VoIP-aware network attack detection based on statistics and behavior of SIP traffic

Jonghan Lee · Kyumin Cho · ChangYong Lee · Seungjoo Kim

Received: 21 January 2014 / Accepted: 14 May 2014 / Published online: 12 June 2014 © The Author(s) 2014. This article is published with open access at Springerlink.com

Abstract VoIP is one of the most popular Internet services. However, VoIP service is vulnerable to several potential security threats. Moreover, existing IP-based security solutions are unable to inspect call setup information. In this paper, we propose a VoIP-aware attack-detection scheme. The proposed scheme is able to detect VoIP network attacks including VoIP DoS and SPAM. It can detect VoIP DoS attacks with low false negatives using a statistics-based detection algorithm and can recognize SPAM with low false positives using a caller behavior-based detection algorithm. We have included experimental results to confirm the proposed scheme.

Keywords VoIP · SIP · VoIP DoS · SPAM · Attack detection · Statistic-based detection · Behavior-based detection

1 Introduction

The early residential Internet used DSL (Digital Subscriber Line) through PSTN (Public Switched Telephone Network) for sending e-mail and browsing web pages. Nowadays, the Internet infrastructure is installed ubiquitously and can be accessed in the street or on public transportation. Furthermore, mobile Internet services are available with wireless communication technologies such as WLAN (Wireless LAN) and LTE (Long Term Evolution).

VoIP (Voice over Internet Protocol) service is one of the most universally popular services. Using the existing IP (Internet Protocol) infrastructure for voice and video communication, it is replacing PSTN. Because of the low cost of the VoIP service, users install VoIP Apps on smart phones.

However, the VoIP service has many security vulnerabilities such as DoS (Denial of Service), VoIP SPAM, and eavesdropping. It inherits every threat from the IP network and includes new threats from its protocols including SIP (Session Initiation Protocol) and H.323. PSTN has a closed network architecture and provides communication using physical line switching. Therefore, the attacker must have a knowledge of the professional PSTN signal process to access the traffic. However, the IP network is open to any Internet user. Consequently, it is less difficult to access VoIP systems such as proxy servers, IP-PBXs (Private Branch Exchange), and IP-phones. The telephone is a real-time service and VoIP is a natural technology to compete with PSTN. However, VoIP must provide PSTN-level security to ensure continued VoIP popularity growth.

In this paper, we focus on two attack methods in the VoIP network. The first is the DoS attack [1] initiated by sending numerous INVITE packets or malformed messages in a short period to designated targets. DoS attacks are already common in other IP services. VoIP uses a unique routing algorithm utilizing a proxy server, SBC (Session Border Controller), and...
IP-PBX; it has several methods such as OPTIONS, INVITE, and REGISTER for telephony communication. Therefore, VoIP-DoS attacks are difficult to detect and mitigate with the existing IP-based security solutions such as firewalls and IPSs (Intrusion Prevention System). The second attack method is VoIP SPAM. Attackers send vast quantities of VoIP SPAM using automated machine processes. In this paper, we describe these attack methods and propose a detection system. Experimental results are included to confirm the proposed scheme.

The remainder of this paper is as follows. In Section 2, the characteristics of an SIP-based VoIP service and consequential security threats are addressed. In Section 3, we propose a novel detection and mitigation scheme to address VoIP network attacks and a system structure based on statistics and call behavior. In Section 4, we evaluate the proposed scheme with experimental results. We conclude this paper in Section 5.

2 Related works

2.1 SIP-based VoIP

Existing security solutions use 5-tuple information (source IP, source port, destination IP, destination port, protocol (TCP, UDP, ICMP)) from the IP traffic to analyze and detect abnormalities. The VoIP application service is based on the SIP (Session Initiation Protocol) protocol [2]. This uses an additional user identifier, URI (Unique Resource Identifier) and the pattern of the traffic is dependent upon the URI information. The URI is a telephone number in PSTN and e-mail address in an E-mail service. However, existing security solutions are not able to capture the URI correctly making the precise analysis of SIP traffic patterns and detection of VoIP abnormalities difficult.

The VoIP service uses a proxy server and URI-based routing algorithm on its overlay network. Call establishment traffic is sent first to the proxy server. The proxy server acquires the corresponding real-IP address from the registration database and forwards the SIP packets to the destination. During this process, the IP information of the IP header is changed at each node. With only the IP and port information from the IP header packet, it is impossible to notify the final destination of the traffic.

As explained above, the VoIP call establishment traffic passes through proxy servers. Figure 1 illustrates a simple example of this process. There are three different sections of the VoIP traffic transmission. They are the section between the callers and server, between servers, and between server and recipients.

The VoIP traffic in each section can be characterized as follows:

1. Section 1: destination IP address is fixed to SIP proxy server;
2. Section 2: source and destination IP addresses are fixed to SIP proxy server;
3. Section 3: source IP address is fixed to SIP proxy server.

Because of these characteristics, the traffic pattern of one user can be analyzed differently depending upon the capture point.

The abnormal pattern of SIP traffic can be characterized by the type of the transmitted messages. A VoIP session is established by a VoIP request transaction and response messages such as INVITE, 200 OK, and ACK. In a normal call condition, the ratio of VoIP messages is maintained as a designated value. For example, the REGISTER method is generally the most frequently sent method. The ratio of INVITE and BYE is almost the same. However, in an abnormal condition, the shape of the ratio is different. Figure 2 is a simple INVITE flooding attack. In this attack, the attacker overwrites the proxy server with numerous INVITE messages in a short period, and the ratio of INVITE messages increases instantly.

2.2 Network attacks of VoIP

As we defined in the Introduction, VoIP network attacks are caused by abnormal VoIP traffic. In this paper, we focus on VoIP DoS and SPAM attacks.

The methodology of a VoIP DoS attack is the same as an existing IP-based DoS attack. The attacker directs numerous VoIP messages to the VoIP proxy server of the VoIP phone. This leads to a deterioration of service quality, and can sometimes incapacitate the service. In the case of VoIP service, a DoS attack interrupts the entire Internet service or normal operation of the target VoIP service.

VoIP application DoS attacks can be classified into three categories. These are invite flooding, registration flooding, and RTP (Realtime Transportation Protocol [3]) flooding. Invite-flooding and registration-flooding attacks are similar to TCP SYN flood attacks on the IP network. A registration-flood attack exhausts the resources of the VoIP proxy server or registrar system by sending a vast number of registration packets to request a connection. An RTP-flooding attack is an attack using voice or media-related RTP packets that are modified at the RTP header and payload level by the attacker. In all three cases, the attacker sends numerous modified VoIP packets to the victim proxy server or VoIP phone software to increase the call execution time and degrade the quality of the call.

SPAM attacks [4, 5] are classified into two categories, one-ring SPAM where the attacker cancels the call before the target answers and induces the target to return the call to the attacker, and SPAM using automatic call systems where the attacker calls random numbers and if the call is established delivers pre-recorded commercial messages. In the former case, one user attempts to call many phone numbers in a short time and the possibility of call establishment is low. If a call is
established, there would still be no RTP media traffic transmission between the attacker and the target. Furthermore, one caller can send many INVITE messages at the same time (in normal conditions, this is impossible). In the latter situation, call frequency is typically lower than one-ring SPAM; however, it is still more frequent than a normal user’s call pattern and the establishment ratio will be lower. The majority of the targets receiving a SPAM call terminate the call quickly. Thus, the call duration will be short.

There is one additional type of SPAM that can be executed. However, it is difficult to determine a factor to distinguish it from a normal call using traffic pattern analysis. There are companies that execute telephone marketing legally. To differentiate this kind of SPAM call, the recipient must listen to the message. We do not consider this option in this paper.

2.3 Previous work

Kim et al. proposed a scheme to detect VoIP SPAM traffic based on caller behavior [5]. It uses seven factors for traffic analysis, including Requests from Administrator, Call Rejection Rate, Number of Call Recipients, Call Duration, Call Traffic, Call Rate, and Inter-Call Rate. As explained in Section 2.2, SPAMers attempt to make phone calls to different people in a short time and each call’s duration is brief. Each of these factor checks if the traffic exhibits this pattern. For example, the factor Call Rejection determines how many calls were rejected by targets per 100 calls.

However, this scheme [5] has the potential to generate false positives. The factors are derived based on a normal distribution. This implies that valid calls can be regarded as SPAM calls. Specifically, it assigns weight to specified factors such as Number of Call Recipients (50 %) and Call Duration (30 %). These weights are excessively heavy. They do improve the efficacy of the detecting function, however, numerous false positive detections can occur. For SPAM detection, a high false positive rate can be worse than a high false negative rate. The aim of the detection process is to increase convenience, not prevent service. For example, if a system allows one SPAM call to be established, it is not a significant issue. However, if a system blocks a valid call through miss-detection, it could cause serious harm to the quality of service.

3 Proposed scheme and system

In this section, we describe the proposed scheme and system structure. The objective is to detect and mitigate VoIP DoS and SPAM attacks. VoIP service is an L7 application service. Therefore, the existing L3-based detection scheme is not acceptable. The proposed scheme executes traffic analysis based on VoIP L7 information. In the first phase, it detects VoIP DoS attacks using a statistical and learning-based detection
algorithm. In the second phase, it performs VoIP SPAM detection using a user call behavior-based detection algorithm.

The assumptions of the proposed scheme are as follows:

1. The proposed scheme must detect and mitigate VoIP-aware DoS and SPAM attacks and cannot consider existing L3-based DoS and PSTN-based SPAM detection methods.
2. The proposed scheme is implemented as a module of inline-based IPS and functions with VoIP packets that are transmitted to and from a VoIP server such as a proxy server.
3. If the VoIP packet is encrypted using an end-to-end encryption algorithm, the system is unable to access the traffic and an analysis is not possible.

3.1 Statistics-based VoIP DoS detection

The statistics-based DoS detection module of the proposed scheme receives VoIP packets and inspects the L7 information such as IP, URI, Call-ID, and Method. It analyzes the VoIP traffic with this call-establishment information to identify abnormal traffic. The threshold values are determined for each detection rule. They are learned by a learning function for each day and each hour and applied to the statistic-based DoS detection module.

The structure of the module is uncomplicated. It compares current VoIP traffic with VoIP-DoS detection rules. Figure 3 shows the features of a rule. Condition 1 is a grouping condition. A group that satisfies Condition 2 is a target of detection. The proposed interval is 10 s (Table 1).

The module analyzes the number of packets having the same IP, To URI, and Call-ID information. Generally, to establish one call, a total of five messages are transmitted, such as “INVITE”, “100 Trying”, “180 Ringing”, “200 OK”, and “ACK”. This is a by-directional transmission. If the system captures only one direction, this number will be less. Attackers generally use INVITE or REGISTER messages for DoS attacks. These two methods are sent only once at the beginning of a session. It is impossible for a human to send several INVITE or REGISTER messages per second. Conversely, attackers typically attempt to send numerous INVITE or REGISTER packets at the same time. To accomplish this, they use duplicated VoIP packets without changing the Call-ID and URI.

The threshold values are updated at 00:00, 08:00, and 18:00 h. The maximum value for the interval becomes the threshold. The module only uses normal traffic (traffic judged as normal) for the study (Figs. 4 and 5).

3.2 Call behavior-based SPAM detection

User call behavior-based VoIP SPAM detection collects and analyzes users’ past call patterns from a Caller Profiling DB to detect VoIP SPAM. The proposed scheme uses five factors for the detection. These factors are from the scheme of [5]. However, to lower the false positive rate, we excluded normal distribution in the calculation of Call Recipient and Call Duration. Only the traffic that is pertinent to all the conditions can be judged as SPAM traffic. The proposed scheme focuses on lowering the false positive rate. The factors for detection are as follow.

- Call Recipient (CR)
  1. Collecting info from most recent 30 (N) calls of each caller
  2. Acquiring call recipient info
  3. Exclusion of duplicated call recipient (Core Operation)

| Role | Factor  | Description                        |
|------|---------|------------------------------------|
| Grouping and count condition | src_ip | source IP                          |
|      | dst_ip  | destination IP                     |
|      | from    | Caller URI                         |
|      | to      | Callee URI                         |
|      | method  | Request Method                     |
|      | status_code | Response Code               |
|      | call_id | Call Identifier                    |
|      | direction | Inbound/Outbound              |
| Interval of detection | interval | DoS detection interval            |
| Threshold | threshold | Threshold for DoS detection (learned by study function) |
4. Deriving the number of call recipients (CallRecipientNum)
5. Deriving call recipient number Rate (CallRecipientRate)

\[
\text{CallRecipientRate} = \frac{\text{CallRecipientNum}}{N}
\]

- Call Duration (CD)
  1. Collecting info from most recent 30 (N) calls of each caller
  2. Acquiring call duration info (Call End Time - Call Start)
  3. Excluding calls having longer Call Duration than Threshold (Core Operation)
  4. Deriving the number of short length Calls (CallDurationNumShort)

5. Deriving Call Duration Rate (CallDurationRate)

\[
\text{CallDurationRate} = \frac{\text{CallDurationNumShort}}{N}
\]

- Call Rejection Rate (CRR)
  1. Collecting info from most recent 30 (N) calls of each caller
  2. Acquiring call rejection info
  3. Counting corresponding Caller Rejections (CallRejNum)
  4. Deriving normal distribution of total caller call rejection count (using CR Mean and CR Std. Dev.)
  5. Deriving the position of corresponding caller call rejection count (CallRejNum) over the accumulated normal distribution.

\[
\text{CallRejectionRate} = \frac{1}{\sqrt{2\pi\text{CallRejStdDev}}} \exp \left( -\frac{1}{2\text{CallRejStdDev}^2} (\text{CallRejNum} - \text{CallRejMean})^2 \right)
\]

- Inter-Call Time (ICT)
  1. Collecting info from most recent 30 (N) calls of each caller
  2. Acquiring Inter-Call Time info (Current Call Start Time - Previous Call End Time)
3. Deriving average of Inter-Call Time (ICT AVG) (Core Operation)  
4. Deriving normal distribution of total caller Inter-Call Time (Using ICT Mean and ICT StdDev)  
5. Deriving the position of corresponding caller ratio over Inter-Call Time Average (ICT AVG) normal distribution

\[
\text{InterCallTime} = \frac{1}{\sqrt{2\pi \text{ICT StdDev}}} \exp \left( -\frac{1}{2 \text{ICT StdDev}^2} (\text{ICT AVG} - \text{ICT Mean})^2 \right)
\]  

\[
\text{CallRate} = \frac{1}{\sqrt{2\pi \text{CR StdDev}}} \exp \left( -\frac{1}{2 \text{CR StdDev}^2} (\text{CR AVG} - \text{CR Mean})^2 \right)
\]  

• Call Rate (CRa)

1. Collecting info from most recent 30 (N) calls of each caller  
2. Deriving the range of time of corresponding caller’s recent 100 calls

3. Deriving average of CallRate (CRAVG) (Core Operation)  
4. Deriving normal distribution of total caller Call Rate (using CR Mean and CR StdDev) (Fig. 6)  
5. Deriving the position of corresponding caller CallRate Average (CRAVG) over the normal distribution

The objective of the scheme is to exclude normal distribution and apply a strong condition for detection. It is designed to perform the SPAM detection process with few false positives.
3.3 System implementation

The proposed scheme is implemented as follows. The received SIP traffic is first analyzed by the statistics-based DoS detection module. The module determines if the traffic has an abnormal VoIP DoS traffic pattern. If the traffic is judged abnormal, the module alerts the administrator and drops the corresponding packets. If it is judged normal, the information is sent to the learning module to be learned and to derive a new threshold. The proposed system then uses the behavior-based SPAM detection module for the second step of the detection. This module includes a function to detect a VoIP SPAM traffic pattern based on each caller’s call behavior such as call frequency and call duration. With these two modules, the proposed system can detect VoIP DoS and SPAM traffic effectively.

4 Evaluation

4.1 Experimental testbed

To validate and evaluate the proposed system, we employed a testbed consisting of a VoIP proxy server with 80 VoIP phones and a VoIP attack emulator. The proposed scheme was implemented on a system with a dual Intel E5130 CPU 2.0 GHz dual-core processor and 4 GB memory. The process ran on a Linux kernel 2.6.25.5.

To ensure reliable test data, we used the real-world VoIP infrastructure of the Korea Internet & Security Agency. To prevent a serious negative influence on the system, a VoIP DoS attack using an enormous number of INVITE messages was not attempted. Figure 7 is a diagram of the testbed.

The VoIP DoS attacks were generated with a VoIP threat emulator, ThreatEX. It is able to generate specified VoIP messages continuously. The user can modify the packet source/destination IP, Call-ID, From/To URI, and Method type. We executed a VoIP DoS attack with INVITE flooding five times. Each attack was maintained for 60 s. Table 2 shows the properties of the attacks.

Table 3 presents the result of the VoIP DoS prevention test. The detection rate was derived using $\frac{\text{Detection}}{\text{TAC}} \times 100$. As shown, the proposed system detected essentially every VoIP DoS attack.
DoS packet regardless of the transmitted packets per second rate. There were a minimal number of false positives. Verifying the raw packets indicated that there were some VoIP packets with poorly formatted information. Some of the attack packets contained invalid data and others did not contain a valid URI field.

Consequently, we re-tested the system with an additional module, a malformed-packet filter. This module is included in the proposed system. Because it does not have a direct relationship to the DoS and SPAM detection, we have not described the details of this function in this paper. Table 4 displays the result of the VoIP DoS prevention test using the malformed-packet filter.

For all the tests, the proposed system detected 100 % of the malicious VoIP packets and the block rate increased. Some of attack packets were not blocked. Blocking is executed after the detection and until to detect the attack it has to pass some packets, even though they are malicious, because it has to collect packets for 10 s to judge if the traffic has a pattern of attacks or not.

The proposed system detected 88.3 % of one-ring SPAM calls and 90 % of automatic call system SPAM calls in the second test. It recorded only a 0.16 % false positive rate. The false positive rate is derived using Eq. (6).

\[
\text{False Positive rate} = \frac{\text{False Positive}}{\text{Entire Calls}} \times 100 \quad (6)
\]

* False Positive indicates the number of valid calls judged to be SPAM.

The proposed system erroneously detected only five calls out of 3,150 normal calls. We expect that as the size of the dataset increases, there will be additional detail information for detection, and the false positive rate will decrease (Table 5).

### 5 Conclusion

In this paper, we proposed novel schemes to detect VoIP network attacks such as VoIP DoS and SPAM. Our contribution increased the efficacy of VoIP detection using a statistical learning-based VoIP DoS detection scheme and decreased the false positive rate of SPAM detection using enhanced analysis factors and strong conditions.

The proposed scheme can be implemented as a module of inline-based IPS systems and can work as a practical VoIP-aware network security solution. We confirmed the efficacy of the proposed system with experimental results.

Our future work is to improve the SPAM detection scheme and increase the detection rate to over 95 %. We may also prepare a more detailed and extensive dataset for the next experiment.

### Acknowledgments

Supported by a Korea University Grant and MSIP (Ministry of Science, ICT and Future Planning), Korea, under the ITRC (Information Technology Research Center) support program (NIPA-2014-H0301-14-1004) supervised by the NIPA (National IT Industry Promotion Agency).

### Open Access

This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

### References

1. Sadiwala R, Sharma M (2014) Security threats of VoIP. J Innov Trends Sci Pharm Technol 1:7–18
2. Lamba RK, Kaur K (2014) Security traits of VoIP. Int J Res Advent Technol 2(3):178–181
3. Chaisamran N, Okuda T, Yamaguchi S (2013) Trust-based VoIP spam detection based on calling behaviors and human relationships. Inf Media Technol 8(2):528–537
4. Yeon K, Kim H (2013) Design of standard VoIP spam report format supporting various spam report methods. J Secur Eng 10.1:1–8
5. Kim H (2009) DEVS-based modeling of VoIP spam callers’ behavior for SPIT level calculation. Simul Model Pract Theory 17(4):569–584

**Jonghan Lee** received his B.S degree in Information Engineering from Sungkyunkwan University (SKKU) of Korea, in 1994 and also received his M.S degree in business school(MBA) from Yonsei University of Korea, in 2009. He is currently working toward the Ph.D. degree in Information Security, Korea University, Korea. His research interests include conditional access system(CAS), policy of broadcasting and communication service and IoT(Internet of Things) service protection.

**Kyumin Cho** received his B.S degree in Computer Science from Seoul National University of Korea, in 1993 and also received his M.S degree in Information Security from Dongguk University of Korea, in 2002. He is currently working toward the Ph.D. degree in Information Security, Korea University, Korea. His research interests include information security, personal data protection and information assurance.

**ChangYong Lee** received his B.S degree in Computer Engineering from Kangwon University of Korea, in 2004 and also received M.S degree in Information Security from Hanyang University of Korea, in 2008. He is currently working in the Korea Internet & Security Agency(KISA) of Seoul, Korea. He works as Senior Researcher of the Korea Internet & Security Agency(KISA). His research interests include mobile network monitoring and mobile malware detection.

**Seungjoo Kim** received his B.S., M.S. and Ph.D. from Sungkyunkwan University (SKKU) of Korea, in 1994, 1996 and 1999, respectively. Prior to joining the faculty at Korea University (KU) in 2011, He served as Assistant & Associate Professor at SKKU for 7 years. Before that, He served as Director of the Cryptographic Technology Team and the (CC-based) IT Security Evaluation Team of the Korea Internet & Security Agency (KISA) for 5 years. He is currently a Professor in the Graduate School of Information Security at KU, and a member of KU’s Center for Information Security Technologies (CIST). Also, He is a Founder and Advisory Director of a hacker group, HARU and an international security & hacking conference, SECUINSIDE. Prof. Seungjoo Kim's research interests are mainly on cryptography, Cyber-Physical Security, IoT Security, and HCI Security. He is a corresponding author.