists. For example, If A user is not white list for B user, non-reputation system helps B user to accept or reject this call. Non-Reputation system is used in more central messaging architecture. Non-reputation score calculates from user's feedback and according to this result system decide about call's intention. From this point, there will be same problems with black list. Because there will be generally positive feedbacks.

6. Address Obfuscation

SPITTers generally find e-mail addresses of SIP users from web pages or public places. In these circumstances, e-mail addresses should be hidden and should be too complex to be non-predicted.

Address Obfuscation is an approach for this situation. It advice that while saving e-mail addresses, these should be formatted differently as a non-predicted. For example, user@domain.com e-mail address should be saved as "user at domain com" or "jdrosenatexampledotcom".

However, under these conditions, after attackers notice pattern of format, can create a tree and it is possible to turn around [12].

7. Limited-Use Addresses

Limited-Use Addresses is about address obfuscation method. It limits number of user's e-mail addresses. For example, the number of e-mail addresses of specific users can be limited within specific time-line

After time-line user's access of e-mail address should be denied. If in this method, user's current e-mail address is used to make SPIT call, protection mechanism works and after that time this e-mail address cannot be used.

A disadvantage of this method is if user's e-mail address reaches maximum user has to notify other users who will be called from this email address. It is an expectation that this will not be welcome by users [12].

8. Payments at Risk

With this approach, for example If A user calls B user, Firstly A user needs to pay for a call to B user. If B user voted this call as a normal call, payment of this call repay to A user. Disadvantages of this method there is an need of transition payment two times and If A user do not have enough money to pay, even if A user is a normal user, he/she will not call B user.

9. Model-Based Filtering

It creates a model based on actual calls over. When the model was created, the frequency and duration of calls are compared with the previous calls. Also while creating a model user based call number, repetitive number of calls, time of calls, and number of unknown caller are saved. After calculation these numbers, decision of forwarding calls to called user depends on these results. This method is also not effective and performance. There can be high false positive rate with just these metrics [14].

10. Circles of Trust

With this approach users voted caller about call's intention. Trust score is applied about joining or rejecting from conversation. Calculating trust score is user-based because of this reason applying this method is proper small network or small providers. Applicability of this method in big networks is not enough to provide scalability.

11. Dendritic Cell Algorithm

DCA (Dendritic Cell Algorithm) is abstract model based classification method and it uses dendritic cell methodology in biology as a prevention mechanism.

Dendritic cells' task is antigen in biology [16]. DC provides to detect bacteria, viruses and other parasites in body. In term of performance and accuracy DC algorithm gives effective results and it is commonly used to detect and prevent for network security [17]. DCs are used as a key for detection and prevention algorithm for SPIT calls.

DCA works real time and is used to detect the anomaly on the data time-series. It is processing the signals and inform about the status of the network. DCA algorithm divide signal into four type [17]. These are;

- **PAMP Signal;** PAMP (Pathogenic Associated Molecular Pattern) signals generate micro-organism. Therefore, if there is PAMP Signal in network, it shows there is a high level anomaly in system.
- **Danger Signal;** This signal Show sudden death in biology. As a detection mechanism If there is an danger signal in the network If show anomaly but smaller than PAMP signal.
- **Safe Signal;** It shows that network is secure.

Antigen signs DC have changed somehow and DC's states can be three type. These are **immature**, **semi-mature** and **mature.**

It is shows transition within states of DC's in Figure 1.

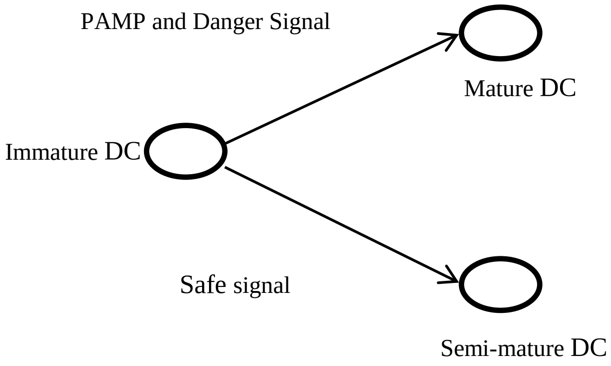

Figure 1: DC State Transition[17]

Inflammation: It shows immature DC have not passed mature state in biology, yet. In anomaly detection, the signals are effective in the formation of the other three signals.

If immature DC receives PAMP or danger signals, it switches to mature DC. Otherwise switches semi-mature.

DC algorithm gives three outputs. These are;

- **CSM Output Signal;** Costimulatory Molecule (CSM) signal show threshold of mature and immature DC. Before moving to the lymph node it is location of the incoming signal.
- **Semi-mature Signal;** It shows the cumulative sum of the safe signals.
- **Mature Signal;** It shows the cumulative sum of the PAMP and danger signals.

DC's states can be shown by 0(semi-mature) or 1(mature). DCA algorithm uses the time difference between the last and first call, daily call numbers to calculate **PAMP** signals, failed call numbers and time duration of calls are calculated for **danger** signal and lastly number of established/successful call numbers is calculated for **safe** signal. Using these numbers CSM and state of DC are determined [17].

After using DCA algorithm for detection and prevention SPIT calls, test results give %93.33 accuracy rates to SPIT calls and with % 96.67 accuracy rate, normal calls are classified. Even though DCA algorithm is complex and costly it can be implemented. And also DCA algorithm can be used to detect other anomaly of the system (flooding, DoS/TDoS, fuzzing, malformed SIP message etc.) not just for SPIT calls. In DCA algorithm to reduce false negative/false positive possibility, there can be defined more metric to calculate state of DC and CSM output. It will be increase accuracy of algorithm. For example, difference between normal and SPIT calls, ID of SPIT calls generally calls someone, not be called too much. ID of SPIT call's incoming call rate is much less.

If there is a big difference between incoming and outgoing call, can be sign PAMP signal. In addition, from historical data, SPITTers generally does not call same number again. However normal users generally call the same number more. The difference between repetitive and different call rate can be used PAMP and danger signal. It can reduce false negative/positive rates By using these metric additionally.

4. Conclusions

SPIT attacks are a threat for VoIP users and infrastructure. There are many method for detection and prevention and as a mentioned above also all of methods have advantages and disadvantages. Therefor there should be combined some of solution approaches. For example in DCA algorithm metric numbers should be increased and there can be used some machine learning algorithm to learn model of system. They can use support vector machines to classify calls [20]. Moreover they can use neural network algorithm and other machine learning algorithm for classification. Additionally they should define distinctive feature [+38] of SIP and use them as a feature vector. After classification traffic can be classify as a normal and bad. After obtained test data realizing SPIT call will be more accurate. And also with this approach users do not need to effort.

5. References

[1] J. Rosenberg "The Session Initiation Protocol (SIP) and Spam," draftietf-
sipping-spam-01.txt, July 2005. Work in progress.

[2] A. Baxter. Shtoom. http://divmod.org/
projects/shtoom, April 2006.

[3] Snoeren, Alex C., and Hari Balakrishnan. "An end-to-end approach to host mobility." Proceedings of the 6th annual international conference on Mobile computing and networking. ACM, 2000.

[4] L. von Ahn, M. Blum, N. J. Hopper, and J. Langford. The Captcha Project. http://www.captcha.
net/, April 2006.

[5] Bila, Nilton. "Dealing with SPAM in Voice over IP." change 34.4: 37.

[6] FELTEN, E. (2005, March 15). Unwanted Calls and Spam on VoIP. Retrieved from https://freedom-to-tinker.com/blog/felten/unwanted-calls-and-spam-voip/

[7] C. Garretson. Qovia ready to take on VoIP spam.
http://www.networkworld.com/news/
2004/071204qovia.html, July 2004.

[8] Prevention of Spam over IP Telephony (SPIT)Juergen QUITTEK, Saverio NICCOLINI, Sandra TARTARELLI, Roman SCHLEGEL

[9] Edelson, E.: Voice over IP: security pitfalls. Network Security, vol. 2005, pp. 4–7 (2005)

[10] Performance analysis of VoIP spoofing attacks using classification algorithms (2014 Application and Innovation in Mobile Computing)

[11] K. Srivastava, H. Schulzrinne, "Preventing Spam for SIP-based Instant Messages and Sessions", Technical report, Columbia, 2004.

[12] https://tools.ietf.org/html/rfc5039

[13] Secure Layered Architecture for Session Initiation Protocol based on SIPSSO. 2015 12th International Conference on Information Technology - New Generations

[14] A comprehensive SPIT Detection and Prevention framework based on Reputation Model on Call Communication Patterns. Farideh Barghi, Mohammad Hossein, Hossein Khosravi Roshtkhari.

[15] VOIPSA. Confirmed cases of SPIT. Mailing list (2006),http://www.voipsa.org/pipermail/voipsec-
voipsa.org/2006-March/001326.html

[16] The Dendritic Cell Algorithm Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy. Julie Greensmith October 2007

[17] Dendritic Cell Algorithm for preventing Spam over IP Telephony. V.Srihari, P.Kalpana , R.Anitha.2015

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

[19] Detection and Filtering Spam over Internet Telephony-A User-behavior-aware Intermediate-network-based Approach. Yan Bai1, Xiao Su1 and Bharat Bhargava3

[20] NASSAR, M., STATE, R., AND FESTOR, O. Monitoring SIP traffic using support vector machines. In RAID '08: Proceedings of the 11th international symposium on Recent Advances in Intrusion Detection (Berlin, Heidelberg, 2008), Springer-Verlag, pp. 311-330.
[21] On the inefficacy of Euclidean classifiers for detecting self-similar Session Initiation Protocol (SIP) messages. Anil Mehta*, Neda Hantehzadeh*, Vijay K. Gurbanit, Tin Kam Hot, Jun Koshiko* and Ramanarayanan Viswanathan*12th IFIP/IEEE International Symposium on Integrated Network Management 2011

Note:

Selin Kamaş was born in Tunceli. Selin is living in İstanbul/Turkey and is graduated from İzmir Instıtute of Technology. Department of her Computer Engineering(06.2015). Because of her interest of network security, she is working cyber security department of NETAŞ as a Cyber Security R&D Engineer. Her first article is published Toll Fraud Article XX.Internet Conference(2015).