

Practical Packet Analysis: Exploring the Cybercriminal behind the LINE Voice Calls

Da-Yu Kao*, Wen-Ying Wu
Department of Information Management, Central Police University,
Taoyuan City 333, Taiwan, ROC

* Correspondence: camel@mail.cpu.edu.tw

Abstract— Cybercrimes are emerging trends on the Internet. Free instant communications, like LINE or WhatsApp, bring close to our family and friends around the globe. People in Taiwan use LINE application more frequently than others. This study uses Wireshark to explore the cybercriminal behind the LINE voice calls. Some algorithms are identified to the criminal's IP address by way of LINE voice calls. It will be helpful for Law Enforcement Agencies (LEAs) to prosecute cybercriminals, and bring them to justice.

Keywords— Packet Analysis, VOIP, LINE Application, Cybercrime Investigation

I. INTRODUCTION

A. LINE Application

A great deal of attention has been placed on the rapidly growing social chat applications recently. Free instant communications, like LINE or WhatsApp, bring close to our family and friends around the globe [23]. LINE is a proprietary application for instant communications on electronic devices such as smartphones, tablet computers and personal computers. LINE Application was developed as a mobile application for a variety of smartphone devices (iPhone, Android, Windows Phone, BlackBerry, and Nokia). The application also exists in versions for laptop and desktop computers using the Microsoft Windows and MacOS platforms [15].

B. Cyber Threats

Cyber threats are real and damaging, and shall continue to evolve at a rapid pace. The impact of cybercrime has forced Law Enforcement Agencies (LEAs) over the world to enter into a new era of collaboration to effectively tackle cyber threats together. The objective of a cybercrime investigation is to identify what happened and by whom. Whether this is for financial gain, the impact on the victims can be devastating [17] [21].

- 1) **Cybercrime Challenges for LEAs:** There are several challenges that investigators encounter when dealing with TCP/IP as evidence. On dealing into an ISP, a criminal will be assigned a dynamic IP address. Only a small group of individuals might have access to a given computer or the ability to use a specific IP address [4] [22].

- 2) **A Lack of Sufficient Information:** Every cybercrime case is unique. Evidence is scattered on the network. Law enforcement officers have to cope with a lack of sufficient information to begin with, and some important decisions may be made quickly and intuitively [1]. Looking for digital evidence on a network is a matter of exploration. It is practical to create a case-oriented checklist of all potential sources of evidence with an associated method of collection.

- 3) **The Need for Network Evidence:** Network packet can be very useful for verifying the sources of the target, corroborating relationships in the logs, or reconstructing a crime [5] [7]. Once a list of live IP addresses is gathered, a target-specific identification is possible.

C. The Need for LEAs

- 1) **Identifying LINE Application User:** The criminal using the LINE application can be anonymous, and the victim may not know who is talking to. In order to find the target in LINE application, this study tries to collect the target's Internet Protocol (IP) address, date-time stamp and port numbers by sniffing packets when the victim uses the free VoIP conversations of LINE application.
- 2) **Packet-Sniffing Software:** There are various types of packet-sniffing software. Each program is designed with different goals in mind. A few popular packet-analysis programs are tcpdump, OmniPeek, Wireshark, A-PacketMan and E-Detective. Tcpdump is a command-line program. OmniPeek, Wireshark, A-PacketMan and E-Detective have Graphical User Interfaces (GUIs) [12] [18]. Both of A-PacketMan and E-Detective are commercial Internet interception systems to collect the network traffic, and trace the user behaviors.
- 3) **Put It All Together:** A criminal will leave lots of clues in different computers/servers over time. Associating an IP address with a criminal may be conjectural ideas. More evidence is necessary to prove without reasonable doubt that the criminal at the crime scene committed the crime [9].

The rest of the paper is organized as follows. Studies of Ampkit VOIP platform, IP address, and hypothesis in LINE experiment are discussed in Section 2. Section 3 describes a novel approach to explore the criminal behind LINE voice calls. The observation on experiment results is presented in Section 4. Conclusions are given in Section 5.

II. LITERATURE REVIEWS

A. Ampkit VOIP Platform in LINE Application

LINE application uses UDP packets to send VOIP messages and the lost packets on the Internet affect media quality. Ampkit was developed as the LINE VoIP module to improve the quality of LINE call feature. The entire flow of Ampkit VOIP platform in LINE messenger is shown in Figure 1 [15] [16] [20].

- 1) **VoIP:** Voice over Internet Protocol (VoIP) systems employ session control and signalling protocols to manage the signalling, set-up, and tear-down of calls. It works with several protocols called SIP (Session Initiation Protocol), H.323, SDP (Session Description Protocol), RTP (Real-time Transport Protocol), Inter-Asterisk eXchange (IAX) and so on. Various codecs in the above protocols optimize the voice communications, which are based on application requirements and network bandwidth [11][20].
- 2) **Jitter Buffer:** There are two distinct types of network delay, called fixed and variable. Because voice, video or media is a constant bit-rate service, the jitter from all the variable delays must be removed before the signal leaves the network. Variable delays are handled through the de-jitter buffer at the receiving router/gateway. LINE media removes network jitter by expending proper de-jitter buffer lengths, and directly contact with their peers [10][16].

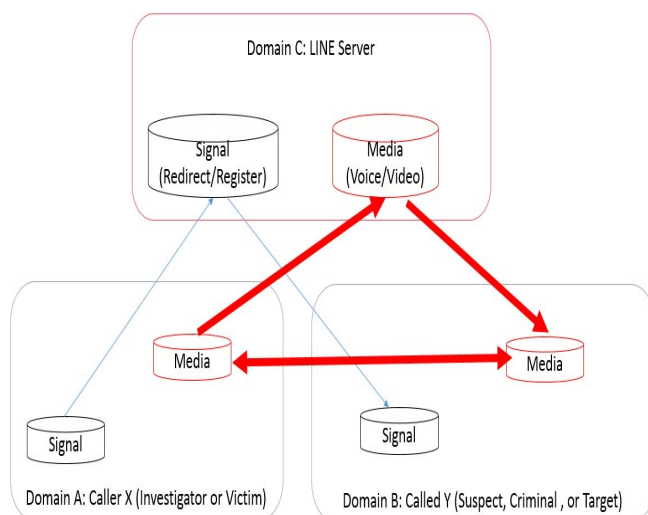


Figure 1. The flow of Ampkit VOIP Platform in LINE Messenger

- 3) **STUN Protocol:** Session Traversal Utilities for Network Address Translator (STUN) is a network protocol to allow an end computer to discover its public IP address, and to permit NAT traversal for applications of real-time voice, message, and other interactive communications. The STUN protocol allows applications operating behind a NAT to discover the presence of the network address translator and to obtain the mapped IP address and port number that the NAT has allocated for the application's User Datagram Protocol (UDP) connections to remote computers [19][20].

B. IP address in Cybercrime Investigation

- 1) **TCP/IP Protocol:** An investigator can use the source IP address to get closer to the point of origin of the criminal. Knowing the point of origin of TCP/IP traffic can also help identify suspects [6]. The TCP/IP packet contains crucial information required for two network interfaces to communicate with each other [13]. It contains fields such as source and destination IP addresses, ports, and the protocol type. Each of these fields is vital for various network layers to function, and especially for the Layer 7 application that makes use of the received data.
- 2) **Network Packet Sniffing:** Using network packets to collect relevant IP address data surrounding the target can give investigators a detailed snapshot in time of the target's location [7]. Network packet sniffing can be used for cybercrime investigation, collect an exact copy of network traffic, and capture every byte transmitted on the network. It is supposed to be used for packet capturing, network troubleshooting, or traffic analysis [3].
- 3) **Log Identification in Cybercrime Investigation:** Essential to the core investigative clues in cybercrime investigation are elements of finding the target. A logical approach to finding the target in cyber investigations is illustrated by the following elements: IP address, date-time stamp, digital action, and response message [14]. Identifying an IP address in an illegal communication is only the very beginning of a search warrant and arrest.
- 4) **Identification of a Suspect:** Identifying the IP addresses of the target is the first step for investigators in the cybercrime investigation [2][4]. Finding a public IP address in a cybercrime communication is sufficient evidence to govern the issuing of a search warrant and arrest. When an IP address is used to potentially identify a criminal by their ISP (Internet Service Provider) to their perimeter router. LEAs will need to subpoena, warrant, or court order to retrieve an IP address user from some ISPs or request them preserve the data in question. The IP address is used by intermediary routers to make a decision on which path data packets should take from source to destination [1]. Users could use dynamic IP addresses to connect their laptops to the Internet. This approach enables an ISP to reassign IP addresses at a

given time to its customers whenever necessary to make efficient use of them. Fortunately for investigators, ISPs often maintain a log of dynamic IP address assignments during a specific period [6]. Sometimes it is also vital for ISPs to identify the user from the following additional information: a particular IP address, date-time stamp, relevant port number, account name or device name.

C. Hypothesis in LINE Experiment

Decisions to effectively progress cyber investigations may have to be based on or guided by a hypothesis. Developing hypothesis is a popular technique among criminal investigators which can be used to try to establish the most logical explanation for why and how a cybercrime has been committed. Cyber investigators would be well advised to give due consideration to the following [8]:

1) Information

- Ensure that sufficient information is available to develop or test the hypothesis.
- Beware of placing too much reliance on a limited number of hypotheses when there is insufficient information available.

2) Consult

- Consult with experts to formulate hypothesis.
- Ensure that hypotheses are kept under constant review and remain dynamic.

3) Define

- Define a clear objective for the hypothesis and develop best fit hypothesis with the known information.
- Ensure a contextual understanding of the relevance, reliability and sufficiency of any clues.

III. A NOVEL APPROACH TO EXPLORE THE CRIMINAL BEHIND LINE VOICE CALLS

A. Experimental Set-Up

This experiment has set up a controlled environment to simulate the LINE voice calls, to collect IP address information, and to find out the personal identity of the target behind the LINE application.

- 1) **Software Environment:** The experiment environment was also varied on different bandwidth and traffic congestion, and affected the transmission time and the packets size. All of the devices used in this study started with the following configuration:
 - a. Laptop at Domain A
 - Wireshark v1.12.7
 - LINE Ver. 3.6.1.931
 - Window 8.1
 - Excel 2013
 - b. Cell phone at Domain B

- Android Operating System v5.0
- LINE Ver. 5.1.11

- 2) **Participants:** There are two kinds of participants in Figure 1: investigator and target. The domain A is configured by the investigator or victim. The domain B is use by the suspect, criminal, or target. All the communication packets were sniffed by Wireshark.

B. Hypothesis Testing

Every hypothesis test contains a set of two opposing statements. A hypothesis is proposed for the relationship between the two data sets, and this is compared as an alternative to a null hypothesis that proposes no relationship between two data sets. The follow-up hypothesis that is testable on the basis of observing the LINE process via a set of random communications.

- 1) **Null Hypothesis (H_0):** The null hypothesis, denoted by H_0 , is usually the hypothesis that sample observations result purely from chance. A simple hypothesis associated with a contradiction to a theory one would like to prove.
- 2) **Alternative Hypothesis (H_1):** The alternative hypothesis, denoted by H_1 or H_a , is the hypothesis that sample observations are influenced by some non-random cause. A hypothesis (often composite) associated with a theory one would like to prove.
- 3) **Can We Accept/Reject the Null Hypothesis?** The comparison is deemed significant if the relationship between the data sets would be an unlikely realization of the null hypothesis. A hypothesis test can have one of two outcomes: Researchers reject the null hypothesis or accept (or fail to reject) the null hypothesis. Acceptance implies that the null hypothesis is true. Failure to reject implies that the data are not sufficiently persuasive for us to prefer the alternative hypothesis over the null hypothesis. This study follows a formal process to determine whether to reject a null hypothesis, based on sample data.

C. The Process of Hypothesis Testing in LINE Application

The process of hypothesis testing in LINE application consists of the following four steps.

- 1) **State the Hypotheses:** This involves stating the null and alternative hypotheses. The hypotheses are stated in such a way that they are mutually exclusive. That is, if one is true, the other must be false.
- 2) **Formulate an Experiment Scenario:** The experiment scenario describes how to use sample data to evaluate the null hypothesis. All packets are sniffed and identified in this evaluation by the researcher. Based on the LINE communication architecture in Figure 1, users can communicate with others through the LINE server. If

users send messages, there is no IP address of the target in sniffing packets. This set of experiment focuses on the target's IP address from voice communication, and examine whether the IP address of the target appears in the LINE voice communication of sniffer packets. The data set of experiments focused on finding the target's IP address behind video calls through the sniffer packets. Experiment scenario was designed to find whether target B's IP address was showed in the packets.

- 3) **Analyze Sample Data:** The related packets on the Target's IP address are analysed and preserved to find the value of the test statistic in the analysis plan by the researcher. Before the investigator A used LINE to call the target B, target B has identified his IP address first and the investigator knew it. After the experiment, the researcher used that IP address to filter the packets, and to figure out whether that IP address appeared in the packets.
- 4) **Interpret Results:** Apply the decision rule described in the analysis plan. If the value of the test statistic is unlikely, based on the null hypothesis, reject the null hypothesis.

IV. EXPERIMENT OBSERVATION

This experiment collects information about the LINE communication. The packets are intercepted, and the contained data includes pictures, videos, passwords, private documents or other data transmitted by network users. This experiment is capable of analyzing the captured packets and converting them into sensible statistical information.

A. Observation Set A

1) State the Hypotheses: $H0a$ and $H1a$

$H0a$: Investigators can't find target B's IP address in the LINE voice communication of sniffer packets.

$H1a$: Investigators find target B's IP address in the LINE voice communication of sniffer packets.

2) Formulate an Experiment Scenario: The proposed experiment scenario is formulated in the following steps:

a. Step 1: Set the environment

The target B told the investigator A his IP address was '223.136.112.115'. The investigator A installed the Wireshark in a laptop, executed it to sniff packets under promiscuous modes.

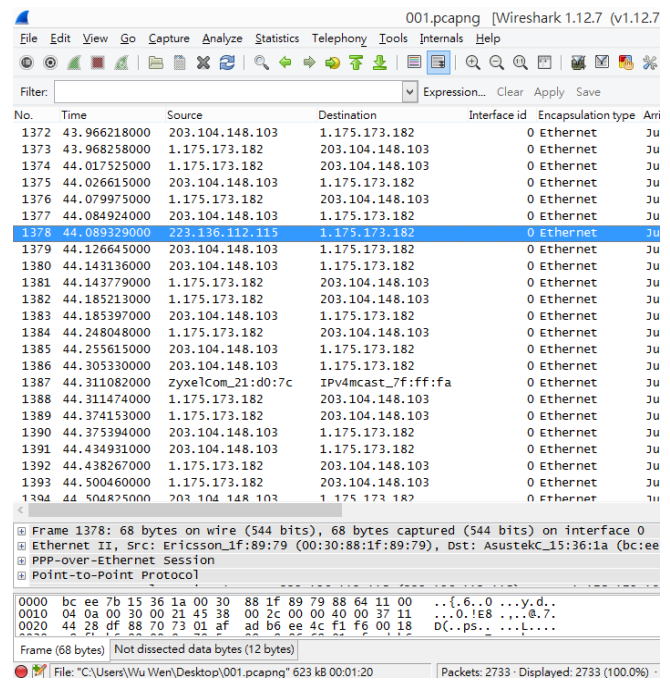
b. Step 2: Call and sniffer all the packets

Investigator A initiated a voice call function in LINE to call target B. After the end of the voice call, investigator A stopped Wireshark and saved the packets for later analysis.

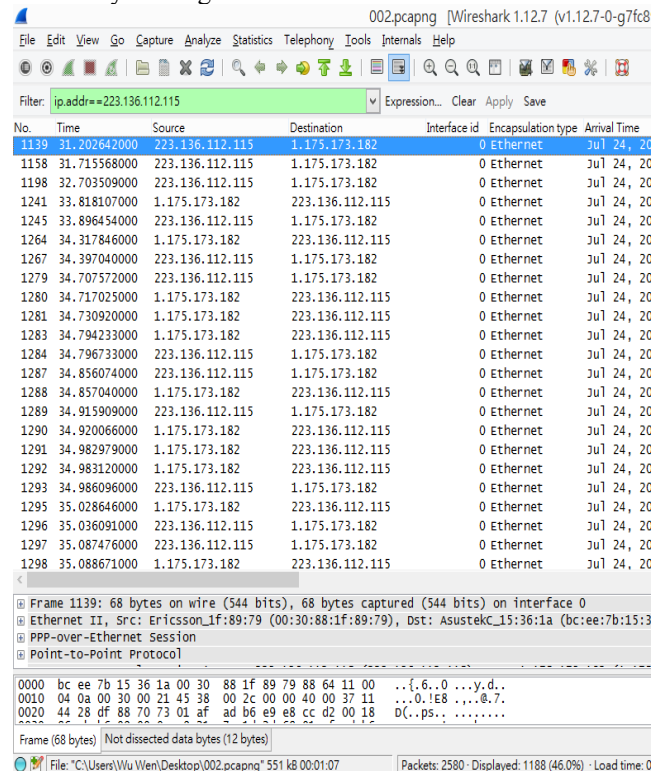
c. Step 3: Find out target B's IP address

In Figure 2, the researcher confirmed whether the target B's IP address did appear in the packets. The filter parameter (ip.addr==223.136.112.115) was set up to find all packets which have contained target B's IP address.

- 3) **Analyze Sample Data:** Investigators should be able to review the captured packets and interpret that information to form an opinion or a conclusion. The appearance of the target's IP address can be as important as the evidence itself. Careful consideration should be made as to the context of what is seen in the captured packets.



a. Identify the target B's IP address



b. Collect all target B's IP addresses

Figure 2. Experiment Observation on Set A ($H0a$ and $H1a$)

- 4) **Interpret Results:** The researcher rejects the Hypothesis H0a, and finds target B’s IP address in sniffer packets. According to experiment results, the sniffer packets did contain target B’s IP address. The researcher can find target B’s IP address through sniffer packets during voice call of LINE. The media communication of line will directly connect with mutual digital devices (computers, laptops, or mobile phones).

B. Observation Set B

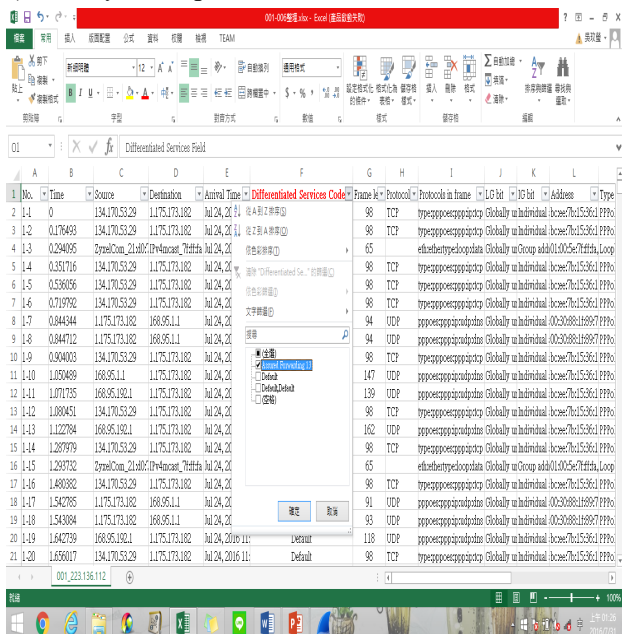
1) State the Hypotheses: H0b and H1b.

H0b: There was no pattern when the target B’ IP address appeared in the LINE voice calls of sniffer packets.

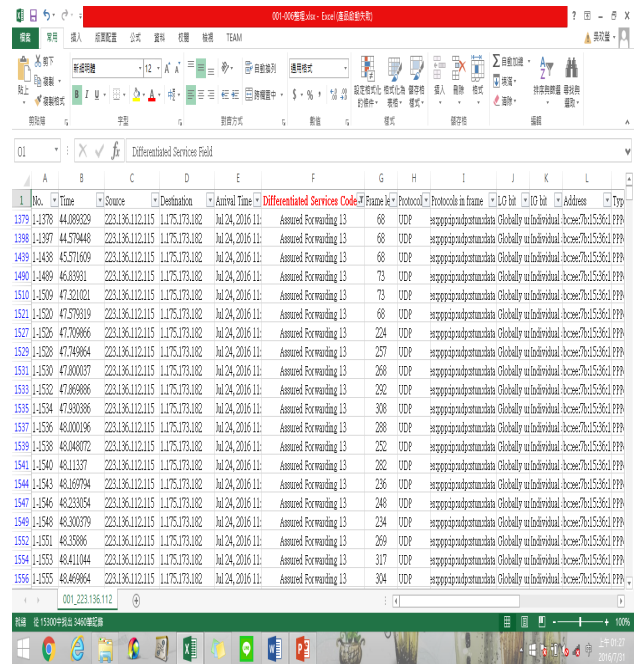
H1b: There were some patterns when the target B’ IP address appeared in the LINE voice calls of sniffer packets.

- 2) **Formulate an Experiment Scenario:** The proposed experiment scenario is formulated in the following steps in Figure 3:
- a. Step 1: Expand all fields of packets, and apply all fields as column
 - b. Step 2: Export packet dissections to a CSV file
 - c. Step 3: Open the CSV file in excel program
 - d. Step 4: Merge all experiments’ outcomes into one file
 - e. Step 5: Find the records that the value of “Differentiated Services Codepoint” column is equal to ”Assured Forwarding 13” (Figure 3a).
 - f. Step 6: Find the records that the value of “Differentiated Services Field” column is equal to ”56” (Figure 3b).

3) Analyze Sample Data



- a. Find the records that the value of “Differentiated Services Codepoint” column is equal to ” Assured Forwarding 13”



- b. Find the records that the value of “Differentiated Services Field” column is equal to ”56”

Figure 3. Experiment Observation on Set B (H0b and H1b)

- 4) **Interpret Results:** The researcher rejects the Hypothesis H0b, and some patterns of the target B’ IP address appeared in the LINE voice calls of sniffer packets. The algorithm for finding the target B’s IP address is listed below:
- a. If the value of field “Differentiated Services Codepoint” was equal to ”Assured Forwarding 13,” or
 - b. If the value of field” Differentiated Services Field” was equal to ”56,”
 - c. All the source IP address in sniffer packets would be target B’s IP address.

C. The Observation on Experiment Results

Since this experiment requires actual user responses to VOIP voice call, investigators have better to use the victim account to call/sniff the target in a controlled environment. The Wireshark was configured on the investigator computer in promiscuous mode. The observation on this experiment results is listed in Table 1. According to the experiment of LINE calls, three opinions are proposed below. First, the target’s IP address did appear in the packets during LINE VOIP calls. Second, if the value of “Differentiated Services Codepoint” column is equal to ”Assured Forwarding 13”, the source IP address would be target B’s IP address. Third, if the value of “Differentiated Services Field” column is equal to ”56”, the source IP address would be target B’s IP address. The above findings are helpful for LEAs to investigate cybercrime if the criminal use LINE application. The resulting digital evidence in this experiment can be used to corroborate other activities on related clues.

TABLE 1. THE OBSERVATION ON EXPERIMENT RESULTS

Hypothesis Testing	State the Hypotheses	Analyze Sample Data	Interpret Results
H0a	Investigators can't find target B's IP address in the LINE voice communication of sniffer packets.	Reject the Hypothesis H0a.	Find target B's IP address in sniffer packets.
H1a	Investigators find target B's IP address in the LINE voice communication of sniffer packets.		
H0b	There was no pattern when the target B' IP address appeared in the LINE voice calls of sniffer packets.	Reject the Hypothesis H0b.	Some patterns of the target B' IP address appeared in the LINE voice calls of sniffer packets
H1b	There were some patterns when the target B' IP address appeared in the LINE voice calls of sniffer packets.		
Experiment Results	(1) If the value of field "Differentiated Services Codepoint" was equal to "Assured Forwarding 13," or (2) if the value of field" Differentiated Services Field" was equal to "56," (3) all the source IP address in sniffer packets would be target B's IP address.		

V. CONCLUSION AND FUTURE WORKS

If an IP address resides in the captured packets, it means that the investigator should identify its provenance and provide context as to when, how, and whether it has been obtained from the target computer. The interpretation of such information may help in proving, or disproving, a location of investigation. This study has tested and demonstrated the way to find the Target's IP address in LINE voice calls. The novel approach of finding the target's IP address will be helpful for LEAs to identify or prosecute cybercriminals. It would be very interesting if researchers compare the contents of the sniffer packets and improve our methodology to find more patterns on various VOIP patterns. Several issues are worth doing in our future research. First, we will investigate more techniques to trace the patterns which are related to the target. Second, we want to investigate the possibility of learning normal patterns among various VOIP communications to improve the verification of correlations and making hypotheses.

ACKNOWLEDGMENT

This research was partially supported by the Ministry of Science and Technology of the Republic of China under the Grants MOST 105-2221-E-015-001-.

REFERENCES

- [1] Akhgar, B., Staniforth, A., and Bosco, F., *Cyber Crime and Cyber Terrorism Investigator's Handbook*, Elsevier Publishing, pp. 88-90, 2014.
- [2] Androulidakis, I. I., *VoIP and PBX Security and Forensics: A Practical Approach* (2nd Edition). Switzerland: Springer International Publishing, pp. 75-99, 2016.
- [3] Baxter J. H., *Wireshark Essentials: Get Up and Running with Wireshark to Analyze Network Packets and Protocols Effectively*, Birmingham, Packt Publishing, pp. 7-142, 2014.
- [4] Bernik, I., *Cybercrime and Cyberwarfare*, John Wiley & Sons Inc., pp. 1-57, 2014.
- [5] Brooks, C. L., *CHFI Computer Hacking Forensic Investigator Certification All-in-One Exam Guide* (1st Edition), McGraw-Hill Education, pp. 13-50, 2015.
- [6] Casey, E., *Handbook of Digital Forensics and Investigation*, Elsevier Inc., pp. 21-208, 2010.
- [7] Casey, E., *Digital Evidence and Computer Crime: Forensic Science, Computers, and the Internet* (3rd Edition), Elsevier Inc., pp. 187-306, 2011.
- [8] Cook, T. and Tattersall, A., *Blackstone's Senior Investigating Officers' Handbook* (2nd Edition). Oxford University Press, Oxford, 2010.
- [9] Graves, M. W., *Digital Archaeology: The Art and Science of Digital Forensics*, Addison-Wesley, pp. 91-110, 2014.
- [10] Hardy, W. C., *VoIP Service Quality: Measuring and Evaluating Packet-Switched Voice*, McGraw-Hill Professional, pp. 3-96, 2003.
- [11] Hartpence, B., *Packet Guide to Voice over IP: A system administrator's guide to VoIP technologies*, O'Reilly Media, pp. 1-120, 2013.
- [12] Jesin A., *Packet Tracer Network Simulator*, Birmingham, Packt Publishing, pp. 7-44, 2014.
- [13] Joseph, V. and Mulugu, S., *Network Convergence. Ethernet Applications and Next Generation Packet Transport Architectures*, Morgan Kaufmann, pp. 467-526, 2014.
- [14] Kao, D. Y. and Wang, S. J., "The IP Address and Time in Cyber-crime Investigation," *Policing: An International Journal of Police Strategies & Management*, 32 (2) , 2009.
- [15] (2016) LINE Corporation website. [Online]. Available: <https://line.me/en/>
- [16] (2015) LINE Engineer, LINE Developer Week 2015, "Ampkit: Mobile VoIP Platform in LINE Messenger by Jungnam Gwock," [Online]. Available: <http://developers.linecorp.com/blog/?p=3718>
- [17] Mehan, J. E., *Cybererror, Cybercrime and Cyberactivism: an In-depth Guide to the Role of Security Standards in the Cybersecurity Environment*, IT Governance Publishing, pp. 25-58, 2014.
- [18] Nath, A., *Packet Analysis with Wireshark*, Birmingham, Packt Publishing, pp. 43-147, 2015.
- [19] Rahbar, A. G., *Quality of Service in Optical Packet Switched Networks*, IEEE Press, pp. 19-43, 2015.
- [20] Roy, R. R., *Handbook on session initiation protocol: networked multimedia communications for IP telephony*, CRC Press, pp. 1-350, 2016.
- [21] Shipley, T. G. and Bowker, A., *Investigating Internet Crimes: An Introduction to Solving Crimes in Cyberspace*, Elsevier Inc., pp. 21-38, 2014.
- [22] Stephenson, P., *Official (ISC)²® Guide to the CCFP CBK*, Auerbach Publications, pp. 293-404, 2014.
- [1] Yeung, J., *Voice over Internet Protocol: A Practical Guide for the Non-telephone Engineer*, Markmonitor Inc., p13-39, 2015.



Dayu Kao is an Associate Professor at Department of Information Management, College of Police Science and Technology, Central Police University, Taiwan. With a Master degree in Information Management and a PhD degree in Crime Prevention and Correction, he had led several investigations in cooperation with police agencies from other countries for the past 20 years. He can be reached at camel@mail.cpu.edu.tw.



Wen-Ying Wu is a student at Department of Information Management, College of Police Science and Technology, Central Police University, Taiwan.