Mobile Malware Classification for iOS Inspired by Phylogenetics

Muhammad ‘Afif Husainiamer, Madiah Mohd Saudi* 
Azuan Ahmad, Amirul Syauqi Mohamad Syafiq
Faculty of Science and Technology
Universiti Sains Islam Malaysia
Nilai, Malaysia

Abstract—Cyber-attacks such as ransomware, data breaches, and phishing triggered by malware, especially for iOS (iPhone operating system) platforms, are increasing. Yet not much work on malware detection for the iOS platform have been done compared to the Android platform. Hence, this paper presents an iOS malware classification inspired by phylogenetics. It consists of mobile behaviour, exploits, and surveillance features. The new iOS classification helps to identify, detect, and predict any new malware variants. The experiment was conducted by using hybrid analysis, with twelve (12) malwares datasets from the Contagio Mobile website. As a result, twenty-nine (29) new classifications have been developed. The experiment was conducted by using hybrid analysis, with twelve (12) malwares datasets from the Contagio Mobile website. As a result, twenty-nine (29) new classifications have been developed. The evaluation conducted, 13% of the mobile applications matched with the developed classifications. In the future, this work can be used as guidance for other researchers with the same interest.

Keywords—iOS; mobile malware; reverse engineering; exploitation; phylogenetic

I. INTRODUCTION

Currently, smartphones based on Android and iOS are commonly and widely used across the world. Yet, they also possess security concerns, especially security exploitation by malware such as ransomware and cryptojacking [1]. Unfortunately, the rapid increase of smartphone users contributed to mobile malware growth in the iOS environment. Malware is referred to as software that can infect devices, software, or networks with malicious attention without the owner’s consent. It can harm the victim with malicious activities such as stealing confidential information, identity theft, and spying on the victim. There are different kinds of malware such as viruses, Trojan, spyware, worms, and ransomware. It will cause a lot of chaos when the malware has successfully penetrated the smartphone system.

Whenever new vulnerabilities are released, Apple will update or patch to fix the weaknesses. By keeping the patch up to date, Apple makes sure the devices are secure enough to use. The malware attacks are carried out by attacking the kernel, giving the attacker private APIs (Application Programming Interfaces) and permission, and eventually gaining confidential information about the user. Unfortunately, there is a growing number of malwares attacking iOS devices. For example, it uses private APIs to implement malicious intent and view and steal its data. Fig. 1 shows statistics on the detection of malware for iOS by Welivesecurity [2].

Compared with Android, iOS is considered more secure. For example, in the iOS platform, the hardware, software, and even their booting process are monitored and secured by Apple procedures [3]. This scenario has an impact where many attackers tend to focus on Android malware rather than on iOS. In addition, based on the McAfee Labs Threat report on June 2018 shows a drastic increase in malware growth, and there were almost 2.9 million samples recorded [4]. Furthermore, high-risk vulnerabilities were detected in 38 percent of iOS mobile apps in 2019 compared to 43 percent of Android mobile apps [5]. Indeed, 40 percent of iOS malware attacks in 2017 targeted banking services [6]. As in Q1 2020, new mobile malware cases have surged by 71 percent, and new iOS malware grew by over 50 percent [7]. Hence, this paper presents a new mobile malware classification for iOS inspired by phylogenetics to overcome the above challenges. Phylogenetics is a term borrowed from biology and has been mapped into the cybersecurity field. It can be used to detect and predict malicious activity. This approach consists of malware behaviour, vulnerability exploitation, and surveillance features [8].

The proposed malware classification developed in this paper can detect any malware attacks against possible social media and online banking exploitation. This new iOS classification aids in the detection, identification, and prediction of new malware variants.
This paper is organized as follows. Section II discusses the related works, while Section III presents the methods used, and Section IV explains the findings. Finally, Section V discusses the conclusions reached by this paper.

II. RELATED WORK

A. iOS Malware Attacks

iOS malware attacks have been increased rapidly from years to years, and many researchers try to invade the issues and solve them to reduce the impact. Attackers evolve with the latest technology to ensure their intention to exploit user data can be conducted smoothly without interruption. Work by [9] found Trident worm and exploited three types of vulnerabilities once the link is clicked. Once it has been executed, the attackers will have the privilege to read, write, and any software in the infected device. Next, work by [10] found iKee Worm, which gathered logs on many jailbroken devices by scanning the OpenSSH port and used the root account and default password, and once infected, it will scan the surrounding IP address to spread the worm. They also found the YiSpecter worm, where the malware used ISP (Internet Service Provider) traffic, Window SNS (Self/Non-self) worm and offline applications installation, and other routes for transmission. It installed malware applications intending to collect private user information. Then, a previous study by [11] used Xcode Ghost worm where malware was sitting in the background of legitimate apps, then it did the data mining and injected malware in the apps when compiled. It possesses a new capability to prompt a fake alert dialogue to phish user credentials, hijack opening specifics URLs, and read and write data in the user clipboard. Work by [12] found AceDeciever that infects any Apple device connected to infected PC (personal computer) were capable of obtaining Apple ID (identification) and password. Finally, work by [13] found Keyraider worm, where it intercepted iTunes traffic and stole user login credentials, GUID (Globally Unique Identifier) devices Apple requests push service certificates and private keys, and iTunes receipts for purchase. It then sends this data to a remote server. Based on these previous studies, it can be concluded that there is a growing number of malwares attacking iOS users, and a solution to overcome these challenges is urgently needed. So, this paper proposes a new mobile malware classification as one of the mitigation solutions for the above challenges.

B. Phylogenetics

The phylogenetics aims to discover the origin of malware genes evolving [8]. It deals with evolutionary history and uses a tree diagram for different organisms and taxonomic groups. Malware phylogenetics emphasizes the similarities and relationships between a set of malwares. For example, a few types of phylogenetic tree models are the minimum spanning tree (MST), the persistent phylogeny tree, and the dendrogram [14]. Works by [15] used process mining which detects temporal logic properties designed to detect Android malware families and track the phylogenetic tree. [16] also used process mining where the program calls trace from a mobile application to classify associations and repeat execution patterns. Work by [17] used fuzzy clustering algorithm, where a malware program's syscalls can be modelled to produce a malware fingerprint with a number of associations and recurrent execution patterns. The author in [18] used discrete time Markov chain (DTMC) due to the paired KLD (Kullback-Leibler Divergence) and JSD (Jensen-Shannon Divergence) track calculation, it is computationally intensive. Bayesian network algorithm used by [19], learn a Directed Acyclic Graph (DAG) from observational data using statistical inference of conditional dependence and an informative antecedent to partial variable ordering. Work by [20] used extension of graphical lasso to discover a precision sparse matrix based on the kernel's combined matrix. An example of the phylogenetics diagram is depicted in Fig. 2. In this paper, there are three features mapped into phylogenetics to develop the classification. The identified features are malware behaviour, iOS version, and surveillance features.

C. Features Mapped to Phylogenetic

Malware behaviour can be classified into five parts: infection, activation, payload, operating algorithm, and mitigation [8]. In dynamic analysis, these five components are significant to classify malware based on their behaviour. The malware behaviour used EDOWA (Efficient Detection of Worm Attack) worm classification as the underlying concept [21]. Apple keeps satisfying their customer by serving them with the best version of iOS. The version must be updated to make sure the user is secured enough from any current security issues. The update also has some new features that can help user’s life easier [22]. Surveillance features used in this research come from 5 basic functions in the smartphone, consisting of call, SMS (Short Message Service), photos, audio, and GPS (Global Positioning System). All these features are dangerous whenever been exploited. The attacker can profit by exploiting either one of its features [23].

As one of the mitigating options for the stated challenges above, this research presents a new mobile malware classification based on phylogenetics. Three features are mapped into phylogenetics, which are malware behaviour, iOS version, and surveillance features.

III. METHODOLOGY

The overall process involved in this experiment is summarized in Fig. 3.

The analysis took place once the malware had been executed. The findings were mapped during the research regarding malware behaviour, vulnerability exploitation, and mobile phone surveillance features to allow malware classification. Malware behavior is referred to as infection,
payload, operating algorithm, activation, and propagation. Vulnerability exploitation refers to the iOS platform version, either iOS 10.x, 11.x, or 12.x, and the type of exploitation used. At the same time, mobile phone surveillance features are the features that attackers could use to exploit a mobile phone in the form of SMS, call log, camera, audio, and GPS. The mathematical formula for the proposed mobile malware classification is as follows:

$$\text{Set up laboratory}$$

$$\text{Extract malware file (.ipa,.deb,.dylib)}$$

$$\text{Create pattern from function and framework from malware files}$$

$$\text{Map result with phylogenetic properties}$$

$$\text{Compare pattern with applications in Apple Store}$$

**IV. RESULTS**

After the exploitation has been discovered, all the exploitation script's functions will be traced to their main frameworks to see what framework they are attacking during the malicious act. Then, the exploit is mapped into phylogenetics. The mapping result showed either the malware might lead to possible social media or online banking exploitation. If SMS or call is being exploited, it can be concluded as online banking exploitation. If any of those five features are being exploited, it is social media exploitation.

Table I shows malware analysis results based on the mapping with the phylogenetics.

| Classification | Description | Social Media Exploitation | Online Banking Exploitation |
|----------------|-------------|---------------------------|-----------------------------|
| E1: Unflod Malware | **Payload:** Phishing. Stole the device’s apple id, password and sent them in plaintext to the server. **Infection:** Host. From Chinese Cydia repositories **Operating Algorithm:** Stealth. **Activation:** Self-activation. Initiated their execution by exploring vulnerabilities in services that are available. **Propagation:** Passive monitoring. The malware infected those jailbroken devices that have download piracy Chinese repositories. | Possible exploitation | Possible exploitation |
| $\alpha_1 - \alpha_5$ (Malware Behaviour) | | | |
| $\beta_1 - \beta_5$ (iOS Chain) | iOS version: Chain 1 (iOS 10.x and below) | | |
| $\delta_1 - \delta_5$ (Surveillance Features) | GPS, SMS, call, audio, and photo | | |

$$(M, T) = S$$  \quad (1)

where

$$M(\alpha, \beta, \delta) = \alpha + \beta + \delta$$  \quad (2)

$$f(M_i, T_j) = S_{ij}$$  \quad (3)

Where $M$ represents the malware classification, $T$ represents the target asset, and $S$ is the detection model.

$M(\alpha, \beta, \delta) = \alpha + \beta + \delta$

$\alpha = \alpha_1 \cap \alpha_2 \cap \alpha_3 \cap \alpha_4 \cap \alpha_5$

$\beta = \beta_1 \cup \beta_2 \cup \beta_3 \cup \beta_4 \cup \beta_5$

$\delta = \delta_1 \cup \delta_2 \cup \delta_3 \cup \delta_4 \cup \delta_5$

$M_i \alpha \beta \gamma$$

where:

$\alpha_1 - \alpha_5$: payload, infection, operating algorithm, activation, and propagation

$\beta_1 - \beta_5$: iOS 10.x, iOS 11.x, iOS 12.x, iOS 13.x, iOS 1x.x

$\delta_1 - \delta_5$: SMS, call log, GPS, audio, and camera.
| E2: Spad Malware | Payload: Destructive. Change the actual owner id for the ads into their id to obtain the revenue. | Infection: Host. Result of some action taken by a user. The criminal hooks the legitimate functions and adds their tweaks. | Operating Algorithm: Stealth. | Activation: Human trigger. The malware will be functioning whenever the ads operate. | Propagation: Passive monitoring. The malware only can infect jailbroken devices. | No possible exploitation | No possible exploitation |
|----------------|-------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------|
| α₁ - α₅ (Malware Behaviour) | | | | | | | |
| β₁ - β₅ (iOS Chain) | iOS version: Chain 1(iOS 10.x and below) | | | | | | |
| δ₁ - δ₅ (Surveillance Features) | No surveillance features are involved as the attacker only exploits affect ads section | | | | | | |
| E3+E4+E5+E6: Inception Malware | Payload: Phishing. Collect device information and recording audio and send it to the c2(Command and Control) server. | Infection: Host. Result of some action taken by a user. The malware will inject into the host whenever the user clicks a malicious link or download any doc files. | Operating Algorithm: Stealth. | Activation: Human activation. The malware will be executed when the file has been opened. | Propagation: Passive monitoring. The malware only can infect those jailbroken devices that click the link or open the trojanized file. | Possible Exploitation | Possible Exploitation |
| α₁ - α₅ (Malware Behaviour) | | | | | | | |
| β₁ - β₅ (iOS Chain) | iOS version: Chain 1(iOS 10.x and below) | | | | | | |
| δ₁ - δ₅ (Surveillance Features) | GPS, SMS, Call Log, Audio, and Photos | | | | | | |
| E7+E8: Xcodeghost Malware | Payload: Destructive. Modifies Xcode, infects and steals some device information, and then sends it to the c2 server. | Infection: Host. Infected the apps that Xcodeghost produced. | Operating Algorithm: Stealth. | Activation: Human trigger. The malware will be functioning whenever the apps have been open. | Propagation: Passive monitoring. The malware only can infect the user of the apps that have been created using Xcodeghost | No possible Exploitation | No possible Exploitation |
| α₁ - α₅ (Malware Behaviour) | | | | | | | |
| β₁ - β₅ (iOS Chain) | iOS version: Chain 1(iOS 10.x and below) | | | | | | |
| δ₁ - δ₅ (Surveillance Features) | No surveillance features are involved. | | | | | | |
| E9+E10+E11+E12: Wirelurker Malware | Payload: Phishing. Exfiltration of user data and exfiltration of application usage and device serial number information. | Infection: Host. Infected through USB (Universal Serial Bus) when the device connects with the infected Mac. | Operating algorithm: Stealth. | Activation: Self-activation. It will start the malicious act once the malware has been injected into the device. | Propagation: Passive monitoring. The malware only can infect users that connect their device with the infected Mac. | Possible exploitation | Possible exploitation |
| α₁ - α₅ (Malware Behaviour) | | | | | | | |
| β₁ - β₅ (iOS Chain) | iOS version: Chain 1(iOS 10.x and below) | | | | | | |
| δ₁ - δ₅ (Surveillance Features) | GPS, SMS, Photos, Audio, and Call | All the surveillance features can be exploited by the attacker if they can brute force the victim’s Apple ID as they already get the Apple ID. | | | | | |
| E13: Zerghelper malware | Payload: Installing backdoor. Installed via a backdoor, requests the user to give Apple ID, shared Apple ID to other users, abuses the Apple ID by running different operations in the background, and abuses enterprise and personal certificates. | | | | | | |
| α₁ - α₅ (Malware Behaviour) | | | | | | | |

www.ijacsa.thesai.org
### E14: OneClickFraud malware

| Infection: | Host. Distributed through an adult site. |
| Operating algorithm: | Stealth. |
| Activation: | Human trigger. The malware will be functioning whenever the play button been clicked. |
| Propagation: | Passive monitoring. The malware only can infect visitors of the adult site. |

| iOS Chain | iOS version: Chain 1(iOS 10.x and below) |
| Surveillance Features | No possible exploitation |

### E15: XSSer malware

| Payload: | Phishing. Act as spyware and harvesting information from user device thus send it to C2 server. |
| Infection: | Host. Installed via a rogue repository on Cydia, the most popular third-party application store for jailbroken iPhones. |
| Operating algorithm: | Steal. |
| Activation: | Human trigger. The malware will be functioning whenever the apps are open. |
| Propagation: | Passive monitoring. The malware only can infect users that use the repositories. |

| iOS Chain | iOS version: Chain 1(iOS 10.x and below) |
| Surveillance Features | No possible exploitation |

### E16+E17: Muda malware

| Payload: | Display advertisements over other apps or in the notification bar and ask users to download iOS apps it promoted. |
| Infection: | Host. Spreads via third-party Cydia sources in China and only affects jailbroken iOS devices. |
| Operating algorithm: | Stealth. |
| Activation: | Human trigger. The malware will be functioning whenever the apps are open. |
| Propagation: | Passive monitoring. The malware only can infect jailbroken devices that using third-party Cydia repositories. |

| iOS Chain | iOS version: Chain 1(iOS 10.x and below) |
| Surveillance Features | No possible exploitation |

### E18+E19+E20: Tinyv malware

| Payload: | Destructive. It connects with its C2 server to get remote commands and install specified IPA file or DEB file(s) in the background, uninstalling specified IPA app or DEB package(s) in the background and changing the /etc/hosts file. |
| Infection: | Hosts. Repackaged into some pirated iOS apps for jailbroken devices. |

| iOS Chain | iOS version: Chain 1(iOS 10.x and below) |
| Surveillance Features | No possible exploitation |
Table I shows 13 of the 29 malware classifications mapped to phylogenetics, which can be used against possible exploitation for social media and online banking. The identified classification are E1 (Unflod malware), E3+E4+E5+E6 (Inception malware), E9+E10+E11+E12 (Wirelurker malware), E13 (Zerghelper malware), E15 (Xsser malware and E28+E29 (Keyraider malware).

In summary, the malwares have been reported based on the classification proposed. The analysis fits with the elements required for the classification based on this classification. Furthermore, every malware examined contains components that can be used for further exploitation.

Next, for the evaluation process, 50 anonymous apps from Apple Store and another 50 from the third-party store were selected. This is to test the practicality of the proposed classification in detecting any possible exploitation in the tested apps. As a result, 13% of the tested apps were identified with possible security exploitation. Two apps were from Apple Store, and 11 apps were from iOS Ninja. This 13% represents the possible exploitation either against social media or online banking.

| Malware Behaviour | Operating algorithm | Activation | Propagation | Payload | Infection | Operating algorithm | Activation | Propagation | Payload |
|-------------------|--------------------|------------|-------------|---------|-----------|---------------------|------------|-------------|---------|
| α1ₐ, α5ₐ (Malware Behaviour) | Stealth. | Human trigger. The malware will be executed whenever the apps are being used. | Passive monitoring. The malware only can infect users that use the repositories. | Destructive. Abusing enterprise certificates, installing malicious apps, uninstall apps, and self-monitoring and updating, collecting, and uploading device information, changing safari configurations, hijacking other apps execution, and pretending to be system apps. | Host. Through hijacking of traffic from nationwide ISPs, an SNS worm on Windows, and an offline app installation and community promotion. | Stealth. | Self-activation. It will start the malicious act once the malware has infected the device. | Passive monitoring. The malware only can infect jailbroken devices that using third-party Cydia repositories. | Destructive. It is stealing Apple account data, certificates, and private keys. | Host. Distributed through third-party Cydia repositories in China. |
| β1 - β₅ (iOS Chain) | iOS version: Chain 1(iOS 10.x and below) | No possible exploitation | No possible exploitation | No possible exploitation |
| δ1 - δ₅ (Surveillance Features) | No surveillance features involved. | |
| E21+E22+E23+E24+E25+E26+E27: Yispecter Malware | |
| β1 - β₅ (iOS Chain) | iOS version: Chain 1(iOS 10.x and below) | Possible exploitation | Possible exploitation | Possible exploitation |
| δ1 - δ₅ (Surveillance Features) | No possible exploitation | GPS, SMS, Call Log, Audio, and Photos |

V. CONCLUSION

Based on these experimental results, the proposed malware classification developed in this paper can be used to detect any malware attacks against possible exploitation for social media and online banking. These new iOS classification helps to identify, detect, and predict any new malware variants. Another important consideration for future improvement would be to revise the frameworks and functions involved in the iOS architecture from time to time and integrate with artificial intelligence-based alarms. Malicious apps will use the combinations of frameworks and functions in the iOS architecture to exploit the targeted feature successfully. Furthermore, as a newer iOS version will be introduced, a new framework and functions may also be offered. Hence, there is a need to add more malware classifications based on the mobile malware classification formulation developed in this paper.

ACKNOWLEDGMENT

The authors would like to express their gratitude to the Universiti Sains Islam Malaysia (USIM) for the support and facilities provided. This research paper project is under the grant: [P1-I7-16120-UNI-CVD-FST].

www.ijacsa.thesai.org
