2011

Security measures for VoIP application: A state of the art review

Rafidah Md. Noor, Dr., University of Malaya
Full Length Research Paper

Security measures for VoIP application: A state of the art review

Mehdi Jahanirad*, Yahya AL-Nabhani and Rafidah Md. Noor

Department of Computer System and Technology, Faculty of Computer Science and Information Technology, University of Malaya, 50603 Kuala Lumpur, Malaysia.

Accepted 16 August, 2011

Voice over internet protocol (VoIP) is being widely used, while the integration of voice and data provide several opportunities. Lower cost and more flexibility are the main advantages of the VoIP which derived the attention of the enterprises to it. Unfortunately in the other extreme, some security issues may come across with the extensive use of VoIP. The purpose of this review paper is to encourage the VoIP industry to consider the importance of the security matters of their application and to find the gaps by discussing the pros and cons of each application and performing the fair comparison. This article begins by discussing the motivation of security measures in VoIP applications and a list of possible threats such as denial-of-service (DoS), Call Hijacking and Interception. These security issues are the most well known attacks that can be carried out in current VoIP deployment. Next, we discuss about two VoIP applications, which are Skype and GTalk, and their security levels have been measured briefly by highlighting some of the security issues in them. These security measurements are used to perform the comparison between the studied VoIP applications. Finally some methods have been proposed to improve the security in the VoIP applications in future state.

Key words: VoIP, network security, skype, Google talk, threats.

INTRODUCTION

Voice over internet protocol (VoIP) is being widely used, and is a new way of voice communication with public switched telephone networks (PSTN) and cellular networks. VoIP can be used to call any PSTN telephone or mobile phone anywhere in the world. Even though particular services can only work on a computer or a special VoIP phone, some allow a caller to use a conventional phone with an adapter. The goal of VoIP is to replace the operating circuit-switched, public switching telecommunication network to a packet-switched network. VoIP has been successful in deriving the attention of the telecommunication markets of all sizes and introducing the advanced features to the market, while on the other side the integration of the voice and data words caused evident security risks. Lower cost and more flexibility are the main advantages of the VoIP which derived the attention of the enterprises to it. As a result, it is very important to consider cautiously all the security issues before installing VoIP. The purpose of this review paper is to encourage the VoIP industry to focus on security issues of their application and to find the gaps in this industry. The other important strength to come out with this paper was to give the technical comparison of the security measures of the VoIP applications to its end-users, giving the ability to select the most applicable application. As a result, this article is prepared to study some of the security issues in VoIP which are the most well known attacks that can be carried out in a current
VoIP deployment and to investigate those in two different VoIP applications. This article discusses the motivation of VoIP, highlights threats of VoIP, provides a comparison between different VoIP application securities and concludes the issues by proposing the idea to improve the security in VoIP applications.

Motivation

VoIP is a technology that allows a user to make a call using a computer over the internet instead of a standard telephony network. VoIP uses protocols such as real-time protocol (RTP) and H.323 to deliver packets over the internet (Cisco Systems, 2002). Each VoIP packet has an internet protocol (IP)/UDP/RTP header with a total size header, 40 bytes. G.711 and G.729 are the two widely used voice encoding standards that are used with VoIP products. G.711 limits the size of voice payload to 160 bytes (20 ms of voice) or 240 bytes (30 ms of voice) at a bit rate of 64 Kbps. A large voice payload size would increase the encoding latency. G.729 limits the voice payload size to 20 bytes or 30 bytes only (Cisco Systems, 2006).

VoIP is a new service which comes in order to improve the legacy voice communication by supporting it with data communication as well. VoIP allows data, images and videos to be transmitted simultaneously (Hens and Caballero, 2008). There are many advantages and benefits (Park, 2009; Raake and Corporation, 2006) that can be gained from VoIP like:

- Cost savings: Reduce the communication cost for the users which means that the communication can be done over private or internet data network line, instead of commercial telecommunications line.
- Extensibility: VoIP can be extended easily to any number of users and without any geographical boundary limitation.
- The available resources can be reused: Available network can be used for VoIP implementation.
- Data and voice service are combined easily: Rich media of services.
- Easy implementation: Speech communications can be designed by computer networking companies within any organization.
- Collaboration and integration with other applications: This is because some protocols can collaborate with other applications easily, so it can take benefits from its properties.
- Mobility of the service: Thus, the users can use the services from anywhere like voice mail, call features and so on.
- User control interface: Thus, most of VoIP have user controls interface or graphical user interface (GUI) like in web, which make it easy to use.
- Phone portability: Thus, the users do not need to change the communication details where ever they go or move

VoIP nowadays is very common for many people who are using IP phones or using client software like Messengers (MSN, Yahoo, Skype, Google Talk, and many more). This popularity of VoIP comes from the services that can be gained by end user. In addition to voice, the users can have video conferencing, instant chat messaging, and fax data over the IP network. Figure 1 shows a conceptual view about VoIP service architecture (Park, 2009). In this Figure 1, the architecture demonstrates three types of networks; a service provider network, a consumer network and an enterprise network which are all connected to PSTN through a media gateway. In consumer network, a digital subscriber line (DSL) router receives connection from the VoIP servers which exist in the service provider network. This network provides voice, video, presence, instant messaging (IM) and internet phone services to its consumers. In the enterprise network, the VoIP servers are connected to a call manager in its center, which provide the services such as the voice using IP phones, video, fax, enterprise IM and presence.

Considering all the advantages of VoIP, bringing together voice and data on the same wire, in spite of the protocols used, employ network security problems. One of the outcomes of this integration could happen when there is a major attack in the network. This event can cause the organization’s entire telecommunications infrastructure to be at risk. To secure the whole VoIP infrastructure, many studies shall be done to perform a proper planning and analysis, and to gain comprehensive information about the details of the implementation.

Literature analysis

Securing business private data is the most important thing that most of the commercial companies and institutions care about. It is because of the new risks accomplished with the new technologies (Porter and Gough, 2007). VoIP is a new technology which emerged since late 90s (Park, 2009). Although, this new technology was accomplished with many security risks; it is still considered more secure than a traditional telephone (e.g. PSTN) communication (Thermos, 2006). Merging voice with data to be transmitted in data network by using IP as an identifier is considered as an attractive area of research which is the basis of VoIP. In order to establish this communication many protocols are used. Each of these protocols has its own properties; which make the area of attack wider.

Many businesses nowadays are migrating from the legacy traditional enterprise telephone PSTN to VoIP;
because of the benefits that can be gained by combining both voice and data together (Wallingford, 2005). It has emerged since late 90s to be one of the telecommunication media in the world (Park, 2009). Figure 2 shows the normal analogue phone (a) and IP phone (b) models, respectively (Hens and Caballero, 2008).

**Figure 1.** VoIP service architecture (Park, 2009).

**Figure 2.** Analogue phone versus IP phone (Hens and Caballero, 2008).

**Threats to VoIP communication systems**

Thermos (2006) addresses in his article that some VoIP service providers confuse what security means in packet based communications. While, VoIP service provider in North America claims that "We are more secure than a regular phone line". The article also discusses two main common attacks in VoIP deployments which focusing in signaling initiation protocol (SIP) such as:

- The ability of the attacker to hijack a VoIP communications.
- The ability of the attacker to eavesdropping VoIP communications.

There are many works that have been discussed on VoIP threats since VoIP implementation. Porter and Gough (2007) also summarise and categorise different types of risks and threats in their article. Dantu et al. (2009) also presented a brief study of attacks on a VoIP infrastructure. They classified attacks into five primary types; this article is a good survey on the previous studies which have been done on VoIP attacks and its responses to give the broad perspective about the topic.
Also they have discussed the solutions to overcome every one of the attacks, and finally they proposed the new infrastructure for the VoIP security to overcome some of the security issues. For selecting the best VoIP application, implementation of security is an important factor to be studied. In his work on threats to VoIP communication systems, Porter (2006) states four types of threat which occur more than the others in VoIP applications,

**Denial-of-service (DoS) or VoIP service disruption**

Denial-of-service (DoS) attacks can affect any IP-based network service, and are the most challenging treat in VoIP applications. One type of attack in which packets can simply be flooded into or at the target network from multiple external sources is called a distributed denial-of-service (DDoS) attack. Porter (2006) shows its architecture in Figure 3. Dantu et al. (2009) in his study on VoIP attacks and responses proposed using Firewalls, special purpose hardware, VoIP aware hardware, effective authentication systems and recovery systems to overcome this type of threat.

**Call hijacking and interception**

Call interception and eavesdropping are other major concern on VoIP networks which cause theft of information and services on VoIP networks (Benini and Sicari, 2008). The existence of this treat in VoIP applications is because of the deficiency or absence of authentication measures. This threat demonstrates the need for security services that enable entities to authenticate the originators of requests and to verify that the contents of the message and control streams have not been altered in transit (Porter, 2006). While address resolution protocol (ARP) is a fundamental Ethernet protocol. This article also states that ARP redirection, ARP spoofing, ARP hijacking, and ARP cache poisoning are related methods for disrupting the normal ARP process (Figure 4).

**H.323-specific attacks**

H.323 is signaling protocol in VoIP communications which is encoded according to ASN.1 PER encoding rules. The implementation of H.323 massage parser, rather than the encoding rules themselves cause vulnerabilities in H.323 suits (Porter, 2006).

**Signaling initiation protocol (SIP)-specific attacks**

SIP is an unstructured text-based protocol which suffers vulnerabilities according to its encoding format, because it is not possible to check all permutations of SIP messages throughout development for security vulnerabilities. Since SIP protocol links other protocols
and services together, it may cause other typical vulnerabilities in services such as SSL, hypertext transfer protocol (HTTP), simple mail transfer protocol (SMTP) to occur in VoIP environment (Porter, 2006).

**Voice over internet protocol (VoIP) security architecture**

Dantu et al. (2009) visualizes the VoIP infrastructure in three layers: end user equipment, network components, and a gateway to the traditional telephone network (Figure 5). This article defines each of these layers as follows: the first layer is the end user equipment which provides the interface for users to communicate with other end users. It can be hard phones (conventional telephones) or soft phones (software that model the telephone). The securities of this end user equipment are related to the method of its installation. While this end user equipment are mostly used in property networks, at home or in hotels, they have provided with very less security characteristics making them vulnerable to any threats. The second layer is the network components. The existence of any vulnerability in this layer, cause the IP network to adopt the vulnerability. The idea is to measure the security of the network components, before installing the VoIP application, and to install it in the most secure location in between the network components. The IP network components, including routers, switches, and firewalls, must also be VoIP aware to provide security features specific to VoIP.

The third layer is the VoIP gateways. The use of this layer is critical to assimilate IP network with the PSTN. As a result, applying the high security in this layer is necessary to avoid vulnerabilities. The main task of the gateway is voice compression or decompression, signalling control, call routing, and packetization. VoIP gateways interface with external controllers such as session initiation protocol (SIP) proxies, H.434 gatekeepers, media gateway controllers (MGCs), network management systems, and billing systems. These are exactly the places where the malicious attackers can intrude to the system in the way to hijack a user's VoIP subscription and subsequent communications. The security should be implemented in the gateway to avoid these kinds of attacks immediately and powerfully. Security issues encountering in VoIP are rare and, mostly, rather complicated.

**Soft phone security**

The extensive popularity of the VoIP and its low cost brought its attention to so many principal Web service firms. Skype, one of the first firms to capitalize on VoIP, provides a point-to-point Internet telephony network. The Skype communications system is famous because of its wide range of attributes, including free voice and video conferencing. Skype is closed-code software and this can cause many problems for its security features. Yahoo and Google also have the same applications that allow users to make calls for a low cost. Lately, Microsoft unveiled its
Office Communicator 2007, a unified communications client with a VoIP soft phone and Web, audio, and video conferencing. The soft phones are application software installed in computer which can have full access to all information in system, as a result the intruder can benefit from the right of the user who run the soft phone. The user with administrative rights can provide the soft phone application the right to use key system information.

**PSTN-VoIP internetworking**

There are many differences in technical assignments of the VoIP and PSTN, and their internetworking has extensive varying infrastructures and protocol, however in the third layer of the VoIP architecture they need to work together. The differences in protocols, trader operations, the delivery service used, and the service offered cause interoperability problems. These interoperability issues need to be addressed at every interconnection point of the network components. Ong et al. (1999) firstly introduced the signaling transport protocol (SIGTRAN), which was the protocol suite proposed by the internet engineering task force (IETF), as a solution to interoperability issues. This suite allows any subscriber in either network to transparently call a subscriber in another network. Unfortunately, this internetworking makes the infrastructure more vulnerable to attacks. Most current research assumes that the PSTN network is already secure and focuses on securing the IP network. The internetworking of VoIP and PSTN, however, poses new threats to the traditional PSTN network. Thus, system administrators must take great care when plugging security gaps created by the internetworking of VoIP and PSTN.

**Comparison between different VoIP application securities**

### VoIP in skype

Skype is one of the most popular VoIP applications that emphasise mostly in a voice communication in addition to a standard instant messaging such as text messages and file transfers. Skype was developed by Niklas Zennström and Janus Friis was the originators of KaZaa (one of the most popular peer-to-peer services). Skype protects the transferred data by encrypting the media channel (Porter and Gough, 2007; Wang, 2005). One of the main reasons for the popularity of Skype VoIP services is its unique set of features to protect privacy of VoIP calls such as strong encryption, proprietary protocols, unknown codecs, dynamic path selection, and the constant packet rate (Zhu and Fu, 2010). However, some of enterprise security groups consider it as threat because it has to skip firewall in order to make call traffic. It supports call quality when establishing a connection with other Skype user. In addition, Skype got the ability to connect to any phone in PSTN (Porter and Gough, 2007; Höffeld and Binzenhöfer, 2008). The Skype user can communicate with anyone anywhere in the world, with either another Skype client or anyone with a phone.

### Skype architecture

Skype architecture is not a very precise peer-to-peer network as long as it uses a centralized server which helps the system sign up new users as well as authenticates existing users with user ID and password information. There are three main types of computers used within the Skype service: a standard node, a super node, and a Skype server (Figure 6). The standard node is any workstation that has the Skype client software installed. The users are able to make and receive calls, send messages, and use all functionalities of the Skype through this workstation. The super nodes are similar in appearance and functionality to the end user, but these workstations have been chosen by the Skype service to handle much of the Skype system’s work. If a workstation has a publicly addressable IP address and extra bandwidth, it is capable of becoming a super node, and the end user has no control over whether their workstation is a super node or not. These super nodes do the heavy lifting for the Skype service, and the service relies on these super nodes, not a centralized server, for...
keeping track of other users in a directory (known as the
Global index) and data from regular nodes. Workstations
that are behind a firewall or a network address translation
(NAT) gateway will never be eligible to become a super
node, the reason is the IP address of the workstation is
not public (Porter, 2006).
Because the communications including text, voice, and
files may be sent to other terminals before reaching the
anticipated receiver, it is crucial to encrypt these
communications to make sure users whose terminals
have access to this information are not able to detect on
the information that is exchanged. The encrypted session
just begins before messaging starts and two clients have
established that they wish to transfer information between
each other. All data that is sent and received between
two clients is encrypted using 256-bit encryption based
on the advanced encryption standard (AES). The key for
this exchange is unique to that particular sessions and
that particular set of terminals exchanging information.
Once the session has been terminated, the key is no
longer valid (Porter, 2006).
The encrypted data which is transferred during Skype
communication prevents users against transit data
modification, eavesdropping, impersonation, man-in-the-
middle attacks (Skype.com, 2011). It uses proprietary
protocols which provide and incorporate heavy encryption
(Slay and Simon, 2008). There is several security policies
applied in Skype in order to guarantee the security
requirements. These policies as they were addressed by
(Berson, 2005) are:

- The usernames in Skype are unique.
- Users must provide username and password before
  they can get the account privileges.
- Only one session can be established at a time, so the
  same account cannot open twice at a time.
- The end to end Skype session communications are
  encrypted, and without any intermediary node.

**VoIP in Google Talk**

Google Talk is another VoIP applications offered by
Google Inc., which is only available for windows platform,
but it can be used in other platform by using internet
browsers. It provides the basic services like data
communication voice mail, chat, file transfer and etc. The
Google Talk client was a huge step forward for Google as
a means of trying in communications with its wide suite of
applications. Using the widely popular jabber protocol
Google Talk is one of the few open-standard chat
services around. It allows connecting through stand alone
client, Google mail, or many of Google’s other web-based
applications (Baskin and Brashars, 2006). In terms of the
VoIP application, there is not any connection to PSTN
with Google Talk, so it has some limitations which can be
considered as a disadvantage (Ahmed and Shaon,
2009). In addition, it does not provide any kind of
encryption in the end to end communication, which
makes it exhibit for attacks like eavesdropping attacks

**Google talk architecture**

Google has announced that a major goal of the Google
Talk service is interoperability. Google Talk uses Jabber
and extensible messaging and presence protocol (XMPP)
to provide real-time extensible messaging and presence
events, including offline messaging (though only through
non-Google clients like Adium). Google Talk now
supports federation with other Jabber servers, allowing
any one to send and receive IMs to other Jabber users
with non-Google Talk accounts (Hester, 2009). On
December 15th 2005, Google released libjingle, a C++
library to implement Jingle, “a set of extensions to the
IETF’s XMPP for use in VoIP, video and other peer-to-
peer multimedia sessions.” Google Talk does not encrypt
the Jabber stream, instead using an undocumented non
standard way of authenticating to the service, retrieving a
token from a secure web server. Other clients than
Google’s own are required to secure their streams with
transport layer security (TLS) before sending the
password, causing them to stay encrypted throughout the
whole session. Google claims that all messages (text and
voice) will be encrypted in future releases (Hester, 2009).

One of the main reasons to use XMPP channel is for
placing the calls. The user is already authenticated in
GMail and Google Talk. So the request is sent through
the authenticated XMPP network to their XMPP
component responsible to offer SIP gateway. Since then
the component will convert the jingle signaling from the
Google Talk to SIP and use the person’s Google
credentials to authenticate in their SIP services in the
backend and stream RTP with standard codecs directly
from the browser (Camargo, 2010). The Google’s future
plan is to be able to delivery mass market a cheap and
alternative method for calling the old fashioned telephone
numbers. This is possible while Google talk is pre
installed in most of the Android phones in the market,
while they are using the same Jingle extension of XMPP
for their service, (Camargo, 2010). Figure 7 shows the
architecture of Google Talk security.

**Skype versus Google Talk**

Garfinkel (2005) argued that “security is not some
abstract quality that can be analyzed in isolation”. As a
result to measure the security of the VoIP applications, it
is essential to study the particular threats and also to
verify if the design or operation of that VoIP application is
Table 1. Comparison table.

<table>
<thead>
<tr>
<th>Security feature (Conversation)</th>
<th>Skype</th>
<th>Google talk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Privacy</td>
<td>The broadcasted Messages are encrypted from Skype-end to Skype-end. In case of any intermediary node, the meaning of the massages will remain secret, and no one can reach to it.</td>
<td>In Google Talk the encryption is not end to end making the entire voice data susceptible of eavesdrop attacks however the connection between the client and the Google Talk server is encrypted. Mail notifications, Friends list, personal settings and messages of others between a client and the Google server can't be seen by others.</td>
</tr>
<tr>
<td>User names are distinctive while users or applications have to enter a Skype username and its selected authentication code such as password. Only after the authentication the session will permit transmit messages such as voice, video, files, or text.</td>
<td>User names are distinctive but in Google Talk the user can use the same authentication password and username for Gmail and Google Talk, and other Google applications. Also it supports inter-operability; the user in Google Talk can communicate with the users in other services.</td>
<td></td>
</tr>
<tr>
<td>Authentication</td>
<td>Since the Skype requires the authentication password and username, whole its network may stop to work if its authentication servers break down or become unavailable.</td>
<td>Since Google Talk also requires the authentication password and username, whole its network may stop to work if its authentication servers break down or become unavailable.</td>
</tr>
<tr>
<td>Availability</td>
<td>Skype's authentication servers can't survive network disruptions or attacks</td>
<td>Google Talk's authentication servers also can't survive network disruptions or attacks</td>
</tr>
<tr>
<td>Survivability</td>
<td>It is generally very resilient to local network disruption</td>
<td>Google Talk is also very resilient to local network disruption</td>
</tr>
<tr>
<td>Resilience (Conversation)</td>
<td>There is no guarantee that the files will be delivered completely as transmitted.</td>
<td>There is also no guarantee that the files will be delivered completely as transmitted, but the risk is higher.</td>
</tr>
<tr>
<td>Integritiy</td>
<td>Skype uses a high limit of bandwidth; it may cause security vulnerabilities to other parts of the system, for example it could be an infection vector for spyware. It has no built in anti-virus protection.</td>
<td>Google Talk uses a less limit of bandwidth; Google Talk may cause vulnerabilities to the system due to remote HTML-injection vulnerability because it has also no built in anti-virus protection.</td>
</tr>
</tbody>
</table>

secured from those threats. Garfinkel (2005) also introduced security features such as privacy, authenticity, availability, survivability, resilience, conversation integrity and system integrity as important keys to be measured in VoIP applications. Table 1 demonstrates the comparison between the Skype and Google Talk as two most popular...
types of VoIP applications as mentioned in the security features above. Skype has been improved its peer-to-peer voice service for the last seven years, while Google voice is started recently well-built. Google’s communications service may offer a different set of tools, but its rich features and brand reputation are derived the attention of the VoIP enterprises. Skype is sure to feel the pressure of the competition after Google has introduced its new adventure by integrating Google Voice with Gmail in-box. Although Google Talk is in low-development stages, many users are connected to the service because they like the simple way for chatting provided by Google Talk. In general Google Talk is more enterprise-oriented than Skype, requires less bandwidth and adopt collaborative features that work well with the implementation of Google Applications in business communication services and integrates with the company’s active directory through third-party tools, to name just a few advantages, however it has more place to improve its security and does not provide end to end voice encryption making the whole voice data susceptible to eavesdrop attacks. Skype, on the other extreme, provides interesting features that Google Talk does not have at present, for instance the ability to establish a conference call with up to five people at a time and the ability to make phone calls to mobiles and landlines worldwide at low rates. However, this software requires more bandwidth than Google Talk and it is not favorable to be used as a means of communication in business communication services. In contrast to Google Talk which is only available for Windows, Skype is now available for Widows, Linux, Mac OS, Widows mobile and some Nokia series platform. However, the main problem with Skype is its proprietary protocol, which bars users to send and receive calls to other VoIP services (Ahmed and Shaon, 2009).

CONCLUSION AND FUTURE WORK

This article discusses the VoIP application security. The previous studies have shown that using VoIP increases the vulnerability of the network, as a result so many efforts has been done to overcome this problem, and to make it as less as possible. However, the low cost of the VoIP encouraged enterprises to produce many different applications that facilitate this technology, in distinct characteristics. As a result, security issues should be one of the most important things that should be taken seriously during selecting the appropriate VoIP application to be used. In this article, we studied the architecture of the VoIP and its protocols, and the security issues regarding to each level, to come with a good understanding of the security condition of the VoIP applications in the market. Finally, two different models of the VoIP application, Skype and Google Talk have been compared together, to find out which application is more reliable in terms of security. The result of comparison shows that Google Talk is more enterprise-oriented than Skype and it is open to improve its security. Skype, on the other hand, provides interesting features that Google Talk does not have at present, for instance the ability to establish a conference call with up to five people at a time and the ability to make the phone calls to mobiles and landlines worldwide at low rates. However, this software requires more bandwidth than Google Talk and it is not favorable to be used as a means of communication in business communication services. In summary, we can see that Skype is more secure in VoIP application but it does not offer interoperability. On the other hand, Google Talk has interoperability but with less level of security. It is one of the areas of the future study which fills the gap between the interoperability and security and designs the new VoIP application which offers interoperability without side tracking the security. In addition to that, security enhancement is an important factor to be considered such as a built-in antivirus protection, VoIP-aware firewalls and VoIP anomaly detection systems. In a future work, the authors might also want to consider about improving the state of the message encryption to prevent the call hijacking and eavesdropping attacks.

REFERENCES