Voice Over IP And Forensics: A Review of Recent Australian Work

Jill Slay[12], Matthew Simon[1], David Irwin[1],

University of South Australia, Mawson Lakes, SA 5095, AUSTRALIA

Polytechnic of Namibia, Windhoek, NAMIBIA.

([jill.slay][matthew.simon][david.irwin]@unisa.edu.au)

Abstract

The popularity of Voice over the Internet Protocol (VoIP) for providing voice communication over IP networks such as the Internet has resulted in VoIP becoming a global telephony service. VoIP applications convert analogue voice signals into a digital format, which is then encapsulated into IP packets for transmission over the Internet. Our research has examined both the security and privacy implications of widespread adoption of VoIP for personal and business telecommunications and also the use of VoIP calls by criminals, as many implementations of VoIP may also use strong encryption to secure both the voice payload as well as to control messages. We have also considered the implication of recovering electronic evidence and information from VoIP since conventional methods of eavesdropping and wire-tapping do not apply to VoIP calls.

This paper provides an overview of the development of investigative processes and forensic tools to enable law enforcement to engage the digital forensic process in a VoIP environment

Introduction

VoIP technology has radically changed the way voice data is communicated, and revolutionised the Australian and International Telecommunications industry. With the growth in popularity and speed of the Internet, this technology emerged rapidly, allowing phone calls to be sent via Internet infrastructure rather than the traditional Public Switched Telephone Network (PSTN). There are many advantages to using VoIP technology instead of the older PSTN system. The primary benefit is cheaper call costs for local, long distance and international calls. VoIP is also an advantage in terms of regional and remote users since it avoids large-scale roll-out of cable and cut costs in large organisations with extensive internal phone systems.

Our original interest in this topic was developed after the Voice over IP Security Alliance (VoIPSA) (Kuhn, Walsh & Fries 2005) released a detailed review of threats faced by VoIP technology. The most serious of the threats identified were denial of service, host and protocol vulnerability exploits, surveillance of calls, hijacking of calls, identity theft of users, eavesdropping and the insertion, deletion and modification of audio streams .

Our work started with a successful pilot study (Simon & Slay 2006) in which we began to examine the potential threat to privacy by the capture and reassembly of VoIP packets by a hacker (or other criminal or terrorist) from a computer or network after a VoIP conversation has taken place. We have applied memory forensics to address some of the concerns with the use of VoIP. The results of the limited number of experiments conducted with one particular implementation of VoIP (SIP) on one specific operating system (Windows XP SP 2) showed that it was possible to recover packets from memory after the completion of a VoIP call. Although very few packets remained in memory in

our pilot study, there was enough evidence with these few packets to prove that a call has actually been placed and between whom.

At the time, we found little other published research in this area of IT Security / Forensic Computing. Neumann, Tillwick, & Olivier (2006) had explored the information exchanged in VoIP call control messages and the implications this has on personal privacy. Chen, Wang & Jajodia (2006) examined the privacy and security aspects of peer-to-peer (P2P) VoIP calls and show how the use of VoIP has substantially shifted the previous balance between privacy and security that exists in traditional PSTN calls. This paper shows the development of our work at the University of South Australia beyond hat identified at our commencement in 2005-6.

In the following sections, we deal with Pilot Study, Memory Forensics, Investigations, Looking Further and Deeper into VoIP and finally the Development of a Software Tool.

Pilot Study

Our original work (Simon & Slay, 2008) was an examination of the privacy implications of VoIP. The result of this work was that we were able to search for packets in a memory image taken from a computer running Windows XP or Mac OS 10 and sort, order and output the packets and eventually, under specific circumstances reconstitute the packets to an audio file (audio could be heard).

We showed that potentially a user's privacy might be breached by a hacker using software similar to that which we developed as part of this original research and that a hacker would be able to copy remnants of the data stream from the computer's volatile memory and recreate small portions of the conversation. We also showed the corollary: it was possible to extend the software developed as a tool for users to mitigate against the threat of a hacker who was able to breach their physical and logical security and extract remnant VoIP packets from memory space (in specific situations). This also meant that we had also developed our own first VoIP forensic tool

A major problem with this research was the inability to verify the resulting memory images obtained during the memory imaging process. It was difficult to ascertain how much of a change the process of acquiring the image makes to the image itself. Investigation into the amount of change that is caused by this factor was difficult. Using special-purpose memory image hardware was the only foreseeable method of conducting such an investigation. It may be possible to conduct a 'before and after' experiment that compares the hardware-acquired memory state before and after the process of software-based memory imaging specified in this research. The development of memory forensic techniques since 2005-6 has made our research considerably easier and we spent some time, slightly distracted, examining memory forensics in a broader context than VoiP.

We saw a need to examine other operating systems, implementations of VoiP including encrypted forms and then later began to consider the effect of the widespread use of VoIP on law enforcement and police analysis and investigations.

Recovery of Remnant Information from Memory

We extended our work on VoIP in 2009 (Smon & Slay, 2009) This research presented a study into the feasibility of recovering remnant information from the physical memory of a target computer about a particular IM setup. The setup emulated in this study, was Google Chat used through the Pidgin-Portable client while employing Tor to add a layer of encryption and anonymity to the entire process.

The primary objective of the experiment was to assess if remnant data does remain in the physical memory of the system after use of the target communication technology. It also aimed to identify some of the influencing conditions that affect the outcome so as to define patterns in the data that was recovered.

The results showed that remnant artefacts do remain in the physical memory during and after the execution of the target communication technology in the context tested. This indicates that physical memory forensics has potential for use in recovering information about the technologies tested in this research.

While the results of this research confirmed that the recovery of communication technology artefacts was feasible, it also helped us understand the limitations of our research. The results showed that the information is not necessarily recoverable in all situations – the termination of the relevant processes was highlighted as one case where data was often subsequently unavailable.

Law enforcement Investigations

We noted in 2011 (Simon & Slay, 2011) that investigation of communication technologies is an important activity that law enforcement agencies carry out. We noted that over the extensive lifespan of traditional communication technologies, methodologies have been built that support the acquisition of information in a legally sound and rigorous manner. In obtaining information about the use of traditional communication technologies, police use a combination of communication interception, access of stored information and use of post-mortem analysis. Two major factors support these methods, legislation and the nature of the technology. The legislation is effective in allowing law enforcement to carry out certain activities but also in forcing service providers to operate in certain ways that support law enforcement (e.g. collecting certain types of information). The 'nature of the technology' is more nuanced. It effectively relates to the low level of control the user has over the technology that largely prevents the users from circumventing law enforcement methods of obtaining information.

We also saw that the methodologies used for investigation of traditional communication technologies are well suited to their purpose. As the target technologies have existed for many years and have evolved little over this time, the methods have evolved and become sound and rigorous. However, they have also become specific to their intended use. This means they are inflexible and do not cope well with altered parameters. Specifically, Internet based communication services are very different to traditional communication technologies and methodologies that have developed for the investigation of traditional communication technologies may not be effective against contemporary communication technologies in many cases.

We argued that law enforcement investigation methods where the carriage service is an Internet application are ineffective in many instances. When carrying out communications interception, the provider has no legal obligation to assist law enforcement. The use of interception at the carrier level circumvents the carriage service provider but the service decoupling property will mean that communications may be missed, and the use of encryption will render this approach ineffective. Another layer of complexity is added by the lower barrier to entry and the borderless supply properties that affect both communication interception and access of stored information. The carriage service provider could potentially be anyone located anywhere in the world. This may make even attempting to interface with the provider very difficult, let alone accessing the required information. These properties potentially allow the provision of services for the direct purpose of secure communication that cannot be intercepted or recovered by third parties; this is contrary to traditional communication services where such a service would be illegal. Even users with low technical skills could employ these services to utilise secure communication greatly increasing the potential user-base.

The increasing complexity of end-user devices is another change that works against law enforcement methodologies. Many devices now have built in encryption that can be easily activated by even technically low-skilled users. This will prevent the use of post-mortem analysis to recover any information stored on the device. More highly skilled users can employ advanced techniques such as data obfuscation or plausible deniable encryption to add a layer of complexity to an investigation.

The legislation supporting current methodologies and the inherent nature of traditional communication technologies provides an almost 'ideal' situation for law enforcement. However, contemporary communication technologies may always provide a challenge due to its very nature. Even with a widened scope of current laws to incorporate Internet-based carriage services in the same way as traditional carriage services, the lower barrier to entry and the borderless supply properties mean that such laws may be very difficult to enforce.

Looking further and deeper into VoIP for Electronic Evidence

As a result of the preceding work we decided to explore further and argued that since:

1. VoIP is global telephony service, in which it is difficult to verify the user's personal identification,
2. the security of placing such calls may also be appealing to criminals, as many implementations use strong encryption to secure both the voice payload as well as to control messages, and
3. monitoring or tracing such VoIP calls is difficult since conventional methods such as wire-tapping is not applicable to VoIP calls.

Therefore, other methods of recovering evidence and information from voice over IP protocol are required and It is essential that forensic computing researchers devise methods to allow law enforcement agencies to overcome some of the aspects of this method of telephony that are advantageous to criminals.

We made the point that VoIP is a telephony service that uses UDP/IP protocols which is one of the foundations of the Internet (James, 2005) using packet switching technology. Since packet switched networks differ from circuit switched networks as a set bandwidth is not reserved for an individual connection, instead, information data are carried by IP packets (each packet carries up to 1500 bytes information data) and each IP packet is forwarded individually from source to destination through a series of routers using the IP addresses (including both of source and destination IP addresses) across the Internet. The call setup and termination in VoIP is implemented by service provider, usually called a VoIP gateway. When a user attempts to make a VoIP call, he/she sends a call request, which contains both calling and called parties information, to VoIP gateway. Upon receiving the request, the VoIP gateway assigns two call IDs, one for the calling party and one for called party, respectively, in which each call ID consists of the IP addresses of the end-user and VoIP gateway, and the corresponding user ID associated with an authentication key for the communications between the end-user and VoIP gateway. Both call IDs are registered in the database located in the VoIP gateway as a pair of virtual VoIP telephony connection.

We noted that an end-user only has one call ID to the VoIP gateway but not the call ID from VoIP gateway to the other end-user. During the call, voice data packets are forward from one end-user to VoIP gateway using its call ID, and then the VoIP gateway uses the user ID to search the other end-user from its database and forwards the voice data packets to the other end-user using the corresponding call ID registered in the database. When a VoIP call is completed, both call IDs are removed from the database. Firstly, due to the global nature of the Internet, the VoIP can be located in anywhere around the world. Secondly, the call IDs are only used during connection period and most calls are randomly occurred. These factors have made the trace and monitor of VoIP calls even

more difficult. However, from a commercial business point of view, most VoIP service providers require their clients, such as IP phone card holders or IP end devices, including software running on a computer or hardware IP phone, be registered in order to obtain a user ID and associate authentication key before using the VoIP service. Also, the VoIP gateway keeps a record including user ID, call transaction number, time that calls are made and call duration for each registered client for billing purpose.

In practice, VoIP telephony relies upon various protocols and methods to establish calls and transmit data. Skype software uses proprietary protocols that are also encrypted. Many VoIP implementations, however, use SIP and RTP. The SIP protocol is used for call initiation, call teardown and other call related data sent during the conversation. This is analogous to the PSTN in-band signalling mechanism, Dual Tone Multi-Frequency (DTMF). SIP is a text based application level protocol and relies heavily on other protocols for transport (such as IP and UDP) (Davidson & Peters, 2000; Ahuja & Ensor, 2004; Sicker & Lookabaugh, 2004). VoIP implementations that use SIP generally rely on a SIP proxy server to which the users must authenticate their login credentials. This proxy is also used to route call and signalling data. Clients can find each other and forward SIP messages via this proxy (Ahuja & Ensor 2004). Communications using SIP are not only used for initiating and to teardown calls, they are also used for changing call parameters or other features such as integrating more callers into a conference session. shows the exchange of SIP messages between two clients. It shows the use of SIP registrars and the proxies which allows callers to create new call sessions. SIP registrars are additional servers used to locate other users; generally, the SIP proxy also acts as the registrar.

At transport layer, VoIP is using UDP protocol, which is a connectionless protocol without providing a mechanism to ensure that data packets are delivered in sequential order. In this case, VoIP implementations face problems due to network latency and jitter. This is especially true when satellite circuits are involved, due to long round trip propagation delay ranging from 400 milliseconds to 600 milliseconds for geostationary satellite. Furthermore, IP packets are lost or delayed at any point in the network between VoIP end-users, there will be a momentary drop-out of voice. This is more noticeable in highly congested networks and/or where there are long distances and/or interworking between end points. It has been suggested that it is necessary rely on the packetized nature of media in VoIP communications and transmit the stream of packets from the source phone to the destination phone simultaneously across different routes multi-path routing. In such a way, temporary failures have less impact on the communication quality. In capillary routing, it has been suggested to use at the packet level Fountain codes or particularly raptor codes for transmitting extra redundant packets making the communication more reliable. Therefore, the receiving node must restructure IP packets that may be out of order due to delayed, delay jitter or packet loss while ensuring that the audio stream maintains a proper time consistency. This functionality is usually accomplished by means of a play-out buffer to temporally store 800 - 1000 milliseconds of VoIP data for the purpose of restructuring the VoIP packets in sequence order and digital voice decoding.

Our more recent research is based on analysis of the very limited amount of VoIP data stored in the play-out buffer located in IP telephone devices. Examples of this kind of telephone devices include both hardware IP phone and software IP phone end-device, i.e., VoIP software running on computer, such as that implemented in Skype . The play-out buffer usually contains of 800 – 1000 milliseconds of VoIP data, which are packetized into variable length (up to 1500 bytes) UDP/IP packet format. This buffer can be used to supply electronic evidence for forensic investigation or intelligence.

Another important feature is routing VoIP traffic through firewalls and address translators. Private Session Border Controllers are used along with firewalls to enable VoIP calls to and from a protected enterprise network by traversing symmetric NATs (network address translators) and firewalls. For

example, Skype is a software program created by the Swedish and Danish entrepreneurs Niklas Zennström and Janus Friis, which allows users to make telephone calls over the Internet to other Skype users free of charge, or to landlines and cell phones for a fee. Additional features include instant messaging, file transfer, short message service, video conferencing and its ability to circumvent firewalls. These Private Session Border Controllers also contain data that might be extracted using a suitably developed software tool to extract electronic evidence, since VoIP users through these Private Session Border Controllers must be registered with the VoIP gateways for billing purposes.

Development of Digital Forensic Tool for extraction of VoIP data

In 2011, Irwin and Slay extended the previous work by the development of a forensic tool with the following functionality:

1. Reconstruction of the VoIP data packet sequence.
    - Detection of the VoIP protocol being used to packetize the voice data.
    - Reconstruct the VoIP data packet sequence order using packet sequence number.
2. Interactive searching functionality.
    - The development of an interactive user interface to help the tool user search for forensic evidence within the captured RAM effectively.
    - Automatically record the search results in a format designed to fit the requirements of a court of law.

3. Analysis of the voice data.

    - Discovery of the VoIP phone application being used, VoIP codec and routing protocols.
    - Identify VoIP audio in RAM using both a statistical searching algorithm to first identify the language type (English, Japanese etc) based on the characteristics of various types of languages and also spectral analysis if enough VoIP voice data is recovered.

4. Trace and tracking information of the VoIP end users.

    - The recovery of VoIP application user names from the control signalling information used to initiate a VoIP call between two parties.
    - The use of a SIP phone as the VoIP application will require both end users to be registered with a SIP provider. These registration details may be recovered at a later date if the unique SIP phone number is extracted from the call setup information.

This research has resulted in the development of a software tool which successfully identifies and extracts VoIP packets based on the Ethernet protocol, the Internet protocol, the User Datagram protocol and Real-Time transport protocol (where implemented) and the payload, the voice component, which is usually encrypted.

The recovered packets may be stored permanently in a database allowing further analysis to be performed upon them, either formatting for legal requirements such as court prosecutions or applying decryption techniques to the payload. This software is not intended to be used for searching huge amounts of physical memory such as hard drives but is more efficient at VoIP packet recovery when applied to RAM captured memory.

In addition to packet recovery, VoIP application control signals used in the process of establishing, maintaining and ending a call may also be captured as well information unique to each VoIP application, such as the way in which contact lists are stored.

Future Work

Performing RAM forensics has successfully demonstrated the ability to recover VoIP protocol artefacts left behind in RAM after a VoIP call has taken place. Having successfully recovered packet sequence information to allow VoIP payloads to be reconstructed correctly, the next phase of the research is threefold:

- Search RAM for unencrypted audio, if it exists.
- Is VoIP injection possible? To perform track and trace of the end-user, the individual at the receiving end of the VoIP call. The recovered IP addresses may not necessarily include the end-user but a node within the network path between the calling and receiving party.
- Extend the research to include VoIP applications on portable devices, primarily, mobile phones.
- Develop a database of contact list structures and control signal information for the most common VoIP applications.

Note that in the process of RAM searches, unencrypted Skype chat logs were recovered. This is beneficial bonus since most VoIP applications also include the element of sending instant messages.

References

Ahuja, SR & Ensor, R 2004, 'VoIP: What is it Good for?' *Queue*, vol. 2, no. 6, September, 2004, pp. 48 - 55.

Davidson, J & Peters, J 2000, *Voice over IP Fundamentals, Cisco Systems*, USA.

Chen, S., Wang, XY. & Jajodia, S. 2006. On the anonymity and traceability of peer-to-peer VoIP calls. IEEE Network 20(5): 32-37.

Irwin, D & Slay, J. 2011.‘ A Digital Forensic Tool For The Extraction Of Electronic Evidence From The Voice Over Internet Protocol’ Advances in Digital Forensics VI. Volume 337 of Advances in Information and Communication Technology; Springer-Verlag GmbH, Boston (accepted 20/1/2011)

James, M 2005, The Internet Telephone: Voice over Internet Protocol (VoIP), updated February 9, 2005, Commonwealth, viewed 20 March 2006.

Kuhn, DR, Walsh, TJ & Fries, S 2005, *Security Considerations for Voice over IP Systems*, National Institute of Standards and Technology, Gaithersburg.

Neumann, H Tillwick , H., & Olivier, MS., 2005. ‘Enhancements to SIP to prevent abuse of Voice-over-IP services. In: Southern African Telecommunication Networks and Applications Conference (SATNAC) Proceedings. (2005)

Simon, M & Slay, J. 2011. ‘ Investigating Modern Communication Technologies: The effect of Internet-based Communication Technologies on the Investigation Process. Journal of Digital

Forensics Security & Law. Vol6n4,

Simon, M & Slay, J. 2011. ' Recovery of Pidgin Chat Communication Artefacts from Physical Memory: A Pilot Test to Determine Feasibility'. ARES 2011.

Simon, M., & **Slay, J**., 2010. What Are You Looking For: Identification of Remnant Communication Artefacts in Physical Memory! International Cyber Resilience Conference, ECU, Perth August 2010.

Simon, M., & **Slay, J.,** 2010 'Recovery of Skype Application Activity Data from Physical Memory'. ARES 2010 - 5th International Conference on Availability, Reliability, and Security 2010, , pp.283-288. (CORE B).
Simon, M., & **Slay, J.,** 2009. 'Enhancement of Forensic Computing Investigations through Memory Forensic Techniques'. WSDF 2009, Fukuoka, Japan. March. (CORE B).

Simon, M & **Slay, J** 2008, 'Voice over IP Forensics'. E-Forensics, Adelaide University, January 21, 2008

Sicker, DC & Lookabaugh, T 2004, 'VoIP Security: Not an Afterthought ', Queue vol. 2, no. 6, pp. 56-64.

Slay, J., & Simon, M., 2009. Voice over IP: Privacy and Forensic Implications. *International Journal of Digital Crime and Forensics.* 1(1).