Voice over IP Forensic Approaches: A Review

Hussein Al-Saadawi
Software Engineering Department, Technology Faculty, Firat University
Elazig, Turkey
husseintaeyb@gmail.com

Asaf Varol
Software Engineering Department, Technology Faculty, Firat University
Elazig, Turkey
varol.asaf@gmail.com

Abstract—Voice over Internet Protocol (VoIP) is one of the highly used applications for voice and data communication purposes. The convenience of using VoIP applications, its acceptable quality, and the relatively low cost of usage have made it a potential replacement for landline and cellular communication. Along with these advantages, the vulnerability of using VoIP is usually considered the most cumbersome con. Various researchers have investigated VoIP risk issues; yet, a very limited research has approached VoIP from digital forensic perspective. A survey of the prominent research proposals and methods that presented VoIP forensic within their context is provided in this paper. In addition, a contemporary discussion on trends for improving VoIP forensic is given in this paper to formulate directions for new and useful research in the area of VoIP forensic.

Keywords—VoIP; digital forensic; VoIP protocols; VoIP vulnerabilities; research trends

I. INTRODUCTION

Communications and ways of contact between human beings are considered as elementary needs in our life. Many communication methods have been assisting in the growth of business popularity by exploiting the benefits of on-time and off-time delivery of voice, message, video, and image. Since the invention of Public Switched Telephone Network (PSTN), voice communication has been widely affecting the evolution technology. Popular development throughout voice-driven technologies via computer, smartphone, and network have become an important part of communication systems. In 1998, the conceptual realization of Voice over Internet Protocol (VoIP) produced a new equipment unit’s hardware and software programs which have allowed the communications between IP networks and PSTN and vice-versa. This had incremented voice traffic usage in the U.S. alone to approximately %2 in less than one year [1]. By the year 2000, the deployment of VoIP as real commercial services took place and it has been widely used ever since. After VoIP was born, the shrinking in PSTN utilization has become obvious in the market [2]. Generally, VoIP is perceived by offering different functionality, dynamic demands aspect, and cheaper rates than PSTN services. This has been reflected on the progressive popularity VoIP had achieved in its early days [3, 4]. This is, in fact, more useful for small size enterprises as it allows them to compete with a minimal cost [4].

Despite all the advantages offered by VoIP, like all IP-based services, the main disadvantage of using VoIP is represented by the generated risk of abusive and harmful applications as well as the attacks initiated by scammers, blackmailers, and cyber-terrorists. The vulnerabilities initiated on the Web are usually affecting VoIP as well and the risk may possibly compromise sensitive or private information related to organizations and/or individuals. In order to alleviate the risk of security issues and optimally respond to the cyber-criminal activities targeting VoIP services, various proposal and solutions have been offered and presented [5]. Essentially, the proposed solutions rely on the use of encryption methods, antimalware programs, and applications, and abiding with recommended security measures and ethics. However, in the case of criminal activity incidents, there should be a proper way to respond in order to backtrack the attackers. The best answer for this need is to investigate and detect legal evidence that can be used in the court of law. This type of procedure is referred to by VoIP Forensic. VoIP Forensic implies the use of digital evidence which is generated by the conducted VoIP transaction that resides somewhere on electronic components of computing devices.

Nevertheless, investigating digital evidence generated by VoIP is not a straightforward and easy task to do as VoIP’s generated traffic uses different types of protocols as well as data encryption techniques [6-10]. This challenging task has been studied and investigated aiming to reach a procedure, software, tool or application that assists in backtracking digital evidence of VoIP traffic or, more formally, a possible VoIP forensic tool. Although these studies unveiled numerous benefits and practical ways assist in VoIP forensic; yet, up to the authors’ knowledge, there is no formal software, application or tool that is specialized in VoIP forensic and can be used for legal evidence extraction and report on a formal basis. Thus, it is justified to review the prominent methods proposed and discussed in the literature related to VoIP forensic in order to classify the used methods, compare the performance and conclude where the gaps reside for further development.

In this paper, a review of the prominent methods and approaches developed and/or applied for VoIP forensic is presented. Beyond the narrated introduction in Section I, the paper is organized as follows: The essential background on VoIP including its mechanism, protocols, as well as the needed information on digital forensic and its importance for VoIP

978-1-5090-5835-8/17/$31.00 ©2017 IEEE
application is provided in Section II. The core element of this paper, that is, a survey of the prominent methods for VoIP forensic is elaborated, investigated and discussed in Section III. The final remarks, conclusions, recommendations and possible directions for enhancing VoIP forensic are summarized and provided in Section IV.

II. BACKGROUND

A. VoIP: The Concept

Voice communication conducted using the Internet or IP network is called Voice over Internet Protocol (VoIP). This application is alternatively known IP telephony or broadband phone [11]. The idea was initiated when the communication over the Internet became possible in the form of email; which was quickly developed to real-time messaging in the form of on-line chatting by texts. Not much later, the idea of online messaging or Instant Messaging (IM) went further to support voice and video communication which was the initiative of VoIP generation [12]. Basically, VoIP is transferring voice data packet over the Internet based on IP address (from one point to another). Traditional telephony, PSTN, employs switching circuitries in order to transfer the voice from one point to another; thus, the signals are arranged over channels throughout the switching network. However, in IP networks, data must be digitized and transmitted packets format. These packets recognize the destination, may be rerouted based on the certain protocols and the network traffic. The main keyword that is always appended to VoIP is the Internet Protocol (IP) which made it VoIP possible [12]. Apart from IP, VoIP needs various protocols, which are discussed in the following subsection, to from the necessary signaling, establishment, maintenance, and termination of the connection between the two points in real-time across the network. In addition, there are other protocols for monitoring the Quality of a Service (QoS), booking resource, etc.

VoIP requires a sort of proper conversion to transfer the spoken words from their audio nature to digital entities. In fact, the conversion here is done using the common sampling, quantization and digitization processes followed for converting the analog voice to digital sequences. However, this method is used for converting and storing the voice where some compression techniques are used to reduce the size of the consumed memory. In the case of VoIP, special compressing/decompressing (CODEC) algorithm is used to provide a real-time compressing/decompressing in order to shrink the needed transmission bandwidth [13]. The whole mechanism is provided in Fig. 1 for the ultimate benefit of the reader. As shown in Fig. 1, the packetized data is transmitted over IP network, mostly the Internet, and for ensuring the transmission process, there is a number of protocols to be used for this purpose, presented in the next sub-section, as well as different gateways to leverage the transform between networks (IP-based, Analog Telephone Adaptor (ATA), PSTN, Private Branch Exchange (PBX), etc.) [14].

In fact, the previous overview is necessary to understand how and what could be traced and/or extract to assist in digital forensic procedure while attempting to formulate or develop a VoIP forensic tool. This fact is obviously reflected by the explained mechanism and it will be further elaborated in section III.

B. VoIP Protocols

As mentioned earlier, VoIP, as network-based service, needs certain protocols in order to establish the proper connection over the network as well as ensuring smooth communication with other telephony services such as PSTN, PBX, etc. In TABLE I, the most common VoIP-related protocols are listed along with a brief of their purposes.

C. Digital Forensic: The Concept

The concept of digital forensic is essentially based on determining, extracting, analyzing, presenting and reporting digital evidence in a way acceptable to the court of law [15]. The digital evidence here refers to any movement of electronic data that can be used in identifying any useful information for the legal case; this data could be of any form but in VoIP, it would be mostly related to either data packets or the protocols associated with VoIP. The process of VoIP forensic can be categorized to the following tenets [15]:

- **Identification**: This step is taken in order to determine which items (computing items) that are considered as part of the crime scene. Generally, any electronic device that has the ability to store digital data temporary or permanently is considered a target for this step.

- **Acquisition**: This step is taken right after the identification where various tools and methods are used to forensically acquire digital evidence resides on identified components. It is important here to mention that it may be necessary in some cases to develop certain algorithms or filtering methods in order to extract the right data related to the needed forensic analysis.

- **Preservation**: This step is taken in order to ensure that valid copies of the extracted data are made and properly saved in a read-only format in order to guarantee no data manipulation will occur.
### Table I. VoIP-Related Protocols

<table>
<thead>
<tr>
<th>Protocol Name</th>
<th>Definition and Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>H.323</td>
<td>A protocol suite that is designed to provide a mechanism for transferring media content (audio, video, and data) over the IP networks [16]. This protocol suite is considered as the core set for VoIP due to the major signaling responsibilities which it takes care of.</td>
</tr>
<tr>
<td>MGCP</td>
<td>Media Gate Control Protocol – is a supporting protocol that can be used as an alternative of H.323 in functionalities.</td>
</tr>
<tr>
<td>RTP</td>
<td>Real-Time Transport Protocol – is a protocol used by both SIP and H.323 in order to maintain media transfer over IP network. The main duties accomplished by this protocol are: detecting packet loss, dynamic media adjustment to compensate for BW issues, media frame timing, source identification for multicasting sessions, and jitter compensation [19].</td>
</tr>
<tr>
<td>RTCP</td>
<td>Real-time Transfer Control Protocol – is a control protocol works in conjunction with RTP to obtain the needed feedback on the QoS, inter-media synchronization, identification, and session control.</td>
</tr>
<tr>
<td>RTSP</td>
<td>Real-Time Streaming Protocol – is a client-server protocol that is used for controlling real-time media delivery control.</td>
</tr>
<tr>
<td>RSVP</td>
<td>Resource Reservation Protocol – is a protocol used for creating a sort of circuit-switching network for IP network service to alleviate the issue of delay.</td>
</tr>
<tr>
<td>SDP</td>
<td>Session Description Protocol – as its name indicates, it provides an information on the conducted VoIP session such as beginning and end, name and purpose, used BW, etc.</td>
</tr>
<tr>
<td>SAP</td>
<td>Session Announcement Protocol – is a multicast specialized protocol used for advertising the session by sending an announcing packet to corresponding port and address.</td>
</tr>
<tr>
<td>Skinny</td>
<td>A proprietary protocol designed by Cisco for skinny system clients in order to provide the needed communication with H.323 systems clients [18].</td>
</tr>
</tbody>
</table>

- **Analysis and Examination**: This step is taken to regenerate the needed useful information that can be used as a legal evidence just ambiguous digital evidence. This step requires special tools and programs that vary according to the application. In some cases, it may be needed to use more than one tool to obtain the needed information. Here, understanding the background and the concept of the application under which the forensic is being conducted is essential.

- **Presentation**: The final step in digital forensic is generally creating a readable report that can be used in the court of law. The produced report shall not be technical and hard to be understood by non-specialized people, rather, it should be in a sort of actual presentation of the followed methods, the obtained evidence and final conclusion on the related legal case.

Digital forensic is generally applied for all computing application cases. Here, the applicability, difficulty, and benefits range based on the orientation of the application, digital evidence, and the available tools. Tracing an attack on VoIP application needs a deep knowledge in networking concept, VoIP protocols, related computer memory and the needed and possible analysis for each of them. There is a number of digital forensic applications available for acquiring, analyzing, extracting and reporting evidence from various types of memory; yet, it is not the case of VoIP. Thus, reviewing the available methods and attempts developed for VoIP forensic is required academic task.

### III. VoIP Forensic Formulating Attempts

VoIP is affected by general threats that may affect other Internet services, especially voice and media communication service. However, some of the vulnerabilities can be generated due to the lack of a real physical connection with the caller or the true identity of the contacting party. This kind of threats is generally called social threats. There could be other types of attacks such as the Denial-of-Service (DoS), interception, service abuse, and eavesdropping [14, 20]. In order to forensically trace such events, various research attempts have been developed. In the following, a review of the most prominent and relevant VoIP forensic attempts:

Leong and Chan took a lead on contemplating and developing and idea based on Skype forensic [21]. They set up the needed Skype event path and they formulated the needed framework for evidence extraction and analysis. They succeeded in detecting sockets associated with active Skype and provided recommended steps for further analysis. This work is considered as one of the early attempts in the field of VoIP forensic. The attempt could provide some answers to whether it is possible to forensically analyze and encrypted peer-to-peer communication and how could be the future development and needed effort. However, the work failed in providing a complete VoIP forensic tools to be used in a broader sense.

Pelaez and Fernandez [22] had presented an attempt to formulate a method for VoIP network forensic. Their idea was essentially based on creating an applicable framework for VoIP forensic. In their work, they proposed VoIP collector design pattern for evidence collection and VoIP evidence analyze pattern for evidence analysis. These two patterns combined were called VoIP forensic pattern and presented in UML for further development. This work was generally useful as pioneer attempt in the field; however, it did not actually offer...
any practical implementation nor did it provided and experimental testing. In fact, the idea of this research attempt was initiated from a previous research proposed by Fernandez et al. [23].

Khan et al. [24] approached VoIP forensic from an important aspect that is, detecting caller identity from an encrypted traffic. The idea of the attempt is interesting and so is the implementation where they used Artificial Intelligence (AI) in the form of machine learning regression model for training the machine on various packets to detect the possibility of the attacker. In their work, they succeeded in detecting anonymous attacker by a success percent of 70-75%. However, such results will not build a legal case in the court of law, rather, it could be used as a complementary approach to verify other extracted evidence.

Francois et al. [25] proposed a generic method for conducting digital forensic in VoIP networks. They focused on detecting the used device specifications from the captured traffic. Their work was based on only signaling protocols and considered all other data as irrelevant to the analysis. Although the idea of extracting device specifications is useful in forensic analysis, as it can assist in backtracking, yet, it is generally neither enough for VoIP forensic nor is it provide an adequate information for legal cases.

Irwin and Slay [26] presented one of the most useful attempts up to that time. In their work, they proposed the idea of analyzing the RAM for discovering evidence related to VoIP usage. This is, of course, based on the used protocols especially UDP and RTP in this case. The work here is assuming that the device is maintained on, otherwise, the contents of the volatile memory, RAM, will not be kept, thus, it is capturing would not be useful. In their work, they managed to create a basic GUI for the tool that they developed and performed the analysis based on special selective criteria such as the memory and protocol to be analyzed. The presented work was able to retrieve information on VoIP call from the RAM with high true rate (97.4% for Skype call for a period of 3 minutes and 99.7 X-Lite call). The main limitation of this work is the status of the analyzed device, as switching off the device means ignoring the possibility of applying this method, which might be the case. In addition, the study focused only on retrieving information on the call rather than other embedded evidence that may be more necessary than the call itself. Moreover, there is no enough information on the claimed implemented forensic tool and interface in order to inspect, investigate and validated its usability. Interestingly, the same thoughts stimulated Irwin et al. [9] to come up with a comparative approach in order to measure the feasibility of forensically extracting VoIP evidence from virtual hard disk and RAM in the same. Yet, the result of virtual environment did not provide on available evidence related to VoIP calls.

Hsu et al. [27] investigated and proposed a collaborative methodology for VoIP forensic. The core concept in this research effort is based on including all possibility for evidence residing on network operator component as well as the Internet Service Provider (ISP) servers. They also proposed the idea of active forensic by discovering any forget header information available in SIP requests. The proposed work considered critical steps in developing VoIP forensic, however, the effort was guided only discovering fraud information embedded in SIP which is usually extremely tricky and may cause misleading results.

Ibrahim et al. [28] developed a model for VoIP evidence based on gathering attack information from different parts of VoIP application system. The main argument of the proposed work is that gathering attack information from different components of VoIP will definitely further assist in extracting a more readable digital evidence for the legal attack case. Even though the idea of the attempt sounds logical, however, the practical implementation and experimental testing are totally missing in the proposed research. This has obviously made the work incomplete could only be used for further development and investigation rather than practical applications.

Le-Khac et al. [29] conducted an experimental research on Tango forensic where they used real-life data to proof the feasibility of their work. Their idea was based on the argument that Tango had a high number of users and it could be targeted by attackers. They employed an iOS mobile device as well as an Android device for the experimental effort. The main aim of their study is to investigate Tango forensic and compare to both Whatsapp and Viber. The authors did elaborate a lot of details in their paper, yet, there is enough information on the followed methodology, the developed forensic tool as well as the future possibility of enhancement.

Psaroudakis et al. [30] presented an extensive work on VoIP forensic using SIP and SDP. They also proposed a framework for considering the readiness of the VoIP forensic for attacker identification based on volatile memory analysis. They managed to formulate a generic framework for VoIP forensic which was able to extract information about possible attackers such as IP addresses, used devices, and network topology. In addition, their readiness model helped in identifying possible threats and attackers as well as alerting and guiding network administrator for these potential risks and the available ways of combating/blocking them. The main limitation of the developed work is its complexity as well as privacy concerns generated by the fact that the application may be misused for revealing private information of legitimate users.

Manesh et al. [31] developed an extensive framework for VoIP based on network forensic procedure that considers SIP and RTP in the analysis. In their work, they proposed the idea of reordering the sequence of the generated RTP packets and used Wireshark and Ethercap along with their developed Network Forensic Analysis Tool Kit (NFATK) for this purpose. The developed system here aims at collecting any possible information on unauthorized VoIP activities, tracing the illegal or malicious content source and then generate the needed report to the proper authorities. The attempt represents a considerable effort for VoIP forensic tool. However, the main issue here is the lack of extensive testing based on various VoIP applications and internetworks scenario to verify the basic obtained results and validate the usability of the proposed framework. In a similar scenario by Sha et al. [32], the same colleagues, proposed what is called VoIP forensic analyzer as a
modified solution of the former work. However, the major limitation of adequate testing and validation remains questionable.

Carvajal et al. [33] approached VoIP forensic from a different perspective where they conducted a research on how to improve VoIP security by detecting the unprotected SIP-based VoIP. Here, the idea is based on detection rather than tracing. This can be useful for protection and early prevention of any risk issue. In simple words, the analysis was based on inspecting SIP header to confirm the status and then sending a warning if the results indicated that the traffic, in this case, is unprotected SIP type. This work can be considered as Intrusion Prevention System (IPS) which improves the security and reduces the vulnerabilities; yet, it doesn’t provide any mechanism for VoIP forensic analysis.

IV. CONCLUSIONS AND FINAL REMARKS

In this paper, the concept of VoIP along with its forensic details was presented and discussed. The discussion went through various points including the basic concept, needed protocols, as well as digital forensic tenets and ways of applying it for VoIP. This work also attempted to present and discuss the most relevant literature and previous prominent work related to VoIP forensic. Throughout the review, various models, framework, assumptions and solutions were discussed along with their limitations being highlighted for further benefit. Essentially, the field of VoIP forensic has seen quite a number of proposals, yet, an optimal solution that can be presented to the market is still to be developed. Most of the research work either was guided to retrieve an audio content, which might not help much for forensic investigation, or performed real forensic analysis but on a limited basis. This implies that an extensive work for VoIP forensic is a necessity and could save a huge loss due to the succeeded illegal activities practiced by the attackers on VoIP applications. Finally, it worth to mention that the recent trend in the area of VoIP forensic is moving from computer-based applications to smartphone and tablet devices. This is actually justified by the fact that the use of handheld devices has recently become a much popular, which in turns, creates a potential target for malicious users.

REFERENCES


