DDoS attacks in VoIP: a brief review of detection and mitigation techniques

Sambath Narayanan*, Selvakumar Manickam, Yu-Beng Leau

National Advanced IPv6 Centre (NAv6), Universiti Sains Malaysia (USM), Penang, Malaysia

A R T I C L E   I N F O

Article history:
Received 20 July 2016
Received in revised form 28 August 2016
Accepted 25 September 2016

Keywords:
VoIP
SIP
DDoS
Security

A B S T R A C T

Voice communication in recent trends has shown rapid growth in homes and businesses with the development of Voice over Internet Protocol (VoIP). The growth in VoIP subscribers was determined by the increase in VoIP flexibility, Quality of Service and monetary savings. The fall in public switched telephone network and raise in phone portability migrated PSTN to VoIP. The Session initiation protocol being an application layer protocol helps to create session between the caller and the called for bidirectional communication using SIP messages. The VoIP became targeted victim of different attacks as internet became the medium of transmission. The security vulnerabilities arise from new protocols and the existing infrastructure of traditional data network. Flood-based attacks are more threatening and annoying than other attacks. This brief review paper discusses on different types of VoIP attacks along with the existing VoIP detection and mitigation techniques based on Entropy, Wavelet, Sketch and Hellinger distance, Sunshine and RQA are presented.

© 2016 The Authors. Published by IASE. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

One of the emerging technology rapidly embraced by the telecommunication market is VoIP. This new technology, deploying the services of Public Switched Telephone Network (PSTN), has the ability to extremely change voice communication services over the internet. Traditional PSTN is now being substituted by VoIP whose services are replaced abundantly in homes and enterprises. In 1995, Vocaltech Inc., introduced VoIP and their internet phone allowed the users to communicate via computers. VoIP helped to transmit multimedia data in a single infrastructure. As stated in (Zhao and Ansari, 2012), VoIP calls are made using peer to peer VoIP through computer, IP telephony as well as traditional phones. It paved way for monetary savings by lowering cost of user services, which increased flexibility and popularity of VoIP. It provides better Quality of Service (QoS) than PSTN at comparatively less cost. The local call rates are reduced up to 40% and international call rates are reduced up to 90% using VoIP technology as stated in (Heckstall, 2016). Hence the voice network along with data network are integrated to lower the management cost and effort.

In 1998 less than 1% of all voice calls was used by VoIP. There was a slow increase in VoIP users, which accounted for 3% in 2000 and raised to 25% by 2003 (Hallock, 2004). In 2013, the Point Topic organization tracked the global VoIP operators and recorded a total of 155.2 million global subscribers (Topic, 2013). Subscriptions for VoIP have increased substantially worldwide (Wansink, 2016) and predicted to grow further by 2020. IBISWorld states that the VoIP industry’s contribution is expected to increase 15.3% every year until the year 2017 as stated in (IBIS, 2015). Due to this estimated increase in the near future, the flexibility for both residential customers and businesses in VoIP technology will substantially increase. On surveying, the fall of fixed-line PSTN subscriptions is compensated by the rise of the VoIP subscriber base. Another advantage of VoIP is phone portability where the device uses the same number all over the world. But in a legacy phone, the device is assigned a fixed number for a fixed location and this device number change when moved to a different location. As a result, migration of PSTN to VoIP is inevitable.

2. VoIP architecture

The components of VoIP include the source agent, ATA (Analog Telephone Adapter), server, gateway...
and destination agent. The VoIP architecture is shown in Fig. 1.

![VoIP architecture](image)

**Fig. 1: VoIP architecture**

The voice signal from the source user agent is transmitted to ATA and then to router to generate IP packets. A dial tone from ATA signifies connection to the internet. When a number is dialed the tone is converted to digital data. These packets passing through the IP network reach the Session Initiation Protocol (SIP) server. The SIP server locates the destination user agent with the help of location server. Here the phone number is checked for validity and then they are mapped to an IP address. The packets are then passed on to the termination carrier which acts as the gateway to be connected to the destination user agent. Thus a session is established between the two agents. To communicate between these agents, a uniform protocol should be used among them. The communication between the source user agent, server and the destination user agent is linked by the SIP protocol. The ATA at receiver end converts the packets back to analog audio signals. The session is terminated by hanging up the phone.

3. SIP architecture

The standardization of SIP by Internet Engineering Task Force (IETF) is used by VoIP and other multimedia bidirectional communication like voice calls, video conferencing and data sharing. SIP is an application layer protocol which creates, modifies and terminates sessions in VoIP communications. Since SIP is a simple and flexible protocol, it can add features to it. It allows multiple multimedia sessions in one call as seen in online gaming, instant messaging and various services. The Uniform Resource Locators (URL) addressing scheme in SIP does not depend on physical location. They are addressed by either a phone number, an IP address, or an e-mail address. It is similar to the HTTP web protocol as the messages comprise of headers and body message. The default port for SIP is 5060 for either TCP or UDP. The user datagram protocol (UDP) over the transmission control protocol (TCP) at the transport layer is favored by SIP because of the connection orientation of SIP and simple behavior of UDP.

The three major components in a SIP communication are User Agent Client (UAC), the SIP proxy server and User Agent Server (UAS). The main network elements involved in the SIP communications are described in this section. Fig. 2 illustrates the structure of SIP.

![SIP architecture for VoIP](image)

**Fig. 2: SIP architecture for VoIP**

The User agent (UA) generates or receives SIP messages. It acts as a UAC for transmitting SIP messages; the receiver act as UAS. The SIP client acts as both a SIP UAC and SIP UAS. The SIP request from user agents is received by the SIP server and forwards them to the corresponding host. The Registrar server processes REGISTER messages and then the users URI are mapped to the present location of the user. The registrar server can be placed separately or inside the SIP proxy server. The Location server helps to store the location of registered users. The proxy finds the user's location using location and registrar server.

The Feature servers provide special treatment to enhance communications experience. The proxy server uses special routing rules to control SIP feature. The Media server record media streams, play back recorded media, collect DTMF input from the user. It uses a media bridge that mixes multiple media streams as in conferencing. When a user is on a different domain the calls are connected using Domain Name System (DNS). The redirect server returns a forwarding address when a user is moved temporarily from the current domain to the next domain. Session Border Controller (SBC) in SIP protects the internal network from malicious attack. Signaling and Media Gateway enables non-SIP interactions. Signaling gateway translates signaling protocol and the media gateway transcodes media data.

The SIP components are provided with SIP address which resembles an email address. This address contains a username followed by hostname. SIP operations are performed between
them by exchanging messages. SIP messages used for communication purpose have message header similar to HTTP. These messages are in the form of request and response. UAC uses request message and UAS uses response message. The SIP request messages are REGISTER, INVITE, ACK, CANCEL, BYE, and OPTIONS. The SIP response messages are PROVISIONAL (1XX), SUCCESS (2XX), REDIRECTION (3XX), CLIENT ERROR (4XX), SERVER ERROR (5XX) and GLOBAL FAILURE (6XX). SIP message being in the form of text-based presentation is vulnerable to attacks. VoIP by its popularity acquires security issues due to new protocols and components of the network which exploits the VoIP services. The first and foremost concern about communication protocol was reliability and efficiency. But security concerns were given less importance and became endangered to multiple attacks. Flood-based DoS attacks (Cha et al., 2007) have been identified as major threats among the other attacks. VoIP switches are more vulnerable to a wide range of network attacks like DDoS, eavesdropping, man in the middle attack as the internet being the transition medium among the internal and external users.

4. Security issues faced by VoIP

The maximum number of VoIP devices and programs has vulnerable spots for intruders with a wide attack space. Confidentiality, Integrity and Availability (CIA) must have high priority while considering the security issues. Various threats affect the CIA of VoIP systems. Malformed Message Attack uses the vulnerability of text-based protocol. These attackers manipulate on SIP header deletion, overflow-space, non-ASCII code to malfunction the proxy server (Sonkar et al., 2012). Spoofing Attack involves an attacker masquerading himself as a legitimate user. Fraudulent emails, fake websites and wireless access point are provided to trick victims in collecting their personal data. Spoofed BYE messages can terminate sessions between the users (Sonkar et al., 2012). Eavesdropping impacts on confidentiality of the VoIP user agent. The attacker monitors both the signaling and data streams between the user agents. They can reply to the conversation and obtain secured information. The Man in the Middle (MITM) Attack encounter the confidentiality and integrity of user agents. The attackers listen to the conversation between the two user agents and masquerade on both the side as a legitimate user. Spam over Internet Telephony (SPIT) involves the generation of unsolicited advertisements and unwanted calls to the users. In terms of bandwidth and cost, SPIT is a potential risk (Ekekwe and Maduka, 2007). Call Hijacking involves the attacker impersonating a user agent by spoofing the identity of the phone device. The VoIP device is setup with victim’s identity. Hence incoming calls can be redirected to the attacker’s phone (Butcher et al., 2007). Registration hijacking involves an attacker replacing the legitimate registration with false data.

5. Denial of service (DoS) attack

DoS is a flooding attack which involves the SIP phones to generate excessive SIP messages to a specific user agent within a short period of time. This discards the services rendered by the user agents. VoIP telephony services are interrupted by the attacker due to excessive requests from the source agent to the destination user agent or the server (Sisalem et al., 2006). When the attack is on a particular user, the user cannot respond to the calls. But when the attack is on the server, the entire network cannot transmit or receive calls. Hence legitimate users are denied services. DoS impacts include exhaustion of resources on the network like RAM usage, CPU performance and bandwidth consumption. Dos attacks are done by Physical, VoIP Signaling and VoIP Media. In VoIP Signaling DoS attack, several call setup requests are created by the attacker. The processing power of proxy server or terminal is consumed by these attacks. VoIP services are interrupted for legitimate users by sending them a large number of INVITE call request per second. The pending call set up signals is cancelled by sending a CANCEL, GOODBYE or PORT UNREACHABLE message. This makes the phone unable to setup calls. Hence the quality of service is degraded quickly to an unacceptable level. INVITE flooding and Bye flooding are the most annoying attacks for VoIP users. In VoIP Media DoS attack, the attackers flood several RTP packets on IP phone, gateway and other VoIP components. If the legitimate user interrupts RTP packets, the voice quality degrades. If any one of the SIP components failed, the whole SIP network halts down. Physical DoS attacks involve power outage and physical damage to network components. Traditional telephone operates even during power outages due to 48 volts input provided by the telephone line whereas VoIP requires power supply. The attacker physically accessing the SIP components may interrupt its normal services by plugging out the power cord or network cable.

6. Distributed denial of service (DDoS) attack

In a DDoS attack, multiple machines generate more attack traffic to the targeted system. This is achieved using botnets or zombies. The attack can be made on network, transport and application layer. The attacker controls and activates the botnets on a particular time using a controller. The targeted victim is attacked by the attacker using zombies via a controller. Several fake requests from spoofed IP are generated by the attacker. The server misunderstands that the senders of fake requests are legitimate users. Hence the target servers are flooded by sending a number of fake requests. The corresponding response message sent by the server floods the target. Hence multiple machines are harder to detect when compared to the detection process in DoS attack which attacks using a single machine.
The strong growth in DoS and DDoS attacks targeting SIP/VoIP services has raised up from 9% in 2014 to 19% in 2015 as reported by Arbor Network in "The 2016 Worldwide Infrastructure Security Report (WISR)". Based on the above threats, it is clearly seen that DoS and DDoS are the most vulnerable attack in real-time VoIP system which disrupts legitimate user's services leading to service unavailability.

7. DDoS detection and mitigation techniques in VoIP

There are several detection and mitigation techniques available for VoIP DDoS attack. Some of them are Entropy (Tritilanunt et al., 2010), Wavelet (Li and Li, 2009), Sketch and Hellinger distance (Tang et al., 2012), Sunshine framework (Hoffstadt, et al., 2014) and Recurrence Quantification based Approach (Jeyanthi et al., 2014). The above methods are contending in detail in the succeeding section.

8. Entropy

The attackers stream a wide range of fraudulent data to distort the services of the server, which leads to DoS on the server end. The volume-based techniques can only detect high volume traffic neglecting short-term DoS attacks. But huge volumetric traffic delivered by legitimate users to the server is undistinguished from higher traffic of bogus messages delivered by the attackers. The entropy detection scheme detects the header field from incoming packets. In entropy-based input-output traffic mode detection technique, the inspection on features of traffic is based on Shannon's function. The entropy $H(t)$ at a time $t$, is expressed as,

$$H(t) = -\sum_{i}^{n} \frac{m_i}{S} \log(\frac{m_i}{S})$$

(1)

where, $n$ is the number of packets with size $l$. The inspection time frame length is denoted by $S$. Packets with similarity in length are gathered and analyzed to calculate the entropy. The Shannon's function inspects a similarity and distribution of traffic in the inspection time frame. During DoS attack the entropy of the traffic in the observation window falls down. This shows the presence of a DoS attack on the network. This detection technique focusing on the entropy of a packet size also examines the packet. This feature helps to eradicate false belief stating a certain legitimate packets are identified as suspicious traffic in some situation. The author in (Tritilanunt et al., 2010) compares entropy of normal and attack traffic. This method accurately detects small DoS/DDoS attack provided with high computation time. The major problem in this method is that the attacking method is modified by the attacker by knowing the detection strategy.

9. Wavelet

The Fourier transform is used for processing stationary signals having frequencies at every period. But wavelet transforms accomplish frequency resolution and time resolution at low frequencies and high frequencies respectively. The continuous wavelet transform (CWT) calculates the correlation for each lag at possible scales. The discrete wavelet transform (DWT) calculates coefficients which are decreased with respect to the scaling factor. Due to cutoff frequencies in network traffic signals, DWT is used. The server total traffic $y(t)$ is the addition of normal traffic $n(t)$ and attack traffic $a(t)$.

$$y(t) = n(t) + a(t)$$

(2)

Network sniffing tool helps to find $y(t)$. During detection, normal traffic is found by daily or weekly cycles. The time cycle value is denoted by $c$ and $y(t)$ is similar to $y(t-c)$. The normal traffic $n(t)$ without attack can be formulated as,

$$n(t) = n(t - c) + \epsilon(t)$$

(3)

where, $\epsilon(t)$ is the noise with mean 0, $c$ is cycle value of traffic. The attack traffic is given as,

$$a(t) = y(t) - n(t - c) - \epsilon(t)$$

(4)

Thus, when the server is under no attack, $a(t) = \epsilon(t)$. The wavelet technique is used on $a(t)$ to decompose it to approximation and detail coefficients to reduce the impact of noise. The data reconstructed by detail coefficients clearly show the attack points. Finally, this method rapidly identifies whether the server is under attacks. The author in (Li and Li, 2009) proposed wavelet analysis for rapidly detecting DDoS attack. But this method depends on statistical traffic pattern before detection. It provides less efficiency in attack detection and also influenced on wavelet basis functions.

10. Sketch and Hellinger distance (SHD)

This method is developed due to inefficiency in detecting low-rate flooding attacks and multi-attribute floods by monitoring a wide range of SIP messages simultaneously. This proposed method helps in detecting and preventing the flood attacks by introducing two techniques, i.e. three-dimensional sketch design and Hellinger distance (HD) detection technique. The sketch is a data summarization technique that summarizes compact and constant-size data summary of high dimensional data streams using probability,

$$a_i = (k_i, v_i)$$

(5)

where, $k$ is the key which acts as the SIP address. The value of $v$ is assigned 1. The Hellinger distance (HD) is the technique used for calculating the distance between two probability distributions. This is the
best approach for calibrating similarities between these two data summaries. They also presented “estimation freeze mechanism”,

$$H^2(P, Q) = \frac{1}{2} \sum_{i=1}^{n} \left( \sqrt{P_i} - \sqrt{Q_i} \right)^2$$ \hspace{1cm} (6)

where, $P = (p_1, p_2, ..., p_n)$ and $Q = (q_1, q_2, ..., q_n)$.

If the two probabilities are totally different, $HD = 1$. If they are same, then $HD = 0$. The author in (Tang, et al., 2012) introduced an online method to detect and prevent SIP flooding attacks by Sketch based detection using Hellinger distance. The attack traffic is scanned for signature. This method cannot accurately detect attack.

11. Sunshine

The Sunshine framework is a detection and prevention technique on network and application level for VoIP fraudulence. The architecture of this framework is a composite of firewall and intrusion detection, a distributed sensing system, a CDR analysis, alarm component and a DNS based blacklist. This framework was designed and implemented with the Distributed Sensor System technique which acts as a scanning system. It includes Sensor Central Service (SCS) and the sensor part. The Sensor component detects for signature and reports the incoming packets on signature mismatch in the SIP network. At the initial stage, the sequences of SIP messages are recognized which are described in XML signatures. They report on recognized message sequences to the Sensor Central Service. The process of misuse detection executes three different steps. The Listener module, Analyser module and Notification module. The Listener module abducts all SIP messages from the network interface and put them in a First in first out order. The Analyzer module approach the order and then evaluate the messages by using pre-defined XML signature. The Notification module triggers SCS only after successful detection of the attack i.e. the signature mismatch. The authors in (Hoffstadt et al., 2014) proposed a comprehensive framework for detecting and preventing VoIP fraud and misuse. It prevents DDoS attack by analyzing the traffic with signature. The attack detection and mitigation time is very less provided with more accuracy. But this method involves more complicated framework.

12. Recurrence quantification approach (RQA)

The RQA detects the different types of DDoS attacks at an early stage. This technique is a mathematical approach and analyze the behavior of non-linear traffic data. The recurrence property in a dynamic system is the change in state during disturbance and restores to original position after disturbance. The traffic in the network is monitored continuously for any deviation from the normal traffic behavior to detect the presence of an attack. Recurrence Plot is a square matrix depicting the collection of pairs of times at which the trajectory is at the same place i.e. showing state x at the time i and j,

$$R(i, j) = u(\varepsilon - ||x_i - x_j||), \hspace{1cm} i, j = 1, ..., N$$ \hspace{1cm} (7)

$$u(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$$ \hspace{1cm} (8)

where, $N$ is the number of states under consideration, $\varepsilon$ is the threshold distance and $||.]||$ is a norm and $u$ is the Heaviside function. The recurrence is indicated by black dots or lines which are diagonal, horizontal or vertical. Diagonal lines show the evolution of state at different times is similar. The vertical and horizontal line shows the states do not change with time. The recurrence area is larger for the maximum norm, smallest for minimum norm and intermediate for Euclidean norm. RQA uses different parameters, for example Recurrence Rate, Entropy, Laminarity, Determinism etc. Recurrence Rate is the probability of a particular state to reoccur. The percentage of the recurrent quantifies the percentage of recurrent points within the given radius ranging from 0 up to 100%,

$$RR = \frac{\text{sum of all } R(i, j)}{N^2}$$ \hspace{1cm} (9)

Determinism is the ratio of recurrence points, forming diagonal structures to all points in the RP which vary according to the types of the signal. Periodic signals make very long diagonal lines; very short diagonal lines represent chaotic signals, whereas stochastic signals have no diagonal lines at all,

$$DET = \frac{\text{sum of all } l \times P(l)}{RR}$$ \hspace{1cm} (10)

where, $P(l)$ denotes the frequency of diagonal lines with length $l = l_{\text{min}}$ to $N$. Laminarity is the ratio of recurrence points forming vertical structures to all points in the RP,

$$LAM = \frac{\text{sum of all } v \times P(v)}{RR}$$ \hspace{1cm} (11)

where, $P(v)$ is the frequency of vertical lines with length $v = v_{\text{min}}$ to $N$. Trapping time is the average length of vertical lines, determines the duration of a system remaining in a specific state,

$$TT = LAM / \text{Total number of vertical lines}$$ \hspace{1cm} (12)

Divergence is the reciprocal of maximal diagonal line length (without LOI), estimating the positive maximal Lyapunov exponent of the dynamical system,

$$DIV = 1 / L_{\text{max}}$$ \hspace{1cm} (13)

Entropy is Shannon’s entropy of the probability $p(l)$ that a diagonal line has length exactly to $l$,

$$\text{Entropy} = -\text{sum of all } [p(l) \times \text{ln } p(l)]$$ \hspace{1cm} (14)

$$p(l) = p(l)/\text{sum of all } P(l)$$ \hspace{1cm} (15)

where,

$l = l_{\text{min}}$ to $N$
The recurrence parameters are analyzed using the above equation. The deviation of these parameters from normal value indicates a DDoS attack. The authors in (Jeyanthi et al., 2014) presented Recurrence Quantification based approach. RQA detects SIP-based attacks more efficiently. This method cannot be compromised even after knowing the detection scheme by the attacker.

<table>
<thead>
<tr>
<th>Features</th>
<th>Wavelet</th>
<th>Entropy</th>
<th>Sketch and Hellinger distance</th>
<th>Sunshine</th>
<th>RQA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
<td>Detect</td>
<td>Detect</td>
<td>Detect and prevent</td>
<td>Detect</td>
<td>Slow</td>
</tr>
<tr>
<td>Scanning signature</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Detection Time</td>
<td>Fast</td>
<td>Slow</td>
<td>Slow</td>
<td>Moderate</td>
<td>Moderate</td>
</tr>
<tr>
<td>Efficiency</td>
<td>Less efficient</td>
<td>More efficient</td>
<td>Less efficient</td>
<td>More efficient</td>
<td>More efficient</td>
</tr>
</tbody>
</table>

The recurrence parameters are analyzed using the above equation. The deviation of these parameters from normal value indicates a DDoS attack. The authors in (Jeyanthi et al., 2014) presented Recurrence Quantification based approach. RQA detects SIP-based attacks more efficiently. This method cannot be compromised even after knowing the detection scheme by the attacker.

The Sunshine and SHD uses signature-based scanning to detect and prevent SIP-based attacks. The sunshine method involves more framework system leading to high computation time. Moreover, the sketch design and Hellinger distance technique being a statistical model can be improved by integrating additional schemes. In future, the integration of sketch and Hellinger distance with load balancing rate limiter can be implemented to accurately detect and prevent multi attribute VoIP attacks.

Acknowledgement

The guidance is endorsed by the USM Fellowship and National Advanced IPv6 Centre. The authors would like to thank the NAv6, USM for supporting this research.

References


13. Conclusion

This paper reviewed the state of the art on the protection of VoIP and SIP servers against DoS/DDoS attacks. The major problem of flooding attack still focuses on the guarantee of QoS in VoIP. Several common detection techniques have been discussed. The comparison table for the discussed security techniques is indicated in Table 1.
INFOCOM: 1161-1169. https://doi.org/10.1109/INFCOM.2012.6195475

