A Phased Framework for Countering VoIP SPAM

Jongil Jeong[1], Taijin Lee[1], Seokung Yoon[1], Hyuncheol Jeong[1], Yoojae Won[1], Myuhngjoo Kim[2]

[1]IT Infrastructure Protection Division, Korea Information Security Agency
78, Garak-Dong, Songpa-Gu, Seoul, Korea, 138-803
{jijeong, tjlee, seokung, hcjung, yjwon}@kisa.or.kr

[2]Department of Computer Science and Engeering, Seoul Women's University
126, Gongreung-Dong, Nowon-Gu, Seoul, Korea 139-774
mjkim@swu.ac.kr

Abstract. VoIP spam will become severe problem preventing from generalization of VoIP service. For the purpose of presenting multi-leveled anti-spit framework, we divided VoIP service domain into three independent domains as an outbound, an intermediary, and an inbound domain. The proposed framework enables administrator to establish anti-spit policies in each domain. In outbound domain, the framework focuses on detecting and preventing spammers. The focus in intermediary domain is to block forged SIP message using sender policy framework. The framework enables victims to directly report spam contents they received to administrator. We showed that the multi-leveled anti-spit framework is enough to mitigate spam attacks.

Keywords: SPIT, SPIM, SPF, Anti-SPIT framework, VoIP SPAM

1 Introduction

According to the "*Hype Cycle for Consumer Technologies in 2007*" [1], the residential Voice over Internet Protocol (VoIP) service has already reached stage of "*Slope of Enlightenment*". This means residential VoIP service became the practical technology and its technical process can be accepted as the actual service model for achieving commercial business goals. The cycle expected the residential VoIP service will reach the next stage "*Plateau of productivity*" within 2 years. At this stage, the related technologies are commercialized and market also grows up based on its technical maturity. In advancing to the next stage, VoIP spam also will become a severe issue like email spam problems. VoIP spam may cause social problems increasing stress at home and in office, and deteriorate performance at work. Since these problems will bring down value of VoIP service, customers will hesitate to use VoIP service consequently; therefore, providing secure solutions is required for continuing growth of VoIP business.

This work was supported by the IT R&D program of MKE/IITA. [2006-S-043-03, The Development of VoIP Security Technology]

This paper divides VoIP service scope into three domains as an outbound, an intermediary, and an inbound domain respectively. We implemented three modules for countering spam attacks in each domain. In the outbound domain, we focus on detecting callers who have abnormal call pattern similar with that of spammers. In the intermediary domains, we focus on preventing from abusing intermediary domain name as the originator of forged SIP messages. In the inbound domain, we focus on reporting spam information immediately.

2 Background Study

Session Initiation Protocol (SIP) facilitates to create, delete, and modify multimedia sessions among devices that want to communicate with other devices over the Internet. Figure 1 shows the barebones architecture consisted of SIP components. User Agent Client (UAC) is a logical component generating SIP request messages. It begins SIP transactions. User Agent Server (UAS) is a logical component generating response messages corresponding to SIP request message requested by UAC. SIP proxy routes SIP messages between UAC and UAS. Registrar is a server to offer personal mobility of SIP. It gives UAC specific information about UAS's connection address [2].

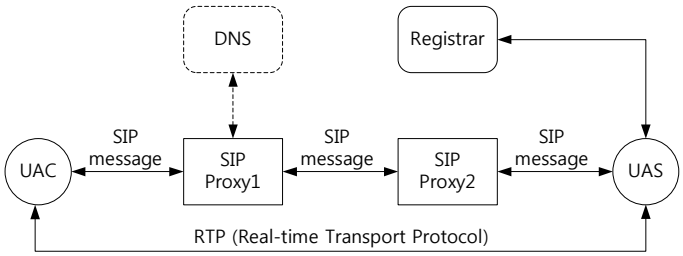

Figure 1. the barebones SIP architecture

2.1 Types of VoIP SPAM

In "*Threat Taxonomy*", which was published by VOIPSA (Voice Over IP Security Association) in 2005 [3], VoIP spam was classified into the "*unwanted lawful contents*". This includes that a seller solicits consumers to purchase lawful contents or goods for adults. Most users, however, strongly want to screen such solicitation. Types of VoIP spam can be assorted as follows. ***Call spam****:* a spammer attempts to send bulk unsolicited set of SIP messages in order to establish a multimedia session. ***IM spam****:* a spammer sends bulk unsolicited set of instant messages. Mainly this spam is sent via the extended SIP message for IM such as MESSAGE, but it is also sent via the "*Subject*" field of SIP Request message such as INVITE, OPTION, and SUBSCRIBE. ***Presence spam****:* using the "SUBSCRIBE" message of SIP, a spammer sends bulk unsolicited set of presence requests to become a member of the 'whitelist' or 'buddy list' of a user.

2.2 Current ANTI-SPIT Solutions

In order to present technical background, we explain current anti-spam solutions [4]. The following solutions were also deployed in our proposed anti-spit framework.

- *Blacklist*: a blacklist includes users that are considered as spammer. This list is usually used to block calls being initiated by the enlisted users.
- *Whitelist*: whitelist contains trusted users. Calls made by white-listed users are never blocked regardless of whether such users are registered into black list.
- *Content filtering*: These filters analyze the contents of messages, characterizing them as spam or not.
- *Challenge-response*: This checks whether the communication is established by a human or a bot. If the caller correctly answers a challenge sent by the callee, he is not a bot.
- *Reputation-based*: This approach is based on the notions of reputation and trust of the callers or the callers' domains. If each domain has the pre-defined threshold of trust level, the communication from a caller with good reputation is permitted, otherwise it is rejected.

3 Implementation of A Phased ANTI-SPIT Framework

Figure 2 shows the anti-spit solutions implemented in each domain.

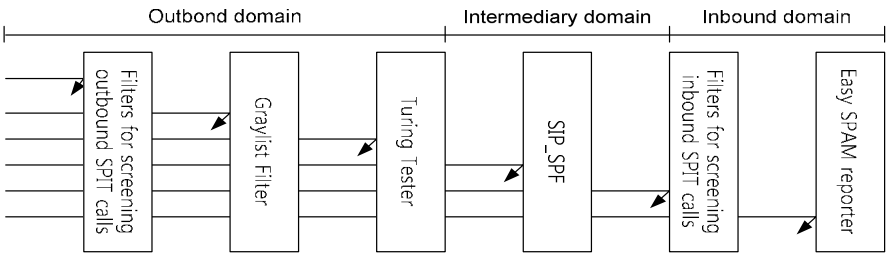

Figure 2. the multilayered anti-spit framework

3.1 Countermeasure in Outbound Domain

Once a call signal is initiated, media spam is directly sent to receivers over RTP (Real-time Transport Protocol). For this reason, it is important to prevent spam call signals before a spammer tries to initiate call signals. In outbound domain, we have a focus on detecting spam call patterns and preventing them.

Module for Detecting SPAM Caller

TLS assures receivers that integrity of incoming messages has been preserved completely, but a spammer can also transmit malicious messages to someone over TLS if he disguises himself as a normal user. So we proposes a graylist scheme to

manage users according to policies for detecting and preventing spam calls. Graylist has three states which are unknown, gray, and black respectively. Administrator configures threshold by which boundary of each state is defined. Graylist calculates SPIT level of each user. If a user's SPIT level exceeds threshold configured by administrator, his current state is transited to the next state. Finally, if a user is transited to black state, his call should be blocked.

SPIT Level Decision Model

SPIT level decision model is defined by three states, each of which means current state of a caller. *Su* means a state that it is difficult to define SPIT level of a caller. *Sg* means a state of a caller transited from *Su* state by change of his specific attributes. *Sb* is a state that a caller is clearly a spammer. SPIT level decision model consists of *T, X, S,* and Δ . Time element *T* is defined for defining state transition that can occur after a regular time. Each element is defined in Figure 3.

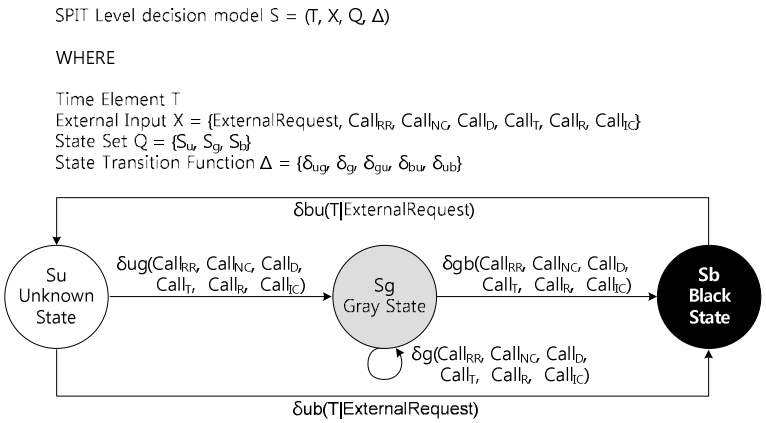

Figure 3. State transition flow of SPIT level decision model

SPIT Level Decision Algorithm

Call Detail Records (CDR) [5] includes several factors such as caller identity, callee identity, call start time, call end time, call traffic, and call rejection. Using these factors, we propose a SPIT level decision algorithm. To define SPIT level, these factors need to be calculated as a quantitative value. Above all, the calculated result should be regulated and reflected to SPIT level. Table 1 shows how each factor is calculated and the portion of each factor. The range has the fixed value from 0 to 1. Note that, each portion can be changed according to the administrator's experience or preference [6], [7]. The administrator needs to find the most reasonable portion through trial and errors.

Tabe 1. Factors to calculate SPIT level and portion of factors

Factor	Calculation expression	Portion
Number of call	NumOfCallRecipient is divided by NumOfConnectedCalls	50%

recipient		
Call duration	The number of meaningful calls that can be recognized as a normal call time among all call attempts of a caller.	30%
Average call traffic rate	If a caller generates traffic over 10% of average call traffic, his call is considered as a spam call.	10%
Call rejection rate	Call rejection rate of a caller can be calculated by getting how many calls were rejected among his call attempts.	5%
Inter call time	The time interval between call attempts of a caller per unit time.	3%
Call rate	The number of call attempts of a caller per unit time.	2%

Turing Tester

Some normal callers, however, could have similar call traffic pattern with that of spammers. In this case, they can be suspicious as a potential spammer regardless of their intention. We implemented turing tester on the hardware phone to prevent such misclassification. Turing Tester is used for screening real spammers from suspicious callers. In general, spammers tend to automatically send unsolicited bulk messages or media to others using software, but software is not able to response the answer from a callee. Therefore, if a suspicious caller provides the correct answer to turing tester, the caller is classified into the normal caller group; otherwise, the caller should be classified into the blacklist and his call is blocked continuously without further test. Although the turing test is easy enough to provide the correct answer, we should consider that some callers could be classified into blacklist due to their mistake. Thus, the administrator should able to recovery the caller's status when the caller requests it.

3.2 Countermeasure in Intermediary Domain

In outbound domain, SPF authorizes hosts to use the domain name as the originator of outbounding the SIP message. By communicating with SPF, a receiver can check whether the incoming messages were sent by one of the authorized hosts [8]. Using Figure 4, we describe how SPF helps to prevent forgeries. First, administrator in outbound domain states IP addresses on SPF records. For example, SPF records might be stated as `IN TXT "v=spf1 mx ip4:outbound.org/24 -all".` This sentence means that if a host wants to send messages using outbound.org domain name, the IP address of the host have to belong to C class subnet of an IP address registered on MX record managed by the outbound.org domain.

(1) SPF records are published to DNS. (2)' A spammer forgeries the return-path as 070xxxxxxxx@outbound.org and forwards it to SIP Proxy_I directly. (3) SIP Proxy_I asks whether the spammer's IP address is authorized to use the domain name "outbound.org" to DNS. (4) DNS gives SIP Proxy_I the answer. (5) If the answer includes "fail", SIP Proxy_I rejects the incoming message.

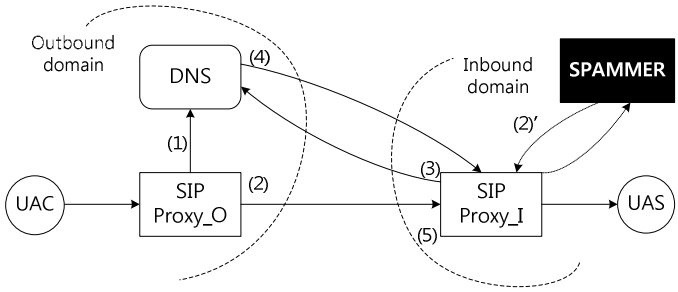

Figure 4. Processing flows of proxy server with SPF module

3.3 Countermeasure in Inbound Domain

The ease of spam victim's feedback enables to update spam information immediately and the feedback will improve performance of spam filters after all. "Easy Spam Report" enables a callee to report spam information to the administrator to check the reported spam contents. If it is clear spam, administrator reflects it to the blacklist.

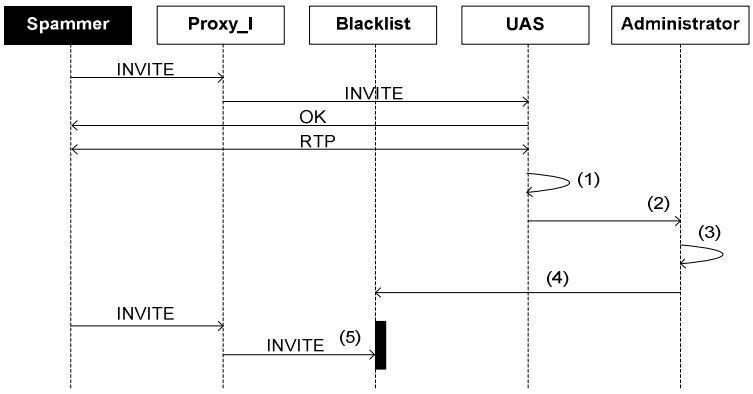

Figure 5. Processing flows of proxy server with SPF module

Figure 5 explains the usage of this system in detail. (1) UAS records the SPIT content received from a spammer as a wav file. (2) UAS reports the wave file to Administrator. (3) By listening it, the administrator checks whether it is spam or not. (4) If it is a clear spam, administrator registered the spammer's number to the blacklist. (5) the spammer's call is blocked by the updated blacklist.

4 Conclusion

This paper proposed a phased framework for countering VoIP SPAM. We showed that the phased anti-spit framework is enough to mitigate spam attacks. Although the proposed framework has multilayered anti-spit solutions, each solution is operated by an administrator's experience and preference. Therefore, it is not easy to ensure the individual user's preference in current anti-spit framework. In the future work, we plan to design and implement user reputation system based on social network analysis. The individual user's reputation information makes anti-spit solutions strong, when it is shared and deployed among domains. We expect that current framework will be enhanced through further study such as the user reputation system and real-time blockhole list.

References

1. Hype Cycle for Consumer Technologies in 2007, http://www.gartner.com
2. Rosenburg, J., et al.: SIP: Session Initiation Protocol, RFC 3261, June 2002.
3. VOIPSA.: http://www.voipsa.org/Activities/taxonomy.php
4. Rosenberg, J., Jennings, C.: The session Initiation Protocol(SIP) and Spam. Fab. 2007. IETF-DRAFT draft-ietf-sipping-spam-04.txt.
5. CDR: Call Detail Records, http://en.wikipedia.org/wiki/Call_detail_record
6. Wang F., Mo, Y., and Huang B.: "P2P-AVS: P2P Based Cooperative VoIP Spam Filtering", The Proceedings of the IEEE Wireless Communications & Networking Conference (WCNC2007), Hong Kong, Mar. 2007, pp 3550-3555
7. Balasubramaniyan, V., Ahamad, M., Park, H.: "CallRank: Combating SPIT Using Call Duration, Social Networks and Global Reputation", The Proceedings of the Fourth Conference on Email and Anti-Spam (CEAS2007), Mountain View, CA, Aug. 2007.
8. Wong. M.: Sender Policy Framework (SPF) for Authorizing the Use of Domains in E-Mail, Version 1, RFC 4408, April 2005.