http://www.icact.org/journal.asp

ICACT-TACT JOURNAL

Transactions on Advanced Communications Technology

Volume 7 Issue 2, March. 2018, ISSN: 2288-0003

Editor-in-Chief
Prof. Thomas Byeongnam YOON, PhD.

GIRI
Global IT Research Institute

Journal Editorial Board

Editor Guide

■ Introduction for Editor or Reviewer

All the editor group members are to be assigned as a evaluator(editor or reviewer) to submitted journal papers at the discretion of the Editor-in-Chief. It will be informed by eMail with a Member Login ID and Password.

Once logined the Website via the Member Login menu in left as a evaluator, you can find out the paper assigned to you. You can evaluate it there. All the results of the evaluation are supposed to be shown in the Author Homepage in the real time manner. You can also enter the Author Homepage assigned to you by the Paper ID and the author's eMail address shown in your Evaluation Webpage. In the Author Homepage, you can communicate each other efficiently under the peer review policy. Please don't miss it!

All the editor group members are supposed to be candidates of a part of the editorial board, depending on their contribution which comes from history of ICACT TACT as an active evaluator. Because the main contribution comes from sincere paper reviewing role.

■ Role of the Editor

The editor's primary responsibilities are to conduct the peer review process, and check the final camera-ready manuscripts for any technical, grammatical or typographical errors.

As a member of the editorial board of the publication, the editor is responsible for ensuring that the publication maintains the highest quality while adhering to the publication policies and procedures of the ICACT TACT(Transactions on the Advanced Communications Technology).

For each paper that the editor-in-chief gets assigned, the Secretariat of ICACT Journal will send the editor an eMail requesting the review process of the paper.

The editor is responsible to make a decision on an "accept", "reject", or "revision" to the Editor-in-Chief via the Evaluation Webpage that can be shown in the Author Homepage also.

■ Deadlines for Regular Review

Editor-in-Chief will assign a evalaution group(a Editor and 2 reviewers) in a week upon receiving a completed Journal paper submission. Evaluators are given 2 weeks to review the paper. Editors are given a week to submit a recommendation to the Editor-in-Chief via the evaluation Webpage, once all or enough of the reviews have come in. In revision case, authors have a maximum of a month to submit their revised manuscripts. The deadlines for the regular review process are as follows:

Evalution Procedure	Deadline
Selection of Evaluation Group	1 week
Review processing	2 weeks
Editor's recommendation	1 week
Final Decision Noticing	1 week

■ Making Decisions on Manuscript

Editor will make a decision on the disposition of the manuscript, based on remarks of the reviewers. The editor's recommendation must be well justified and explained in detail. In cases where the revision is requested, these should be clearly indicated and explained. The editor must then promptly convey this decision to the author. The author may contact the editor if instructions regarding amendments to the manuscript are unclear. All these actions could be done via the evaluation system in this Website. The guidelines of decisions for publication are as follows:

Decision	Description
Accept	An accept decision means that an editor is accepting the paper with no further modifications. The paper will not be seen again by the editor or by the reviewers.
Reject	The manuscript is not suitable for the ICACT TACT publication.
Revision	The paper is conditionally accepted with some requirements. A revision means that the paper should go back to the original reviewers for a second round of reviews. We strongly discourage editors from making a decision based on their own review of the manuscript if a revision had been previously required.

■ Role of the Reviewer

Reviewer Webpage:

Once logined the Member Login menu in left, you can find out papers assigned to you. You can also login the Author Homepage assigned to you with the paper ID and author's eMail address. In there you can communicate each other via a Communication Channel Box.

Quick Review Required:

You are given 2 weeks for the first round of review and 1 week for the second round of review. You must agree that time is so important for the rapidly changing IT technologies and applications trend. Please respect the deadline. Authors undoubtedly appreciate your quick review.

Anonymity:

Do not identify yourself or your organization within the review text.

Review:

Reviewer will perform the paper review based on the main criteria provided below. Please provide detailed public comments for each criterion, also available to the author.

- How this manuscript advances this field of research and/or contributes something new to the literature?
- Relevance of this manuscript to the readers of TACT?
- Is the manuscript technically sound?
- Is the paper clearly written and well organized?
- Are all figures and tables appropriately provided and are their resolution good quality?
- Does the introduction state the objectives of the manuscript encouraging the reader to read on?
- Are the references relevant and complete?

Supply missing references:

Please supply any information that you think will be useful to the author in revision for enhancing quality of the paperor for convincing him/her of the mistakes.

Review Comments:

If you find any already known results related to the manuscript, please give references to earlier papers which contain these or similar results. If the reasoning is incorrect or ambiguous, please indicate specifically where and why. If you would like to suggest that the paper be rewritten, give specific suggestions regarding which parts of the paper should be deleted, added or modified, and please indicate how.

Journal Procedure

Dear Author,

- **You can see all your paper information & progress.**

- **Step 1. Journal Full Paper Submission**

Using the Submit button, submit your journal paper through ICACT Website, then you will get new paper ID of your journal, and send your journal Paper ID to the Secretariat@icact.org for the review and editorial processing. Once you got your Journal paper ID, never submit again! Journal Paper/CRF Template

- **Step 2. Full Paper Review**

Using the evaluation system in the ICACT Website, the editor, reviewer and author can communicate each other for the good quality publication. It may take about 1 month.

- **Step 3. Acceptance Notification**

It officially informs acceptance, revision, or reject of submitted full paper after the full paper review process.

Status	Action
Acceptance	Go to next Step.
Revision	Re-submit Full Paper within 1 month after Revision Notification.
Reject	Drop everything.

- **Step 4. Payment Registration**

So far it's free of charge in case of the journal promotion paper from the registered ICACT conference paper! But you have to regist it, because you need your Journal Paper Registration ID for submission of the final CRF manuscripts in the next step's process. Once you get your Registration ID, send it to Secretariat@icact.org for further process.

- **Step 5. Camera Ready Form (CRF) Manuscripts Submission**

After you have received the confirmation notice from secretariat of ICACT, and then you are allowed to submit the final CRF manuscripts in PDF file form, the full paper and the Copyright Transfer Agreement. Journal Paper Template, Copyright Form Template, BioAbstract Template,

Journal Submission Guide

All the Out-Standing ICACT conference papers have been invited to this "ICACT Transactions on the Advanced Communications Technology" Journal, and also welcome all the authors whose conference paper has been accepted by the ICACT Technical Program Committee, if you could extend new contents at least 30% more than pure content of your conference paper. Journal paper must be followed to ensure full compliance with the IEEE Journal Template Form attached on this page.

➢ How to submit your Journal paper and check the progress?

Step 1. Submit	Using the Submit button, submit your journal paper through ICACT Website, then you will get new paper ID of your journal, and send your journal Paper ID to the Secretariat@icact.org for the review and editorial processing. Once you got your Journal paper ID, never submit again! Using the Update button, you can change any information of journal paper related or upload new full journal paper.
Step 2. Confirm	Secretariat is supposed to confirm all the necessary conditions of your journal paper to make it ready to review. In case of promotion from the conference paper to Journal paper, send us all the .DOC(or Latex) files of your ICACT conference paper and journal paper to evaluate the difference of the pure contents in between at least 30% more to avoid the self replication violation under scrutiny. The pure content does not include any reference list, acknowledgement, Appendix and author biography information.
Step 3. Review	Upon completing the confirmation, it gets started the review process thru the Editor & Reviewer Guideline. Whenever you visit the Author Homepage, you can check the progress status of your paper there from start to end like this, " Confirm OK! -> Gets started the review process -> ...", in the Review Status column. Please don't miss it!

Volume. 7 Issue. 2

Analyzing WannaCry Ransomware Considering the Weapons and Exploits

Da-Yu KAO*, Shou-Ching HSIAO**, Raylin TSO***

*Department of Information Management, Central Police University, Taoyuan City 333, Taiwan

**Haishan Precinct, New Taipei City Police Department, New Taipei City 220, Taiwan

***Department of Computer Science, National Chengchi University, Taipei 116, Taiwan

dayukao@gmail.com, oliver84312@gmail.com, raylin@cs.nccu.edu.tw

Abstract— **As ransomware has increased in popularity, its creators are using our fears to their advantage. The rapid proliferation of ransomware attacks indicates the growing tendency of ransomware-as-a-service (RaaS) and the integration of hacking weapons. This paper presents the analysis of the infamous WannaCry ransomware, which is one of the most propagated and damaging malware in 2017. The anatomy of ransomware attacks is discussed to understand the multi-phased execution of WannaCry, including the deployment, installation, destruction, and command-and-control. The chain of WannaCry's execution comprises several hacking weapon components. WannaCry not only embeds the binary in the resource section for multi-phased execution, but also implements a strong encrypting algorithm and a key structure. A reverse engineering analysis of each component, along with the network analysis of WannaCry's exploits offers an insight into the inner design of WannaCry. The observations of this research contribute to recent security systems and future defense strategies.**

Keywords—**Ransomware, Reverse Engineering Analysis, Network Analysis, Hacking Weapons, WannaCry Exploits**

I. INTRODUCTION

During previous decades, malware has evolved in terms of the sophisticated obfuscation of malicious software and the diversity of attack vectors [4]. Ransomware is one of the greatest and most rapidly growing threats to the digital world [11]. Ransomware typically operates by locking the desktop of a computer and by rendering it to be inaccessible to users or by encrypting, overwriting, or deleting the user's files [6]. Ransomware can cause global catastrophes using encryption to hold the victims' data for ransom. Further, ransomware attacks continue to target out-of-date systems as the recent WannaCry ransomware (also known as WCry, WannaCrypt, WannaCrpyt0r, or WannaCrpytor) has spread in a tragic scenario containing thousands of computers [5, 10]. The emergence of malware creation tools has facilitated the creation of new variations of the existing ransomware [1]. Ransomware can easily to modify its ability to propagate quickly [10]. The dark web is a repository of the hacking weapons. By installing the TOR (The Onion Router) browser, criminals can access the dark web to realize their intentions of conducting ransomware attacks, which requires only a few hundred dollars. Easy access to hacking weapons lowers the barrier to initialize a cyberattack. After a hacking weapon was newly developed and implemented during a hacking campaign or malware outbreak, it has become a component in the circular chain of hacking weapons. An observation of the hacking weapons in WannaCry has revealed that some of the modular code was obtained from public source or covert hacker channels while the other parts of code were observed to be designed by the creator. Additionally, the hacking weapons are reusable by nature.

A literature review is presented in Section 2. A reverse engineering analysis of WannaCry components is discussed in Section 3. A network analysis of WannaCry exploits is presented in Section 4, and our research observations about the execution of multi-stage WannaCry are described in Section 5. Our conclusions are given in Section 6.

II. LITERATURE REVIEWS

WannaCry contains various modular hacking weapons in its composition. (Fig. 1).

Fig. 1. Hacking weapons in weaponized ransomware

Manuscript received Dec 2, 2017. This work was a follow-up of the invited journal to the accepted & presented paper of the 20th Conference on Advanced Communication Technology (ICACT2018), and this research was partially sponsored by the Executive Yuan of the Republic of China under the Grants Forward-looking Infrastructure Development Program (Digital Infrastructure-Information Security Project-107) and the Ministry of Science and Technology of the Republic of China under the Grants MOST 106-2221-E-015-002.

Da-Yu Kao is with the Department of Information Management, Central Police University, Taoyuan City 333, Taiwan (Corresponding Author phone: +886-3-328-2321; fax: +886-3-328-5189; e-mail: dayukao@gmail.com).

Shou-Ching Hsiao is with Haishan Precinct, New Taipei City Police Department, New Taipei City 220, Taiwan (phone:+886-2-964-0329; e-mail: oliver84312@gmail.com).

Raylin Tso is with the Department of Computer Science, National Chengchi University, Taipei 116, Taiwan (phone: +886-2-2939-3091; +886-2-2937-8629; e-mail: raylin@cs.nccu.edu.tw).

A. Weaponized Malware

Weaponized malware deserves its name in two folds: the sophistication in composition and the intent for malicious purposes. As the more complexity and scale of malware attacks increase, malware developers tend to weaponize the malicious binary using different hacking weapons. WannaCry is a compound example of malware that not only contains dropper, resource loader, and ransomware binary for multi-execution flow, but is also weaponized with the Eternalblue exploit to ensure worm propagation capability. WannaCry is a type of worm-enabled ransomware and can be used as a weapon of digital destruction, which has cast a gloom over hospitals, banks, and enterprises all over the world, forcing individuals, enterprises, and public agencies alike to cease operation as people they attempt to cope with their infected computers.

B. RaaS: Ransomware-as-a-service

As cyber attackers increasingly use various Internet referrals to acquire ransomware modules, the convenience and service in RaaS have made it a new trend for people who intend to commit cybercrimes [1]. Since cybercriminals simply release these malicious codes on open source platforms, cybercrime has nothing to do with programming ability or hacking techniques anymore. Further, anyone can implement a ransomware attack easily, thereby drastically increasing the number of ransomware attacks. The unprecedented scale of RaaS has made cybercrimes more achievable and attainable, causing the wide spread of ransomware.

C. Anatomy of a Ransomware Attack

The objective of any ransomware attack is to extort the victims. Hackers who intend to conduct any type of attack tend to follow typical attack techniques and procedures. The life cycle of a general ransomware attack comprises the following stages [1] [7]:

1) ***Deployment.*** The first stage of a ransomware attack is to enter into targeted machine and execute its files. Several deployment methods, including phishing emails, malicious websites, vulnerable exploits are observed to vary from one to another.

2) ***Installation.*** After accessing the system initially, the ransomware will install and attempt to take complete control of the infected host. After successful control, the ransomware may either add its autorun registry key, create itself as a service, or dll load-order hijacking to achieve persistence.

3) ***Destruction.*** The ransomware blocks users' access to documents or systems by locking and encrypting files on the compromised device. Usually, ransomware will use public key algorithm along with private key algorithm to form complex encryption structure.

4) ***Command-and-Control.*** The actions of ransomware attack depend on the form of command-and-control systems. The metamorphic ransomware families and variants may differentiate the command-and-control channels, which may sometimes be as simple as web-based communications using HTTP protocol to as complex as the complicated TOR service connections.

III. Reverse Engineering Analysis Of WannaCry Components

The main reason for applying reverse engineering to the WannaCry ransomware is to reveal the actual functionality of the binary, which is a module of code, and why it comes as such designation. The "IDA Pro" is a useful reversing tool for disassembling the WannaCry binaries and offers a deep insight about the manner in which the WannaCry was developed and about the details of its execution flow. Different components, such as the launcher, dropper, resource loader, main ransom body, and encryption dll, implement the functionality in each phase. The chain of reverse engineering analysis is explored by extracting the main components during execution. The details of WannaCry components are listed in Table I.

A. Deployment Phase: Export PlayGame

A WannaCry ransomware attack exploits the MS17-010 vulnerability to inject the initial binary "launcher.dll" through the Eternalblue exploit and Doublepulsar backdoor. WannaCry exploits the SMB driver "srv2.sys" in the kernel module to access the compromised devices and inject the malicious payload [5]. Further, "launcher.dll" is injected into the lsass.exe system process and serves as the loader for mssecsvc.exe (Fig. 2).

Fig. 2. Launching mssecsvc.exe within lsass.exe process

The "launcher.dll" is executed only in memory and leaves no file artifacts on disk. This paper examined the lsass process memory from memory dumps using the "RWX (Read, Write, and Execute)" permission attributes. After the dumped memory was loaded into IDA Pro, the exported entry exhibited that this DLL can be accommodated within PlayGame, which is tasked to start up the ransomware execution. The PlayGame function mainly calls two sub-functions, "ExtractResource" and "CreateProcessMSSECSVC" (Fig. 3).

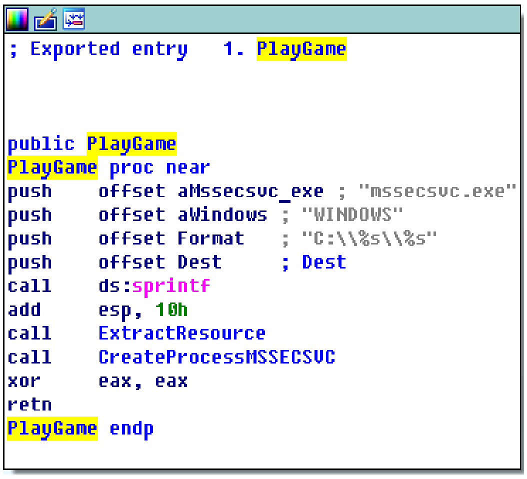


Fig. 3. Export PlayGame

TABLE I
Main Components of WannaCry

Phase	Execution Component	WannaCry		
		File (Internal) Name	***File Description***	***SHA256***
Deployment	Export PlayGame	launcher.dll	Inject through Doublepulsar backdoor	9411c59a83c8c32a925d53a902bef168ebe5b403a88ab4d8dfe807fd7435dd9e
Installation	Dropper and Infection	mssecsvc.exe (lhdfrgui.exe)	Microsoft® Disk Defragmenter	24d004a104d4d54034dbcffc2a4b19a11f39008a575aa614ea04703480b1022c
	Resource Loader	tasksche.exe (diskpart.exe)	DiskPart	ed01ebfbc9eb5bbea545af4d01bf5f1071661840480439c6e5babe8e080e41aa
Destruction	Encryption DLL	kbdlv (3.13)	Latvia Keyboard Layout	1be0b96d502c268cb40da97a16952d89674a9329cb60bac81a96e01cf7356830
Command-and-control	Trace Infection and Payments	@WanaDecryptor@.exe (LODCTR.EXE)	Load PerfMon Counters	b9c5d4339809e0ad9a00d4d3dd26fdf44a32819a54abf846bb9b560d81391c25

The function "ExtractResource" aims for extracting "W/101" resource to create the mssecsvc.exe file at "C:\WINDOWS\mssecsvc.exe" and to launch it. The Windows API "CreateFileA" and "WriteFile" are used to write the loaded resource to "mssecsvc.exe," which is followed by a series of resource extraction routines, including "FindResourceA", "LoadResource", "LockResource", and "SizeofResource". Finally, the mssecsvc process is launched through calling "CreateProcessA" in the "CreateProcessMSSECSVC" sub-function.

B. *Installation Phase: Dropper, Infection, and Resource Loader*

The infection begins when the ransomware payload is delivered to the victim's machine. Once the payload successfully injects the launcher.dll into the lsass.exe system process, the dll launches mssecsvc.exe, which analyzes the system to determine whether it is located on a real computer or in a virtual sandbox [11]. Before any operation, two Windows API "InternetOpenA" and "InternetOpenUrlA" are used to query a hard-coded domain name, which in this sample, checks "www.iuqerfsodp9ifjaposdfjhgosurijfaewrwergwea.com" (kill-switch URL). A successful connection will cause the mssecsvc.exe to terminate. Otherwise, it will proceed with the dropping of "tasksche.exe" and the infection (Fig. 4). The kill-switch in the execution flow provides an opportunity to slow down the malware [10].

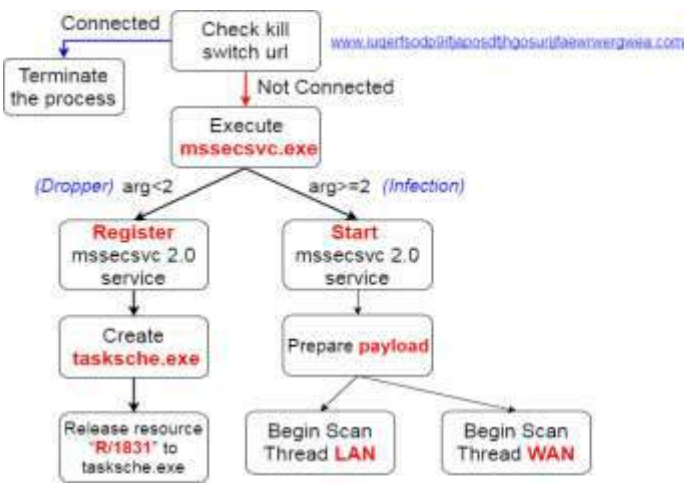


Fig. 4. The flow chart of mssecsvc2.0 installation

In the installation phase, "mssecsvc.exe" and "tasksche.exe" are two focused binaries. The "mssecsvc.exe" comprises two main execution functions, dropper and infection, and has a different entry point of execution depending on the command parameters (Algorithm1). The "tasksche.exe" is responsible for resource loading and the encryption environment setting.

Algorithm 1:

if (*argc* < 2)

then

InstallMssecsvc2.0Service();

ExtractResourceToTasksche(); *(Dropper Phase)*

else

Call StartServiceCtrlDispatcherA() to start mssecsvc2.0 service; *(Infection Phase)*

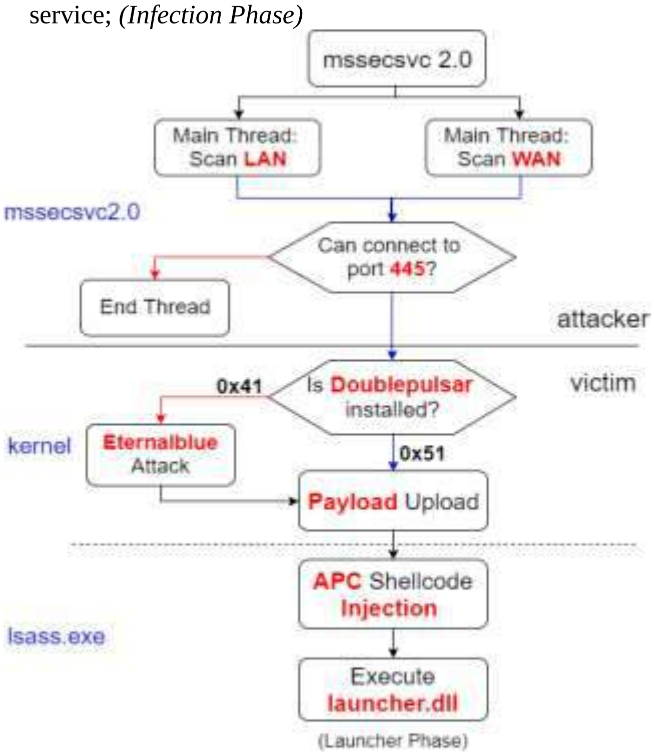


Fig. 5. The flow chart of infection

1) ***Dropper.*** At the beginning of execution, the "mssecsvc.exe" is run without any parameter. Two sub-functions are called to install the "mssecsvc2.0" service and to drop the next-stage binary "tasksche.exe". WannaCry is highly modular in a multi-stage campaign. This resource extraction routine is exactly a modular

example. After the complete extraction of the resource binary, the contents of the binary are written into “tasksche.exe”.

2) ***Infection.*** If mssecsvc runs with a parameter “-m security”, the execution falls to the infection function. The mssecsvc service is created to abuse the exploit of MS17-010 and the Doublepulsar backdoor for infection [8]. Fig. 5 summarizes WannaCry’s infection flow, including the initial stage of the mssecsvc2.0 service that is running on the attackers’ machine and the kernel Doublepulsr backdoor’s implantation on the victims’ machine. On the attacker machine, the mssecsvc2.0 service probes SMB protocol and port 445 [5]. If successfully connected, the attacker will be able to transmit the crafted packets with specific opcode to verify whether the Doublepulsar backdoor was set up to upload the payload. If the target machine has not had Doublepulsar backdoor installed (0x41 in response), the exploit code will proceed to initialize the Eternalblue attack. As soon as the setup of Doublepulsar is confirmed, the payload will be uploaded directly. The payload contains the kernel shellcode, userland shellcode, and launcher.dll with mssecsvc binary embedded in the resource section. The anatomy of the infection can also be revealed by analyzing the network packet using Wireshark.

3) ***Resource Loader.*** The main ransomware “tasksche.exe” is thrown by “mssecsvc.exe” into the dropper phase. It extracts the compressed XIA resource from its resource section, which contains several specific WannaCry files. While analyzing the tasksche reversing code, we organized the ransomware execution flow as depicted in Fig. 6. First, it generates a randomized unique ID for naming the folder that was prepared to contain the extracted resource. Second, the tasksche process verifies if a parameter exits prior to the execution of any operation. The command parameter “i” represents the installation mode of “tasksche.exe”. After the installation, tasksche.exe is run without any parameter; further, a chain of functions is called to prepare for the encryption phase. It creates an autorun registry key as persistence mechanism, releases the resource zip file, and unzips it into the installation folder. Additionally, WannaCry uses the rand() function to randomly select one of the three hardcoded bitcoin addresses and writes it to c.wnry. Then, WannaCry adds the hidden attributes to the installation folder, grants complete access to all users, and decrypts the t.wnry binary to generate the encryption dll.

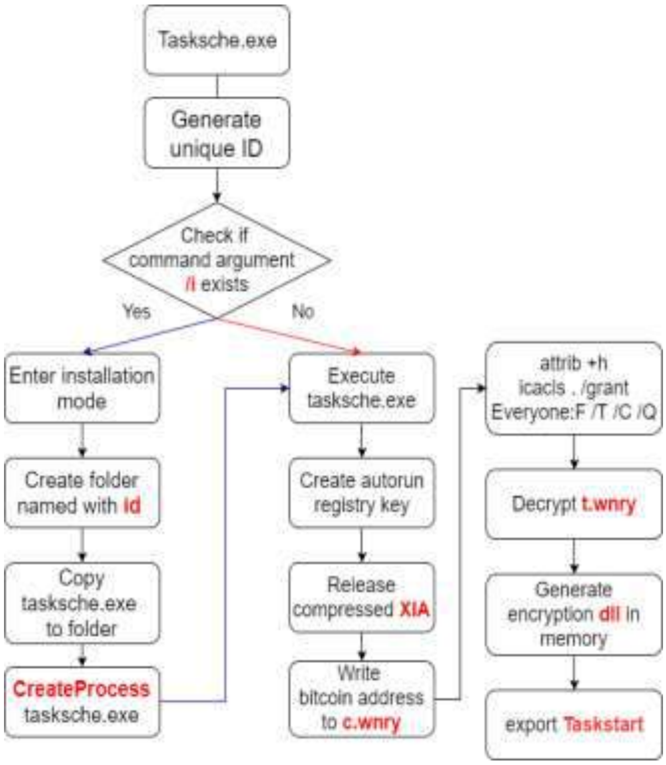


Fig. 6. The flow chart of main ransomware

The resource zip file “XIA” embedded in the resource section is extracted as the same resource loading process in the launcher and dropper phase. As soon as the resource loading completed, the code will unzip the resource file using the password “WNcry@2ol7”. The XIA resource contains several WannaCry files, which are presented in Table II.

Table II

Files in XIA Resource

File Name	XIA Resource	
	File Description	***MD5***
msg\m_*.wnry	ransom notes in different languages	
b.wnry	display instructions for decryption	c17170262312f3be 7027bc2ca825bf0c
c.wnry	target address and TOR information	c17170262312f3be 7027bc2ca825bf0c
r.wnry	ransom note	c17170262312f3be 7027bc2ca825bf0c
s.wnry	TOR software executable	ad4c9de7c8c40813 f200ba1c2fa33083
t.wnry	encrypted ransomware DLL	ad4c9de7c8c40813 f200ba1c2fa33083
u.wnry	“@WanaDecryptor@.exe” decrypter file	7bf2b57f2a205768 755c07f238fb32cc
f.wnry	decrypt for demo	c17170262312f3be 7027bc2ca825bf0c
taskdl.exe	Enumerating and deleting temp files	4fef5e34143e646d bf9907c4374276f5
taskse.exe	Enumerate active RDP sessions and run a process on connected remote machines	8495400f199ac778 53c53b5a3f278f3e
@WanaDecry ptor@.exe	Present user interface, C&C communication, and volume shadow deletion.	7bf2b57f2a205768 755c07f238fb32cc
00000000.eky	generated private key	6317124f38c33cce 36291ec3bc835db4
00000000.pky	generated public key	6f4e6640a2bc54a0 778130f7a25cb1b1
00000000.res	TOR/C2 information	168d54591c029609 959eb4256cbcea26

C. Destruction Phase: Encryption DLL

All the files on the victims' machines begin to be infected, encrypted, or locked by WannaCry. In the resource loader phase, tasksche.exe will throw a zip file from the resource section, which includes the t.wnry file. After a series of pre-processing tasks, tasksche.exe will decrypt the t.wnry into a dll exported TaskStart as the beginning of the encryption. The encryption flow and the key system are the two main themes that are closely related (Fig. 7). The encryption operation is heavily dependent on the management of the key system. WannaCry uses various pairs of keys to successfully form the encryption flow, including the RSA (Ron Rivest, Adi Shamir and Leonard Adleman) and AES (Advanced Encryption Standard) algorithms.

The key system begins from the RSA root public key, whose corresponding root private key is in the hands of WannaCry author. It is difficult for others to source the root key and to solve the encryption knot. A pair of RSA-2048 public and private keys is generated by the DLL, which respectively saves as 00000000.pky and 00000000.eky files. Prior to the RSA-2048 private key being saved to the 00000000.eky file, the private key is encrypted by the root public key in advance. For each targeted file, the encryption routine is initiated by the random creation of AES-128 encryption keys for different files. A unique AES-128 encryption key is also encrypted by the public key read from the 00000000.pky. During the generation of the encrypted file, the unique encrypted AES-128 key is embedded into the encrypted file's header followed by the 8-byte magic value "WANACRY!" and the 4-byte length of AES key as shown in Fig. 8.

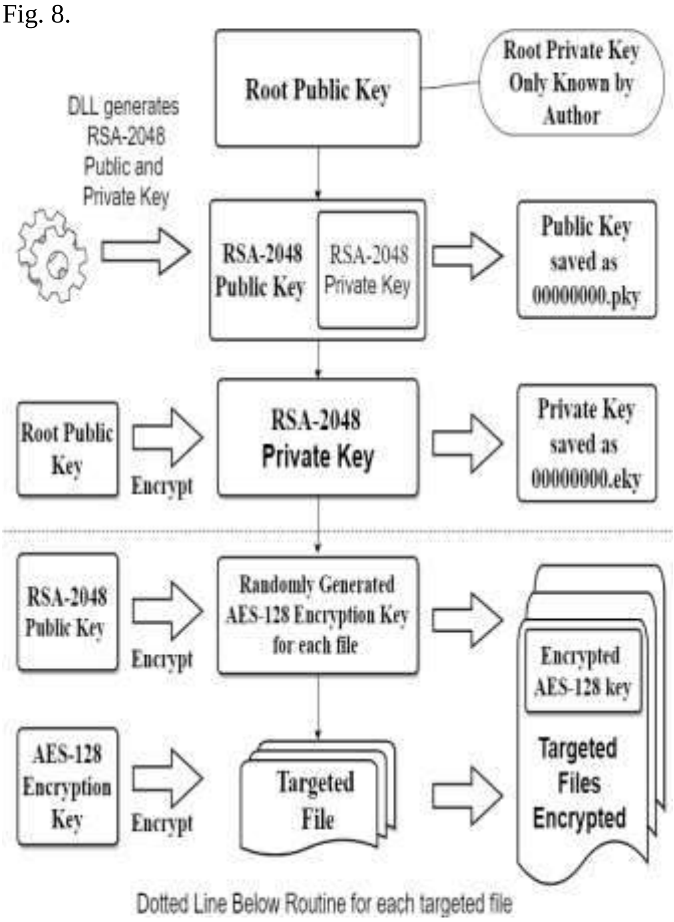


Fig. 7. The encrypting flow and key structure

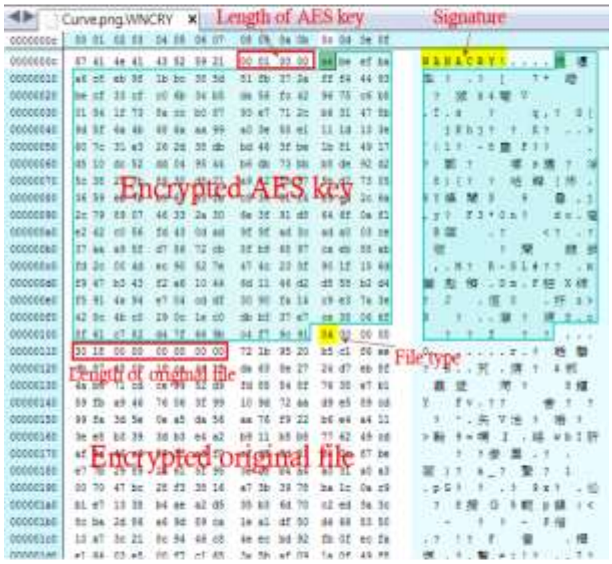


Fig. 8. The file structure of encrypted files

1) ***Delete or Erase Original Files.*** For the original files, the WannaCry will either delete or erase them depending on the file location in the system. WannaCry uses the wiping technique to prevent the user from restoring the files that are saved in the following folders: user desktop, user document, all user desktop, and all user document. The CryptGenRandom() function is called to generate random numbers, which are used to overwrite the original file contents. WannaCry will directly delete the files or move to a hidden recycle folder. The file deleting function taskdl is responsible for the deleting routine.

2) ***Demonstrate Decryption.*** When the machine is compromised with WannaCry, a random number of encrypted files in the folder C:\ProgramData\<randomized unique ID>\f.wnry can be decrypted for free as a demonstration. The user can retrieve up to 10 original files in the decrypting demonstration. There is no guarantee that all the encrypted files can be decrypted through the ransom payment.

D. Command-and-Control Phase: Tracing of infection and management of payments

All actions require some form of command-and-control processes to determine the succeeding actions to be undertaken [7]. The "@WanaDecryptor@.exe" is the binary for the user interface, C&C communication, and volume shadow deletion. This binary can be run using one of three parameters: "fi", "co", or "vs". The malware installs the necessary library dependencies to execute the TOR service. If the "@WanaDecryptor@.exe" runs with the parameter "fi," it attempts to connect to the onion server (C&C) and send the user name, host name and some other information about the infected system. If the parameter "co" is delivered, it launches "taskhsvc" as a sub-process to communicate with the onion server. The response from the C&C server should include a unique bitcoin address,which will update the string in c.wnry. The onion server domains are listed in the "c.wnry" as follows:

1) ***gx7ekbenv2riucmf.onion***

2) ***57g7spgrzlojinas.onion***

3) ***xxlvbrloxvriy2c5.onion***

4) ***76jdd2ir2embyv47.onion***

5) ***cwwnhwhlz52maqm7.onion***

IV. Network Analysis on WannaCry Exploits

To dissect WannaCry's exploits, network analysis was conductd by examining the network packets that were observed between the propagated machine and an infected machine. To perform such analysis, a VMware Workstation was adopted to build two host-only machines and to configure them in the same LAN. The captured packets were analyzed through Wireshark.

Once a machine with an open NetBIOS port was observed, WannaCry will gain a TCP socket for port 445, connect to SMB socket, and obtain an SMB tree ID for later use. Another characteristic is WannaCry's transmission of three NetBIOS session setup packets to it. One has the proper IP address (192.168.135.131 in our experiment) of the machine being exploited. Others contain two IP addresses (192.168.56.20 and 172.16.99.5) hard-coded in the malware body. The phenomenon and characteristic of the hard-coded IP addresses were probed for the target system's exploit status [11].

A. MS17-010 SMB RCE Detection

The detection method of information disclosure was used to determine if the MS17-010 has been patched [11]. WannaCry connected to the IPC$ tree and attempted a transaction on FID 0. If the status returned is "STATUS_INSUFF_SERVER_RESOURCES", it indicated that the machine did not have the MS17-010 patch (Fig. 9).

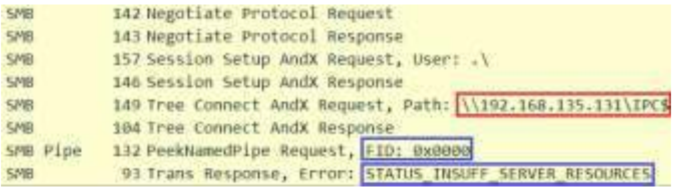


Fig. 9. The detection packets

B. SMB Doublepulsar Probe

The intent of the SESSION SETUP Trans2 Request was to verify if the system had already been compromised with the Doublepulsar backdoor (Fig. 10).

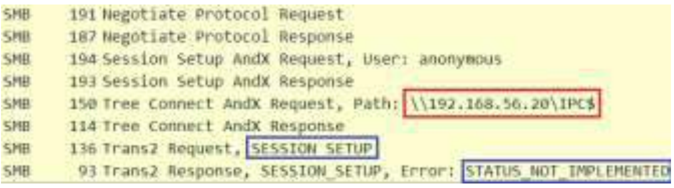


Fig. 10. The probing packets by Wireshark

If the field "Multiplex ID" is equal to 65(0x41), it indicates the current system is normal systems (Fig. 11). Otherwise, "Multiplex ID" that is equal to 81(0x51) indicates that the system has already been infected with Doublepulsar backdoor.

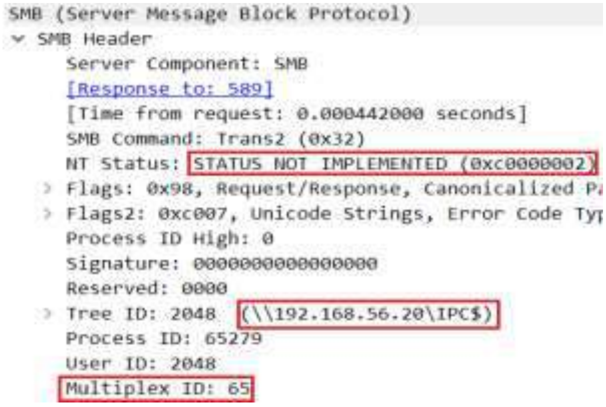


Fig. 11. The SMB Header of Trans2 Response

C. Triggering the Vulnerability

If the detection result shows that the target contains MS17-010 vulnerability and it is not yet infected with the Doublepulsar backdoor, it will proceed to install a Doublepulsar backdoor through the Eternalblue exploit (Fig. 12).

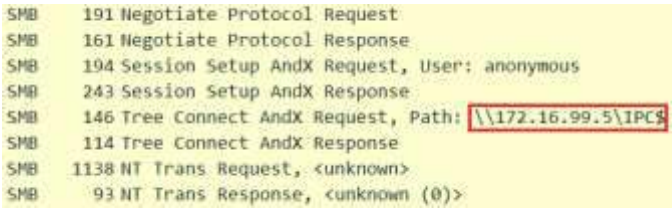


Fig. 12. The vulnerability triggering packets by Wireshark

An initial NT Trans request comprised a sequence of NOPs, which sought for the vulnerabilities in the compromised devices. The attacker could leverage a specialized-crafted packet to exploit targets' SMB protocol (Fig. 13). The large NT Trans request caused multiple Secondary Trans Requests and served as indicators for attackers to trigger the vulnerabilities.

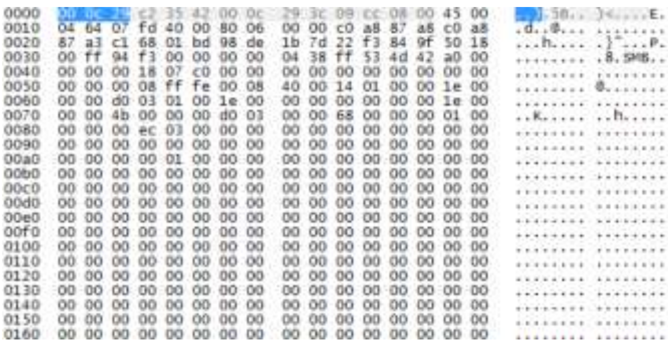


Fig. 13. The large sequence of NOPs

D. Doublepulsar Instrcution

After the completion of Eternalblue attack, the execution control was transferred to the Doubepulsar backdoor. A series of SMB packets were transmitted between the WannaCry propagating machine and the targeted victim, and the Doublepulsar instructions were hidden in specific fields (Fig. 14 and Table III).

Fig. 14. Doublepulsar instruction process

TABLE III
Doublepulsar Instruction Details

Doublepulsar Instruction	**Details of Instruction**		
	opcode	**Hidden Field**	**Description**
Ping Request	0x23	Timeout	Check if Doublepulsar backdoor exits.
Ping Response: Infected	0x81	Multiplex ID (MID)	Respond from Doublepulsar backdoor
Exec Request	0xC8	Timeout	Upload payload and inject.
Exec Response: Completed	0x82	Multiplex ID (MID)	Complete

1) ***Ping Request.*** After the initial negotiation and session setup, WannaCry will send a ping request to the target by sending multiple ping packets to the compromised system. The purpose of ping request was to check if the hook of Doublepulsar was installed successfully. The "ping" instruction was hidden in the "Timeout" field, which was originally the amount of time that the client had to wait for the server to respond to an outstanding request (Fig. 15). According to the Microsoft Open Specifications, the default value of Timeout field was set to 45 seconds. In the WannaCry network packet, the Timeout field was set to 4 hours 20 minutes 10.881 seconds (0x00ee3401). This abnormal Timeout value did not actually refer to the time out set but it implied the Doublepulsar instruction opcode. The algorithm of calculating this opcode is adding each byte and removing the overflow as result. If the Doublepulsar backdoor has successfully installed on the infected system, it will send back a crafted packet with "Multiplex ID" field purposely set..

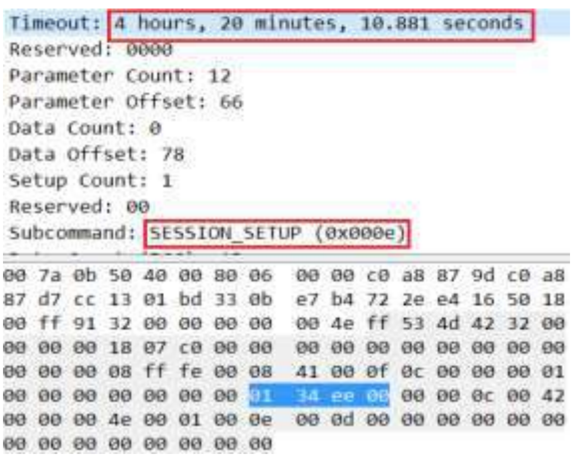


Fig. 15. "Ping" command in hidden Timeout field

2) ***Ping Response: Infected.*** While the Doublepulsar backdoor responded to the "ping" command with the field "Multiplex ID (MID)" set to 0x81, it implied the presence of itself. This packet had another implication using the "Signature" field (Fig. 16), which was set to value 0x011f7a1332. For little-endian, the first byte was set to 0x01, which indicated the machine was developed on an x64 platform. The WannaCry will prepare the payloads according to this probing result. For the remaining four bytes (0x1f7a1332), the encrypted XOR key was used to encode the payload during the uploading stage. The XOR key decrypting routine was conducted before WannaCry started using the XOR key to encode payload. The decrypting algorithm is demonstrated through IDA Pro reversing (Algorithm2).

Algorithm 2: (a1 = encrypted XOR key)
Decrypted XOR key = 2 * a1 ^ ((((a1 >> 16) | a1 & 0xFF0000) >> 8) | (((a1 << 16) | a1 & 0xFF00) << 8))

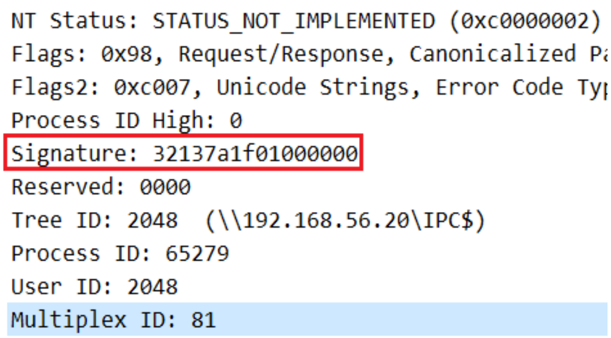


Fig. 16. Hidden Response in MID field with Signature field set to contain XOR key

3) ***Exec Request.*** After the confirming the presence of the backdoor, WannaCry will resume sending the "exec" Doublepulsar command to the target and ordered the backdoor on the target to start the injection of the ransomware into the lsass process. As indicated in the "ping" command, the packet set the "Timeout" field to an abnormal value. In the "exec" command, the "Timeout" field was again set to the value 0x001a8925 (Fig. 17).

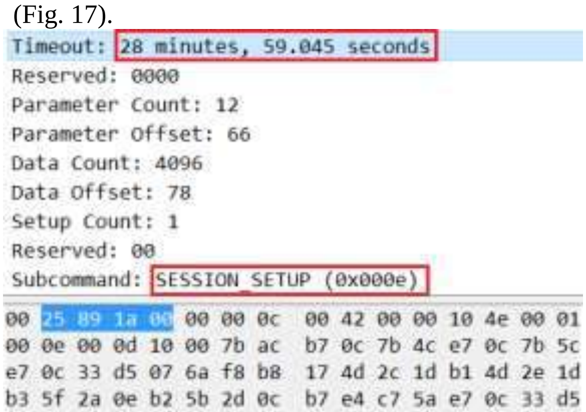


Fig. 17. "Exec" command hidden in Timeout field

4) ***Exec Response: Completed.*** As the shellcode completed, the Doublepulsar backdoor will send a packet with the field MID set to 0x82, to signal the completion of the task (Fig. 18).

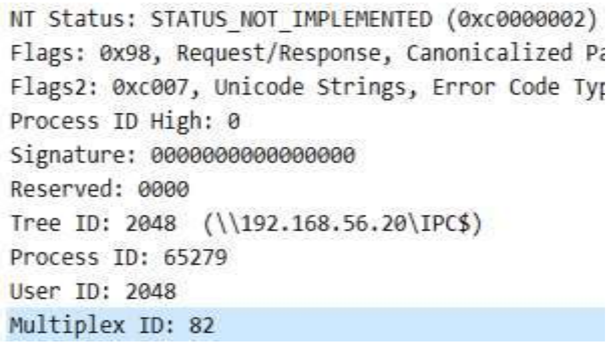


Fig. 18. Task completed

V. Research Finding

A. Multi-Phased WannaCry Execution

The increasing modularization of ransomware has encouraged security researchers to explore the inner design of each binary module. The typical resource extracting module used by WannaCry is the key to enable multi-phased execution. The exploit module provides WannaCry with an opportunity to rapidly propagate across the Internet. The WannaCry ransomware follows an execution flow when it gains access to a system and starts the propagation and encryption of files. Moreover, it inflicts damage by executing a series of tasks [10]. A cyclical life cycle exists throughout the entirety of the WannaCry code (Fig. 9). In this research, the anatomy of ransomware attack has been grouped into four phases in this research finding [1, 7]: *deployment, installation, destruction, and command-and-control.* In each phase, the component behavior is determined by the process parameter. The processes with their parameters are summarized in Table IV.

1) ***Deployment Phase.*** The launcher.dll is remotely injected into the lsass process through the infamous Eternalblue exploit and Doublepulsar backdoor. Launcher exports the PlayGame function, which uses resource-manipulation API functions, such as FindResource, LoadResource, LockResource, and SizeofResource to initialize the embedded mssecsvc binary in the launcher.dll resource section. And then, the mssecsvc process is launched through the path of “C:\Windows\mssecsvc.exe”.
2) ***Installation Phase.*** This phase comprises two components, “mssecsvc.exe” and “tasksche.exe”. The mssecsvc.exe starts up the mssecsvc2.0 service for propagation and drops the “tasksche.exe” using the same resource-manipulation API functions and routines in the deployment phase. The tasksche.exe is responsible for resource loading, environment setting, and the decryption of t.wnry. The “mssecsvc.exe” starts propagating while the parameter “m security” is identified. The propagation process has been categorized into the following four stages: *MS17-010 SMB RCE detection, SMB Doublepulsar probe, triggering the vulnerability, and Doublepulsar instruction*.
3) ***Destruction Phase.*** In the destruction phase, tasksche decrypts t.wnry from its resource section and loads the encryption dll in memory to execute the tasks. The encryption dll exports TaskStart to initiate the encryption. The management of the key system creates a complex encryption knot. For each encrypted target, an AES-128 encryption key is generated, which is also encrypted by the public key read from the 00000000.pky

Table IV

Execution Phase with Main Processes and their Features

Execution Phase	Processes	Features	
		parameter	*operation*
Deployment	launcher dll in lsass	*N/A*	Export PlayGame which loads resource into the mssecsvc binary and launch it.
Installation	mssecsvc	*N/A*	Install mssecsvc2.0 service for propagation and load resources into the tasksche binary before launching it.
		m security	Scan for devices both locally and on the Internet, exposes port 445, exploits the MS17-010 vulnerability, and installs the Doublepulsar backdoor.
	tasksche	*i*	Imply the installation mode of tasksche. It first creates a working directory C:\Windows\ProgamData\<randomized_id>\tasksche.exe to store its released binaries. After installation, the tasksche will then be executed without parameters.
		N/A	Create the registry key: HKLM\Software\Wow6432Node\Microsoft\Windows\CurrentVersion\Run\ <randomized_id>, release XIA resource, get TOR configuration from c.wnry, and run command “attrib +h” and “icacls . /grant Everyone:F /T /C /Q”.
Destruction	encryption dll in tasksche	*N/A*	Decrypt t.wnry into encryption dll and export TaskStart to run.
Command-and-Control	@WanaDecryptor@	*N/A*	Present the ransomware user interface.
		fi	Attempt to connect to the onion server (C&C) in the dark web and send the user name, host name, and some information about the infected system. The response may include an updated bitcoin address in c.wnry.
		co	Launch “taskhsvc” as sub-process to do the communication with onion server (C&C) and send some information about encrypting the users’ files from 00000000.res, including end time of encryption, the amount, the size of encryption
		vs	Delete volume shadow copies utilizing the Windows built-in vssadmin utility. It will launch the following command as sub-process “vssadmin.exe delete shadows /all /quiet” to implement the shadow deleting utililty.

4) **Command-and-Control Phase.** @WanaDecryptor@.exe is responsible for command-and-control. WannaCry tracks the payment and transmits the encryption information back to the onion servers in the command-and-control phase. The main execution flow is shown in Fig. 19.

B. WannaCry exploit signatures

During the initial exploitation, WannaCry will do the SMB tree connection, which contains the packet contents with the SMB header of "SMBr" (0x534D4272), "SMBs" (0x534D4273), "SMBu" (0x534D4275), and "SMB2" (0x534D4232). In addition, two hardcoded IP addresses are used to do the null connection for information disclosure. Therefore, the unique patterns of packets, and the hardcoded IP addresses can be used to generate the Yara rule (Rule: WannaCry_exploits).

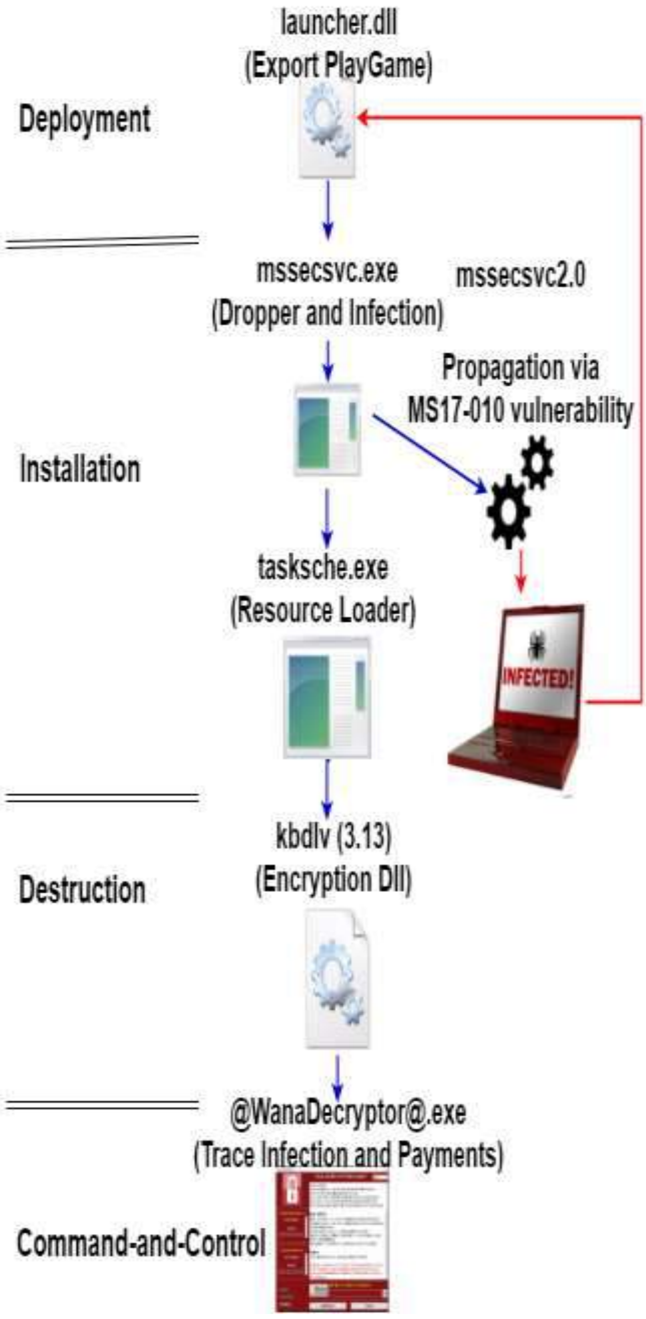


Fig. 19. Main execution flow of WannaCry ransomware

```
Rule: WannaCry_exploits{
   meta:
     description = "Detect WannaCry propagation"
   strings:
     $op1 = {53 4D 42 72 00 00 00 00 18 53 C0 00 00 00 00 00 00 00 00 00 00 00 00 00 00 FF FE 00 00 40 00  00 62 00 02 50 43 20 4E 45 54 57 4F 52 4B 20 50  52 4F 47 52 41}
     $op2 = {53 4D 42 73 00 00 00 00 18 07 C0 00 00 00 00 00 00 00 00  00 00 00 00 00 00 FF FE 00 00 40 00 0D FF 00 88  00 04 11 0A 00 00 00 00 00 00 00 01 00 00 00 00}
     $op3 = {53 4D 42 75 00 00 00 00 18 07 C0 00 00 00 00 00 00 00 00 00 00 00 00 00 00 FF FE 00 08 40 00 04 FF 00 5C 00 08 00 01 00 31 00 00 5C 00 5C 00 31 00 39 00 32}
     $op4 = {53 4D 42 32 00 00 00 00 18 07 C0 00 00 00 00 00 00 00 00 00 00 00 00 08 FF FE 00 08 41 00 0F 0C 00 00 00 01 00 00 00 00 00 00 00 01 34 EE 00 00 00 0C 00}

     $s1 = "\\\\192.168.56.20\\IPC$" fullword wide
     $s2 = "\\\\172.16.99.5\\IPC$" fullword wide

   condition:
       uint16(0) == 0x5a4d and all of ($s*) and 2 of ($op*) and pe.imports("ws2_32.dll", "connect") and pe.imports("ws2_32.dll", "send") and pe.imports("ws2_32.dll", "recv") and pe.imports("ws2_32.dll", "socket")
}
```

VI. CONCLUSIONS

The WannaCry outbreak is a significant security incident that spurs everyone to seriously consider the fundamentals of patching computers to current status. As malware developers tend to apply modular hacking weapons to new variants of malware, the detection technique adopted by security defenders becomes more granular based on the composed hacking weapon binaries. A thorough malware analysis was conducted to (i) identify the malicious binary, (ii) examine the exploits, (iii) collect malicious patterns, (iv) understand the indicators of compromised situation, and (v) report the observations to ensure the formulation of future defense strategies.

This paper conducted the reverse engineering analysis on WannaCry's components, and a network analysis on the WannaCry exploits. The modular hacking weapon in each component and its execution flow were dissected and analyzed. In addition, the techniques used by WannaCry exploits were unveiled by examining the packet details. The research findings, including representative hacking weapon modules and network signatures, can be documented to develop future defense strategies. The trend of integrating the artificial intelligence into unknown malware detection has become a popular issue. It may become a possible extension of our research for future ransomware detection based on the features of these hacking weapons.

ACKNOWLEDGMENT

The authors would like to thank Enago for the English language review.

REFERENCES

[1] Awad, R. A. and Sayre, K. D., "Automatic Clustering of Malware Variants," *2016 IEEE Conference on Intelligence and Security Informatics (ISI 2016)*, pp. 298–303, 2016.

[2] Ceron, J. M., Margi, C. B., and Granville, L. Z., "MARS: An SDN-based Malware Analysis Solution," *2016 IEEE Symposium on Computers and Communication (ISCC)*, pp. 525–530, June 2016.

[3] Fujino, A., Murakami, J., and Mori, T., "Discovering Similar Malware Samples Using API Call Topics," *2015 12th Annual IEEE Consumer Communications and Networking Conference (CCNC)*, pp. 140–147, Jan 2015.

[4] Hansen, S. S., Larsen, T. M. T., Stevanovic, M., and Pedersen, J. M., "An Approach for Detection and Family Classification of Malware Based on Behavioral Analysis," *2016 International Conference on Computing, Networking and Communications (ICNC)*, pp. 1–5, Feb 2016.

[5] Islam, A., Oppenheim, N., and Thomas, W., "SMB Exploited: WannaCry Use of Eternalblue." [Online]. Available: https://www.fireeye.com/blog/threat-research/2017/05/smb-exploited-WannaCry-use-of-Eternalblue.html

[6] Kharaz, A., Arshad, S., Mulliner, C., Robertson, W., and Kirda, E., "UNVEIL: A Large-Scale, Automated Approach to Detecting Ransomware," *25th USENIX Security Symposium (USENIX Security 16)*, pp. 757–772, USENIX Association, 2016.

[7] Liska, A. and Gallo, T., *Ransomware: Defending Against Digital Extortion*, 1st Edition, O'Reilly Media Inc., pp. 1-22, 2016.

[8] Microsoft, "Microsoft Security Bulletin MS17-010 - Critical: Security Update for Microsoft Windows SMB Server (4013389)." [Online]. Available: https://technet.microsoft.com/en-us/library/security/ms17-010.aspx

[9] Mosli, R., Li, R., Yuan, B., and Pan, Y., "Automated Malware Detection Using Artifacts in Forensic Memory Images," *in 2016 IEEE Symposium on Technologies for Homeland Security (HST)*, pp. 1–6, May 2016.

[10] Rousseau, A., "WCry/WanaCry Ransomware Technical Analysis." [Online]. Available: https://www.endgame.com/blog/wcrywanacry-ransomware-technical-analysis

[11] Rudman, L. and Irwin, B., "Dridex: Analysis of the Traffic and Automatic Generation of IOCs," *2016 Information Security for South Africa (ISSA)*, pp. 77–84, Aug 2016.

Da-Yu Kao received the B.S. and M.S. degree in information management from Central Police University, Taiwan, in 1993 and 2001, the Ph.D degrees in Crime Prevention and Correction from Central Police University, Taiwan, in 2009, respectively. From 1993 to 1996, he was with Taipei City Police Department, Taiwan, where he was an information technology police officer involved in the development of policing information systems. From 1996 to 2007, he was with Criminal Investigation Bureau, National Police Administration, Taiwan, where he was a detective and forensic police officer in cybercrime investigation and digital forensics. From 2007 to 2013, he was with Maritime Patrol Directorate General, Coast Guard Administration, Taiwan, where he was an information technology section chief in the department of information and communication. Since 2013, he has been with Central Police University, Taiwan, where he is currently an associate professor in the Department of Information Management. His research interests include cybercrime investigation, digital forensics, digital evidence, information management, criminal profiling, and cyber criminology.

Shou-Ching Hsiao received the B.S. degree in information management from Central Police University, Taiwan, in 2016. Since 2016, she has been with Haishan Precinct, New Taipei City Police Department, Taiwan, where she is currently an information lieutenant and is responsible for information system management, information security, malware analysis, and real-time video for security control. In 2018, she is also working toward the M.S. degree in the Department of Computer Science, National Chengchi University, Taiwan. Her current research interests include malware analysis, cybercrime investigation, digital forensics, digital evidence, information management, criminal profiling, and cyber criminology.

Raylin Tso received the B.S. degree in Industrial Engineering from National Tsing Hua University, Hsinchu, Taiwan and M.S. degree in Management Science - Division of Management and Public Policy, University of Tsukuba, Tsukuba, Ibaraki, Japan, in 2002, M.S. degree in Risk Engineering, Graduate School of Systems and Information Engineering, University of Tsukuba, Tsukuba, Ibaraki, Japan, in 2004, the Ph.D degrees in Risk Engineering, Graduate School of Systems, University of Tsukuba, Tsukuba, Ibaraki, Japan, in 2006, respectively. From 2006 to 2008, he was with Graduate School of Systems and Information Engineering, University of Tsukuba, Japan, where he was an academic researcher. Since 2008, he has been with National Chengchi University, Taiwan, where he is currently an associate professor and chairman in the Department of Computer Science. His research interests include cryptography, information security, post-quantum, blockchain and privacy enhancement.

Extracting Suspicious IP Addresses from WhatsApp Network Traffic in Cybercrime Investigations

Da-Yu KAO*, En-Cih CHANG*, Fu-Ching TSAI**

*Department of Information Management, Central Police University, Taoyuan 333, Taiwan

** Department of Criminal Investigation, Central Police University, Taoyuan 333, Taiwan

dayukao@gmail.com, dorislovesnoopy@gmail.com, fctsai@mail.cpu.edu.tw

***Abstract*—Sniffers are among the commonest approaches for capturing network traffic activities and collecting digital evidences in cybercrime investigations. The ubiquity of instant messaging (IM) apps on smartphones has provided criminals with communication channels that are difficult to decode. Moreover, investigators and analysts of cybercrimes are encountering increasingly large datasets. To combat criminal activity, law enforcement agencies (LEAs) often rely on call-record analysis. In this paper, cybercriminals are investigated by network forensics and sniffing techniques. Retrieving valuable information from specific IM apps is difficult because the criminal's IP address records are not easily recognisable on the Internet. Here, a criminal's identity is located more effectively by a packet filter framework that isolates the WhatsApp communication features from huge collections of network packets. A rule extraction method for sniffing packets is proposed that retrieves the relevant attributes from high-dimensional analysis based on geolocation and a pivot table. The utility of this methodology is illustrated on real-time network forensics and a lawful interception system in Taiwan. The methodology also meets the ISO/IEC 27043:2015 standards of fear, uncertainty, and doubt avoidance. Besides supporting LEAs in discovering criminal communication payloads, prosecuting cybercriminals and bringing them to justice, it improves the effectiveness of modern call-record analysis.**

***Keyword*—Cybercrime Investigation, Network Forensics, Packet Analysis, VoIP, WhatsApp, Lawful Interception, ISO/IEC 27043: 2015**

Manuscript received Dec. 19, 2017. This work was a follow-up of the invited journal to the accepted & presented paper of the 20th Conference on Advanced Communication Technology (ICACT2018), and this research was partially sponsored by the Executive Yuan of the Republic of China under the Grants Forward-looking Infrastructure Development Program (Digital Infrastructure-Information Security Project-107) and the Ministry of Science and Technology of the Republic of China under the Grants MOST 107-2221-E-015-002.

Da-Yu Kao is with the Department of Information Management, Central Police University, Taoyuan 333, Taiwan (Corresponding Author phone: +886-3-328-2321; fax: +886-3-328-5189; e-mail: dayukao@gmail.com).

En-Cih CHANG is with the Department of Information Management, Central Police University, Taoyuan 333, Taiwan (phone: +886-3-328-2321; fax: +886-3-328-5189; e-mail: dorislovesnoopy@gmail.com).

Fu-Ching TSAI is with the Department of Criminal Investigation, Central Police University, Taoyuan 333, Taiwan (phone: +886-3-328-2321; fax: +886-3-328-5189; e-mail: fctsai@mail.cpu.edu.tw).

I. Introduction

WhatsApp is a cross-platform application enabling instant communications on electronic devices such as smartphones, tablet computers and personal computers. More than 1.5 billion active WhatsApp users were estimated in December 2017 [11]. The worldwide popularity of WhatsApp is attributable to a range of attractive features at low subscription cost. New features allow people to group chat and send texts, pictures and other multimedia elements along with their messages. Since WhatsApp was acquired by Facebook in 2014, more users have communicated through this platform by the snowball effect [11]. Unfortunately, the convenience and high functionality of WhatsApp has facilitated effective and secret communications among criminals. The present study attempts to recognise WhatsApp communication features among huge collections of network logs and packets, and thereby locate criminal activities more effectively. Discovering criminal communication contents among vague connections helps law enforcement agencies (LEAs) to better filter criminal activities.

Call-record analysis ranks among the critical criminal investigation strategies of LEAs. Call records provide important information for crime-scene investigations, such as the dates, times, and lengths of outgoing and incoming calls [1]. However, the ubiquity of instant messaging (IM) apps on smartphones has provided criminals with communication channels that are difficult to track by traditional investigation technologies. Nowadays, most criminals communicate through IM apps rather than voice phones to prevent detection by LEAs. Identifying a cybercriminal without the help of foreign authorities is difficult on the Internet, which provides complete anonymity and privacy and consequently hinders an investigation [2]. New techniques for analysing modern call records are urgently required.

The main difficulty of retrieving valuable information from specific IM apps is filtering the massive volume of network connection records on the Internet. Raw data captured from the Internet are full of packets produced by different apps from various devices, each with differing protocols, ports, and connection frequencies. Moreover, smartphones can establish connections through different network interfaces. Despite the challenges of retrieving call records or network connection logs from smartphones, Internet data provide more advanced and detailed information than traditional phone records. For example, the geographic information system or Internet protocol (IP) address reveals the call locations, while the captured network packets provide the multimedia content of the communications.

The remainder of this paper is organised as follows. Section 2 reviews packet analysers, the Voice-over Internet Protocol (VoIP), and WhatsApp. Section 3 describes the research design. Section 4 proposes a cybercrime investigation framework of network traffic compliant with ISO/IEC 27043:2015, and experimentally demonstrates its effectiveness. The last section concludes the paper and suggests ideas for future work.

II. *Literature Review*

A. Packet Analysers

Packet analysers are widely applied to raw-traffic analysis, attack detection, sniffing and network troubleshooting in the network security field [6]. As shown in Fig. 1, a packet analyser performs several functions [3]: reverse engineering, storing and accessing packets, detecting improper data transfer, monitoring network statistics, assisting intrusion detection systems, and handling network problems. Packet analysers can play different roles in various applications. From a moral perspective, packet analysers assist with security audits of data packets; for network administrators, they provide diagnostic tools for network problems. White-hat hackers study the reports of packet analysers to find vulnerabilities in software applications, and thereby issue an early warning before cyber-attackers can launch serious attacks. Protocol developers use packet analysers to diagnose protocol-related issues. Packet analysers can also be used in immoral ways, for example, inspecting packet payloads to decrypt passwords or sniffing traffic to deploy a man-in-the-middle attack. Packet analysis is the process of capturing and interpreting live data flowing across a network, and hence understanding the network dynamics. Most packet analyses are performed by a packet sniffer, which captures the raw network data traversing wires or wireless interfaces. Packet analysis can help with understanding the network characteristics, determining who or what is utilising the available bandwidth, finding unsecured and bloated applications, identifying summit network usage times, and detecting malicious activities.

Packet-sniffing programs are varied in type, and can be free or commercial. Each program is designed for different goals. A few popular packet-analysis programs are Tcpdump, OmniPeek, and Wireshark. Tcpdump is a command-line program, while OmniPeek and Wireshark have graphical user interfaces (GUIs) [10]. Wireshark, one of the most well-known open-source packet analysers, provides both an easy-to-use GUI and a command-line utility with very active community support [7]. It also supports offline and online modes for flexible capturing operations. The features of Wireshark are live-packet capture, a user-friendly GUI and command-line interface, data filtering, GNU open-source software, generation of various statistics, and decoding of sets of protocols [7] (see Fig. 2).

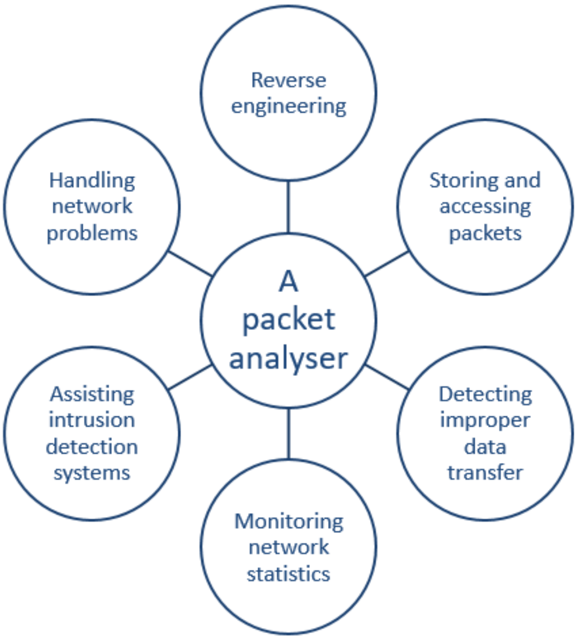


Fig. 1. Functions of a packet analyser

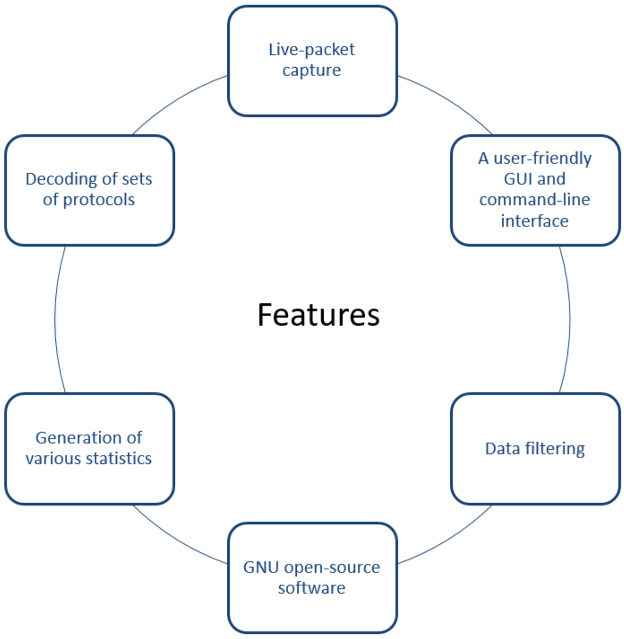


Fig. 2. Features of the Wireshark packet sniffer

B. VoIP and WhatsApp

1) VoIP

VoIP, which sends voices over an IP-based network, totally differs from circuit-switched public telephone network [8]. Whereas circuit switching allocates resources to each individual call, IP networks are packet switched. Each packet is semi-autonomous, with its own IP header and forwarded separately by the routers. VoIP manages the signalling, set-up, and tear-down of calls by session control and signalling protocols. It cooperates with several protocols such as Session Initiation Protocol, H.323, Session Description Protocol, Real-time Transport Protocol, and Inter-Asterisk eXchange. A traditional system requires much control signalling to accomplish the various tasks, but VoIP collects these signalling messages and places them inside IP packets.

It is worth mentioning that because an IP can and does run over almost all types of low-layer communication architectures, VoIP can as well. Researchers can compare the topologies of different VoIP architectures, and can short-list the basic skills required to work on VoIP and traditional telephony. Both VoIP and telephony serve the same functions with the same equipment, but using different techniques with completely different sets of protocols [3].

2) WhatsApp

Network sniffing is a vital strategy in modern crime investigation [1]. With the rapid evolution of the Internet, communication has transformed from traditional phone calling to network-based VoIP interactions. The low cost and interactive features (with delivery of multimedia elements) of IM applications have encouraged a large number of users to almost abandon traditional phones. The most commonly used feature of WhatsApp is voice calling. When a user starts a call to a private IP address behind a network address translation (NAT) firewall, the packet routing should be assisted by a STUN (Session Traversal Utilities for NAT) protocol, which allows the end computer to discover the public IP address, and permits NAT traversal of real-time voices, messages, and other interactive communications [9]. The anonymous nature of the Internet limits the abilities of LEAs to monitor the communications of criminal activities. Therefore, efficient network sniffing technologies are demanded for cybercrime investigation.

3) ISO/IEC 27043:2015

The purpose of network sniffing is to discover criminal activities. To bring criminals to justice, the integrity of digital evidence should be maintained by procedures that collect and analyse network packets. The 2015 ISO/IEC 27043 standard provides readiness, initialisation, acquisitive, and investigative guidance for criminal investigations [4]. However, the ISO standards have been rarely applied in practical solutions. This study simulates a network sniffing scene that collects packets between the victim and suspect following the recommendation processes in ISO/IEC 27043: 2015. The present paper demonstrates the framework of the network sniffing strategy for LEAs operating under lawful interception warrant procedures.

III. RESEARCH DESIGN

This paper simulates the communications between the victim and suspect, and extracts the likely incriminating features in the communication. Using these features, it proposes filtering rules by which LEAs can effectively target suspects in WhatsApp packets. Our research design comprises four phases: data collection, data preparation, feature recognition, and result evaluation (see Fig. 3).

A. Research Experiment

The sniffing of WhatsApp voice calls, collection of IP address information, and personal identification of the WhatsApp application target, were conducted in a controlled environment.

1) Software Environment

Within the experimental environment, the transmission time and packet size were controlled by varying the bandwidth and traffic congestion. All devices were initially configured as follows [5]:

a) Victim: Cellphone at Domain A

- Android Operating System v5.0
- WhatsApp Ver. 2.17.146

b) Investigator: Computer at Domain A

- Wireshark v2.2.5
- Windows 10.0.14393
- Excel 2013
- I2 Analyst's Notebook 8 v8.5.5
- NodeXL Basic Excel Template 2014

c) Suspect: Cellphone at Domain B

- iOS 10.3.2
- WhatsApp Ver. 2.17.146

2) Participants

The experimental participants included a victim, an investigator and a suspect (Fig. 3). Domain A was configured by the investigator or the victim. Domain B was used by the suspect, criminal, or target. All communication packets were sniffed by Wireshark.

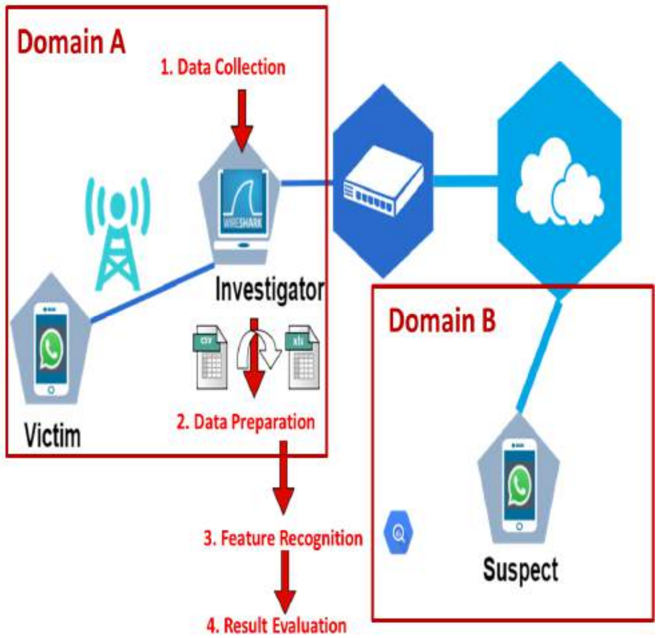


Fig. 3. Research design

B. Experimental Phases in the Research Design

The four experimental phases in our research design are discussed below.

1) Data Collection

The hotspot of the investigator computer in Fig. 3 shared its network connections with the victim's cellphone. The researchers monitored and eavesdropped (by copying) the traffic to and from the investigator's computer. The eavesdropping included the packets from the victim's phone to the Internet. By setting a midpoint in the investigator host, the researchers were able to use Wireshark, capture all network traffic, and investigate the criminal behaviour along the victim–suspect route. Routine data transmission was prioritised over the copying process. This priority might have caused dropped Ethernet frames when collecting the incriminating evidence.

2) Data Preparation

To capture general traffic, the researchers installed the packet-sniffing software, configured the network interface controller (NIC) in promiscuous mode, and collected all network traffic addressed to the MAC address of the NIC. From the collected data, the researchers could overview the WhatsApp performance and tentatively identify the suspect. For this purpose, the data passing through the investigator computer were captured and analysed, then presented in an easy-to-read format.

3) Feature Recognition

Common tools for collecting network traffic, such as pcap (for Unix-like systems) and libcap (for Windows systems), collect thousands of small data packets that are sent across the Internet. Such numerous small packets can be difficult to navigate. The main purposes of the present study are listed below:

- Assess the overall traffic flow through the network
- Exactly copy the network traffic for predictive analysis
- Identify how WhatsApp applications generate the VoIP traffic
- Identify the IP address of the suspect WhatsApp user
- Highlight the features of the WhatsApp packets in the suspect's IP address

4) Evaluation Results

We monitored only the traffic to and from the investigator computer. While two users conducted voice calls through WhatsApp, the researchers assumed that the Wireshark deployment node was lawfully intercepted by the warrant procedures of the victim's agreement. To start a sniffing procedure, the investigator computer must be on the same network as the sniffed cell phone. Packets can be very useful for tracking suspects or offenders in cybercrime investigations.

IV. *Proposed Cybercrime Investigation Framework of Network Traffic*

The storage and handling of network traffic requires the processing of massive numbers of packets, maintaining the integrity of the digital evidence, and preserving the digital evidence during the investigative period. These requirements present significant challenges for LEAs. The ISO/IEC 27043: 2015 international standards provide instructional guidance for the readiness, initialisation, acquisitive, and investigative processes. Our network-based sniffer framework helps to address the above challenges and formalises what should be logged for an appropriate cybercrime investigation.

A. Materials and Methods

The collected digital evidence should increase the conviction rate and restore the truth. The following standardised procedures are vital to the validity and reliability of the collected digital evidence. The network-based sniffer experiments in this study were based on the ISO/IEC 27043: 2015 international standards of incident investigation processes. The various process classes are shown in Fig. 4 and discussed below.

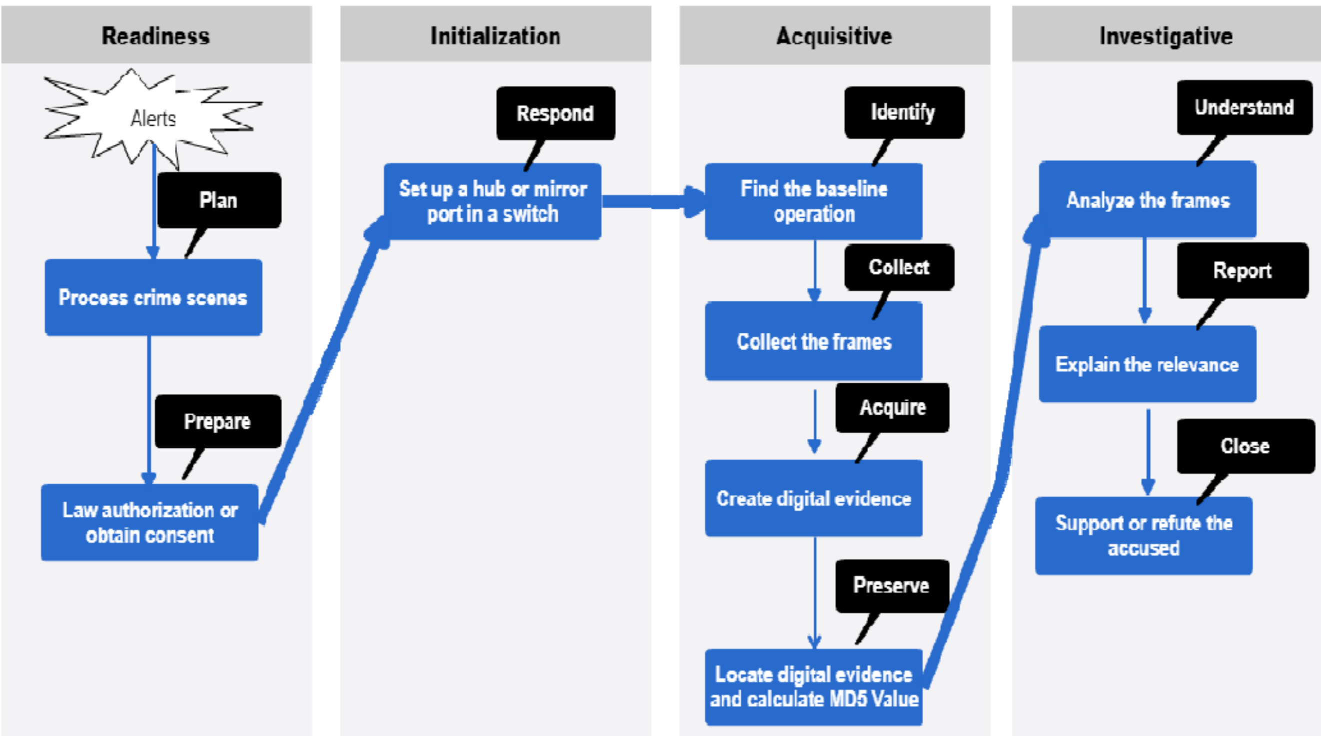


Fig. 4. Network-based sniffer framework for cybercrime investigation

B. ISO/IEC 27043:2015 Process Class

In recent years, network sniffing in criminal investigations has been conducted under lawful interception warrant procedures. Network sniffing poses great challenges to LEAs because unlike traditional call interception, it lacks any systematic procedure. The network sniffing framework proposed in this study is guided by the ISO/IEC 27043: 2015 standards. In particular, it follows the ten steps in ISO/IEC 27043: 2015 to preserve the integrity and prevent damage of the digital evidence. LEAs can adopt the framework as a standard operation procedure to facilitate an efficient network traffic analysis.

1) Readiness Process Class

a) Plan

The plan phase consolidates the scope and purpose of the investigation. Using Wireshark, the researchers captured all network traffic along the victim–suspect route. The WhatsApp network traffic was sniffed to identify the calling and receiving phones. The personal computer was configured as a hotspot for sharing network connections to the cell phone, and as the node for capturing the network packets utilised by Wireshark. The study included the detailed routing information, such as the IP address, protocol, time, and packet length. The investigative tasks were assisted by careful planning.

b) Prepare

Network sniffing is an interdisciplinary process. The team members responsible for this task should possess knowledge of packet analysis, technology devices construction and criminal investigation. Therefore, LEAs should provide multi-domain training courses for their team members. Good preparation ensures that criminal investigators can cope with various crime senses. Once the collection process is complete, the data integrity can be documented by the MD5 value of the pcap file, and the data can be preserved on a write-only medium [1].

2) Initialisation Process Class

a) Respond

One law enforcement strategy in criminal investigations is a dedicated middle node for packet sniffing. To this end, we set up a network sniffing framework that efficiently responds to a crime case. To handle the massive volume of network packets on the Internet, we erected a hub or mirror port in a switch that probed the routing nodes containing the targeted WhatsApp connections. The collected information provides LEAs with quick responses to various crime scenes.

3) Acquisitive Process Class

a) Identification

LEAs should convert the huge number of Internet packets to readable information. To identify the criminal activities, we imported the connecting information to the pivot table. Having identified the facts of the crimes, LEAs can process the investigation by various systematic approaches.

b) Collection

Internet packets were collected by Wireshark software. The collector should be placed en-route between the caller and receiver. Under the lawful interception warrant procedures, the network sniffing node should be the Internet data centre owned by either caller, or the telecommunication service provider of the receiver.

c) Acquisition

Having confirmed the sniffing nodes, the LEA should deploy the packet analyser that collects the network packets. Most of the packet analysers store the packets in their own formats. To analyse the payload information more effectively, we imported the files produced by various packet analysers into a normalised database table using extract–transform–load tools.

d) Preservation

Digital evidence is commonly acquired by live investigation or dead forensics. Live investigation is conducted on a system running at the scene, and dead forensics is usually performed in a trusted laboratory environment. In both investigation modes, the data should be preserved to maintain its integrity. In this study, the data integrity was verified by the MD5 hash value.

4) Investigative Process Class

a) Understanding

This stage analyses the collected digital evidence. The modus operandi of criminal activities in a huge database is probed by forensic tools. Open-source toolkits, data mining, and machine learning approaches that reflect the features or contextual information in a crime case, are also available.

b) Reporting

A criminal investigator must document the processes and results of the case. The report should not only detail the crime case, but should also provide testimony in court. Moreover, it should be precise, easily read, and clearly understandable. The report can also contain multimedia elements such as video, audio and pictures.

c) Close

After checking that all evidence is well-protected and safely deposited, the criminal investigation is closed. The storage should be regulated by strict rules, preventing the evidence from been changed, lost, stolen or destroyed.

C. Research Findings

1) Feature Recognition in Frequency Distribution Analysis

a) Frequency Distribution Analysis

The data packets were captured in the pcap file format, and imported to Wireshark for demonstrating their header and payload information. The pcap files were then exported to excel, where the high-dimensional data were viewed from different angles in a pivot table. To investigate the features of the WhatsApp communications, we imported the headers and payloads of the captured packets into the pivot table. In a frequency distribution analysis, most of the packet fields consisted of random values with no relevance to communication features. However, the values of several attributes, such as Differentiated Service Field, Flags, and Differentiated Services Codepoint, were fixed. These fixed-value attributes were selected as the criteria of feature recognition in the WhatsApp communications. The derived packet attributes and their contents are shown in Table 1.

TABLE I.
CRITERIA AND CONTENTS OF FEATURE RECOGNITION IN WHATSAPP COMMUNICATIONS

Packet Attribute	**Content**
Differentiated Services Field	0x38
Flags	0x00
Differentiated Services Codepoint	Assured Forwarding 13

b) Feature Recognition

The feature-recognition phase identifies the features generated by WhatsApp in the packet records. This phase increases the efficiency of finding the suspect's IP address. The present study found 90 attributes in the flat file (Fig. 5). After viewing the data from different angles in the pivot table, we identified the most important attributes of WhatsApp as "Differentiated Services Field", "Flags", and "Differentiated Services Codepoint". Differentiated services ensure low latency of the critical network traffic such as voice and streaming media while providing simple best-effort service to non-critical services such as web traffic and file transfers. The field value contained in the flag is often explained in the section related to data structure, and the bit field is usually associated with a property or privileges.

2) IP Geolocation in STUN Packets

a) STUN Packets

The STUN packets, which are related to voice call operations, were analysed by the STUN protocol- filtering function of Wireshark. Fig. 6 shows the list of STUN packets in the imported pcap files, ordered by timestamp.

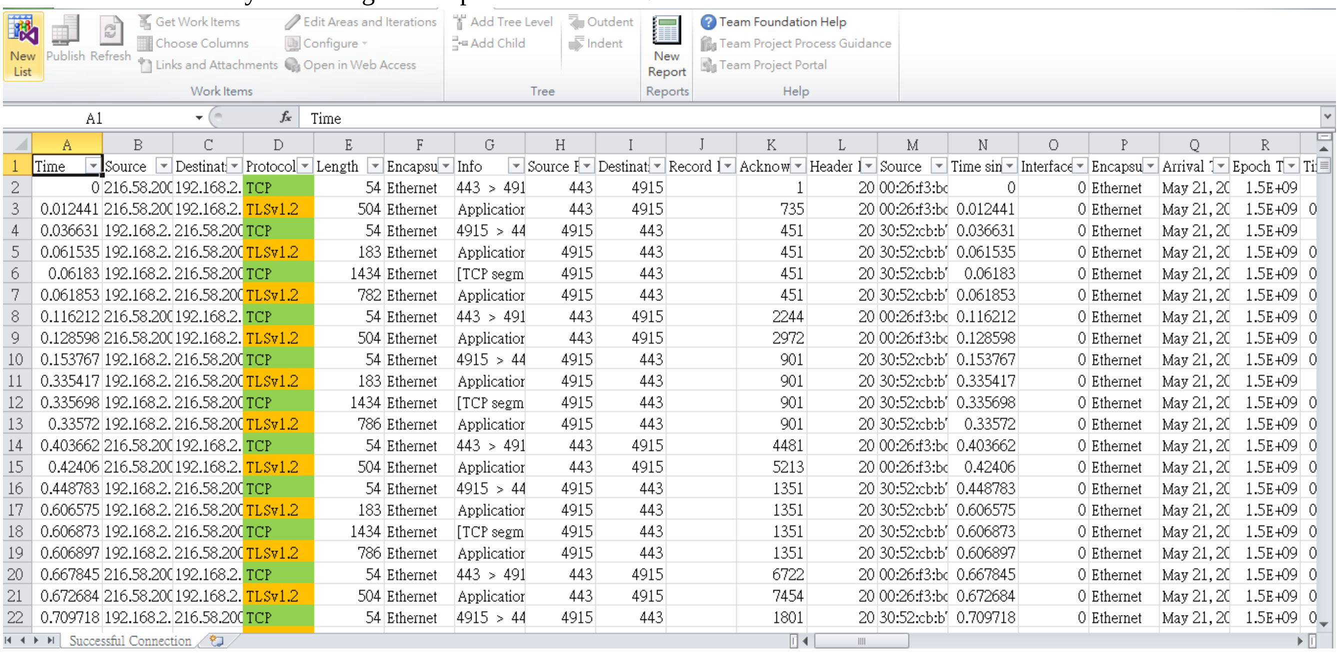


Fig. 5. A flat file for feature recognition

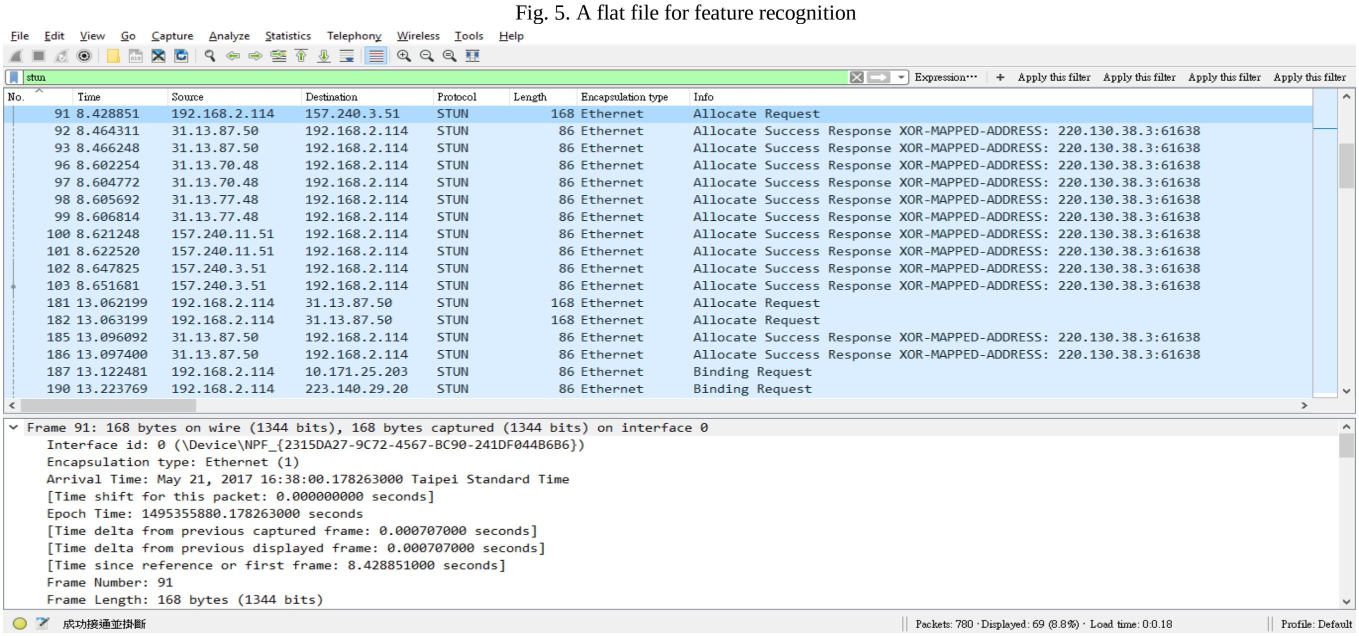


Fig. 6. Packets of STUN Protocol

b) IP Geolocation

The geolocation of an IP address is important for transforming the location from the network space to physical space. The Whois database (originally designed for Unix) has become the commonest mechanism for locating the registration information of IP resources registered in Internet-number resource organisations. After querying a registry, an open-source Whois, Lookup, or IP location tool returns rich geolocation information, including the domain ownerships, addresses, locations, and phone numbers of the queries. To evaluate the effectiveness of our proposed framework, we collected network packets during the 34.505698-second period containing the WhatsApp communications. Although the traffic was only captured from the local area network, the IP list was rendered complicated by additional connections with its software companies and Internet service providers. The geolocations of the IP addresses in this experiment were transformed by whois lookup tools and are listed in Table 2.

TABLE II.
GEOLOCATIONS OF THE IP ADDRESSES IN THE PRESENT EXPERIMENT

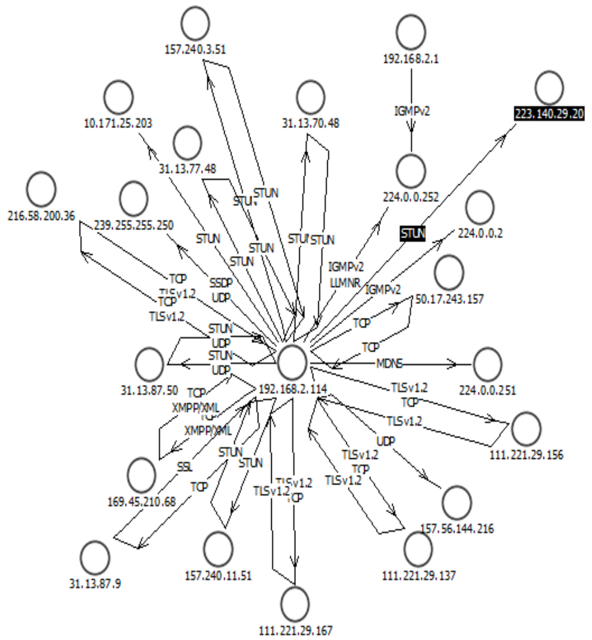
IP Address	Whois Lookup
192.168.2.114	Victim
224.140.29.20	Suspect(EMOME-IP.hinet.net, Taiwan)
157.240.4.51	United States Menlo Park Facebook Inc.
31.14.87.50	Taiwan, Province Of China Taiwan, Province Of China Taipei Facebook Ireland Ltd
31.14.70.48	United States United States Los Angeles Facebook Ireland Ltd
157.240.11.51	United States United States Menlo Park Facebook Inc.

3) Observation Rules on Suspect IP Addresses

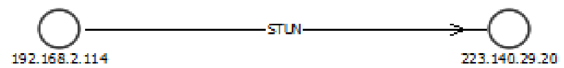
Several phenomena were observed in the WhatsApp network traffic (see Table 3). The rules were established to improve the efficiency of selecting WhatsApp communications from the Internet data. The rules should filter out the noise connections produced by the software companies, manufactures of network devices and ISPs. To demonstrate the effectiveness of our approach, we applied the generated rules to the collected packets containing the WhatsApp communications. Fig. 7 shows the network topology before and after implementing the rules. The proposed rules successfully pruned the complicated topology and revealed the victim–suspect connections. The above findings will assist LEAs in their cybercrime investigations of criminals contacting through WhatsApp.

TABLE III.
THE RULES

Rule 1	IF Differentiated Service Field = 0x38, Flags = 0x00, and Differentiated Services Codepoint =Assured Forwarding 13 THEN Source IP address = suspect
Rule 2	IF Protocol = STUN, Length = 86, and Info = Binding Request THEN Destination IP address = suspect



(a)Before the rule implementation



(b) After the rule implementation

Fig. 7. Network topology of WhatsApp communications after pruning by our proposed rules

V. CONCLUSION AND FUTURE WORKS

Modern call-record analysis is an expected future trend of criminal investigation strategies. Recognising the communication features of IM software on smartphones is essential for revealing the locations of suspects, providing clues that improve the efficiency of investigative work by LEAs. This study has developed a rule extraction framework that reveals the WhatsApp communications and eliminates the impact of disordered Internet connections. In an experimental test, the generated rules successively simplified the complexed network topology to simple connections between the suspect and victim. The criteria discovered in the sniffed packets provided instructive information for identifying the characteristics of WhatsApp communications. Furthermore, the proposed framework can explore the features produced by other IM software. Extracting the connections made by criminals will improve the prosecution and conviction rates by LEAs. To keep pace with the rapid developments of IM software, future research should consider the software upgrade problem, and minimise the impact of updating the IM software. The application of the proposed framework to encrypted communication should be examined from various perspectives. As the proposed methodology is intended to help LEAs, it should also provide a more complete coverage of IM feature-recognition applications. For this purpose, additional IM software should be considered in future work.

ACKNOWLEDGMENT

The authors would like to thank Enago for the English language review.

REFERENCES

[1] Casey, E., *Digital Evidence and Computer Crime.* Waltham, MA: Elsevier Inc., pp. 727-735, 2011.
[2] EC-Council, *Computer Forensics: Investigating Network Intrusions and Cyber Crime.* Boston, MA: EC-Council Press, pp. 27-60, 2010.
[3] Hartpence, B., *Packet Guide to Voice over IP: A System Administrator's Guide to VoIP Technologies.* Sebastopol, CA: O'Reilly Media Inc., pp. 2-5, 2014.
[4] International Organization for Standardization (ISO), "ISO/IEC 27043: 2015 Information Technology – Security Techniques - Incident Investigation Principles and Processes," Switzerland: ISO Office, pp. 5-20, 2015.
[5] Kao, D. Y. and Wu, W. Y., "Practical Packet Analysis: Exploring the Cybercriminal behind the LINE Voice Calls," 2017 19th IEEE International Conference on Advanced Communications Technology (ICACT), Pyeong Chaung, South Korea, Feb. 19-22, 2017, 2011.
[6] Kizza, J. M., *A Guide to Computer Network Security (3rd Edition).* Swindon, UK: Springer-Verlag London Ltd., pp. 299-324, 2015.
[7] Nath, A., *Packet Analysis with Wireshark.* Birmingham, UK: Packet Publishing Ltd., pp. 56-146, 2015.
[8] Rahbar, A. G., *Quality of Service in Optical Packet Switched Networks.* Danvers, MA: IEEE Press, pp. 19-43, 2015.
[9] Roy, R. R., *Handbook on Session Initiation Protocol: Networked Multimedia Communications for IP Telephony.* Boca Raton, FL: CRC Press, pp. 1-350, 2016.
[10] Sanders, C., *Practical Packet Analysis: Using Wireshark to Solve Real-World Network Problems (3rd Edition).* San Fransciso, CA: No Starch Press, pp. 53-102, 2017.
[11] Statista–the Statistics Portal, "the Number of monthly active WhatsApp users worldwide from April 2013 to December 2017 (in millions)." [Online]. Available: https://www.statista.com/statistics/260819/number-of-monthly-active-whatsapp-users/

Da-Yu Kao received the B.S. and M.S. degree in information management from Central Police University, Taiwan, in 1993 and 2001, the Ph.D degrees in Crime Prevention and Correction from Central Police University, Taiwan, in 2009, respectively. From 1993 to 1996, he was with Taipei City Police Department, Taiwan, where he was an information technology police officer involved in the development of policing information systems. From 1996 to 2007, he was with Criminal Investigation Bureau, National Police Administration, Taiwan, where he was a detective and forensic police officer in cybercrime investigation and digital forensics. From 2007 to 2013, he was with Maritime Patrol Directorate General, Coast Guard Administration, Taiwan, where he was an information technology section chief in the department of information and communication. Since 2013, he has been with Central Police University, Taiwan, where he is currently an associate professor in the Department of Information Management. His research interests include cybercrime investigation, digital forensics, digital evidence, information management, criminal profiling, and cyber criminology.

En-Cih CHANG received the B.S. degree in information management from Central Police University, Taiwan, in 2018. In 2018, she is studying in the College of Communication and Information, Florida State University, Tallahassee, FL, USA. Her current research interests include information security, incident response, cybercrime investigation, digital forensics, information systems management, criminal profiling, cyber criminology, and machine learning.

Fu-Ching TSAI received the B.S. degree in Information Management from Central Police University, Taiwan, in 2001, the M.S. and Ph.D degrees in Institute of Information Management from National Cheng Kung University, Taiwan, in 2005 and in 2012, respectively. From 2001 to 2010, he was with Pingtung County Police Bureau, Taiwan, where he was a lieutenant involved in the development of policing information systems. From 2010 to 2014, he was with National Police Agency, Taiwan, where he was a division assistant in network & security incident investigation. From 2014 to 2017, he was with Changhua County Police Department, Taiwan, where he was an Information Management Division chief. Since 2017, he has been with Central Police University, Taiwan, where he is currently an assistant professor in the Department of Criminal Investigation. His research interests include data mining, text mining, digital forensics, social network analysis, and cyber criminology.

Volume 7 Issue 2, March. 2018, ISSN: 2288-0003

ICACT-TACT

JOURNAL

GIRI

Global IT Research Institute

1713 Obelisk, 216 Seohyunno, Bundang-gu, Sungnam Kyunggi-do, Republic of Korea 13591

Business Licence Number : 220-82-07506, Contact: tact@icact.org Tel: +82-70-4146-4991