iCaMs: 안티 콜 피싱 및 메시지 사기를 위한 지능형 시스템

Manh-Hung Tran, 양희규, Thien-Binh Dang, 추현승
성균관대학교 정보통신대학
e-mail :{hungtm, huigyu, binhdt, choo}@skku.edu

iCaMs: An Intelligent System for Anti Call Phishing and Message Scams

Manh-Hung Tran, Thien-Binh Dang, Hyun-Seung Choo
Dept. of Electrical and Computer Engineering, Sungkyunkwan University

Abstract

The damage from voice phishing reaches one trillion won in the past 5 years following report of Business Korea on August 28, 2018. Voice phishing and mobile phone scams are recognized as a top concern not only in Korea but also in over the world in recent years. In this paper, we propose an efficient system to identify the caller and alert or prevent of dangerous to users. Our system includes a mobile application and web server using client and server architecture. The main purpose of this system is to automatically display the information of unidentified callers when a user receives a call or message. A mobile application installs on a mobile phone to automatically get the caller phone number and send it to the server through web services to verify. The web server applies a machine learning to a global phone book with Blacklist and Whitelist to verify the phone number getting from the mobile application and returns the result.

1. Introduction

Nowadays, using a telephone is indispensable for every person to communicate with others. Everyday people often receive quite plenty of phone calls and text messages. Along with that, the problem of receiving impersonation calls and unwanted calls or spam message is increasing such as voice phishing and scam. Besides, phishing attacks have spread to cover wide areas of services as shown in Fig 1. The most target of phishing attack is payment and related to financial. For example, in Korea, a total of 48,743 voice phishing victims were reported [1][2] in 2018 and total financial damage from voice phishing scams up to 444 billion won. According to the numbers and figures are shown here, we conclude that phishing attacks on mobile devices are increasing dramatically with the huge financial damage to the economy. Because of these warning numbers, it is necessary to have a solution to prevent people from phishing attack on mobile.

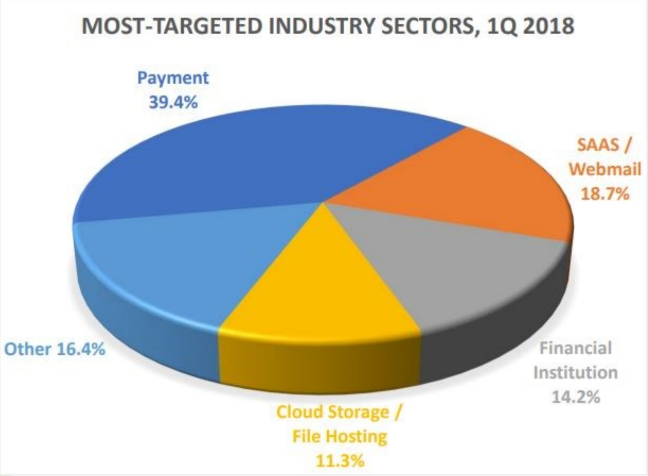

Fig. 1 Distribution of Phishing Attacks [3].

Phishing is pronounced fishing. It is just a scam where a criminal uses text message, phone calls, and other contact methods to pretend to be someone they are not, in order to get access to important information of a victim. Two common ways of phishing attacks on mobile are voice phishing and message or email scam is shown in

Fig. 2. With voice phishing, the victim receives directly fraudulent calls and request to provide financial-related information or transfer money to a phisher. With the second way, a victim receives a message/email with a malicious link. By clicking on this link, users may lose information or pay for an unknown item. Besides, this link can also be a fake website to ask users to submit their important information.

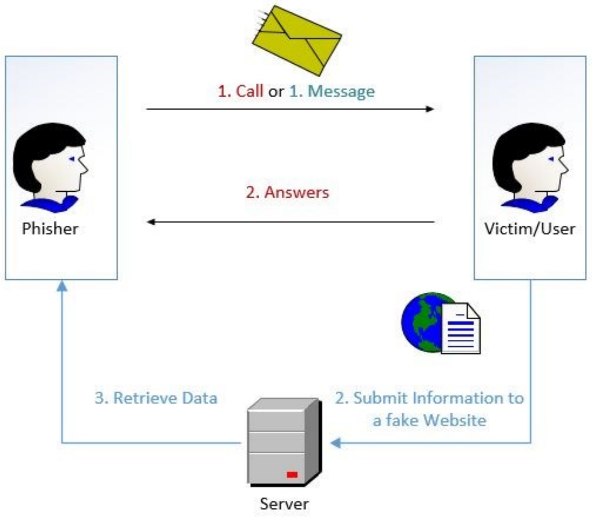

Fig 2. Mobile Phone Phishing Attack

In this paper, we propose an intelligent system – **iCaMs** – to identify the caller and aims to prevent the mobile phishing attack by using a proactive approach. **iCaMs** uses a global phone book with Whitelist and Blacklist. It is an automatic system using AI to learn and analyze the phone number of a caller based on their data in a global phone book and internet. A global phone book is collected by using information in contact of each user. Moreover, this system is a full solution with various function against the mobile phishing attack.

2. Related works

There are some existing researches with a reactive approach on solving phishing attach. In [4], Belal Amro provided an analysis of different types of phishing attacks and analyzed a summary of anti-phishing techniques on mobile devices. Ankit Kumar et al. [5] proposed an approach to checks the legitimacy of a webpage using hyperlink features. When users click a link of a website which is not existing in the white-list, the browser warns users not to submit their sensitive information. Madhuresh Mishra et.al.[6] proposed a preventive anti-phishing technique to avoid to be victims of phishing attacks. This technique helps people identify the true website compared with a fake link received from a phisher. However, this research only deals with preventing users to provide information to the phisher through a fake web page like email/message phishing. It cannot solve the voice phishing problem.

Many mobile applications with proactive approach aim for blocking calls from unwanted numbers and spam numbers such as Should I Answer, Blacklist Plus, Call Blocker Free. These applications protect a user from unknown calls related to telemarketing, advertisement, etc... It allows users to access Blacklist, Whitelist and block spam calls with call reminder and notifications. It also provides various blocking options like blocking calls from hidden numbers, foreign countries, premium rate numbers and from numbers that are not in the contact list of users. Additionally, it displays notifications like phone number rating, information, and user reviews once the phone starts ringing. However, these applications only have a local phone book. Moreover, users have to manually add phone numbers to Whitelist and Blacklist. It cannot help with a strange number from phisher that not have in their phonebook.

T-Mobile provides a Mobile security application [7] to help users against unwanted calls, scams by using the information of users on their own and their partner database. This application only works with the customer of T-Mobile in the USA. In Korea [1], an AI mobile application that auto-detect "voice phishing" scam was introduced in March 2019. It uses machine learning base on voice recognition to detect some words that scammer usually uses when victim/user receive a call. If a scammer is detected, it will notify an alert to the user on their mobile screen. However, it only helps user against voice phishing. And, in my opinion, this application only works with a small number of users because the number of phone calls every day is tremendous.

3. Proposed Solution

we proposed an intelligent system using a mobile application and machine learning-based to prevent the mobile phishing attack. The working flow of the system is shown in Fig. 2. Firstly, a mobile application is installed on the mobile phone of

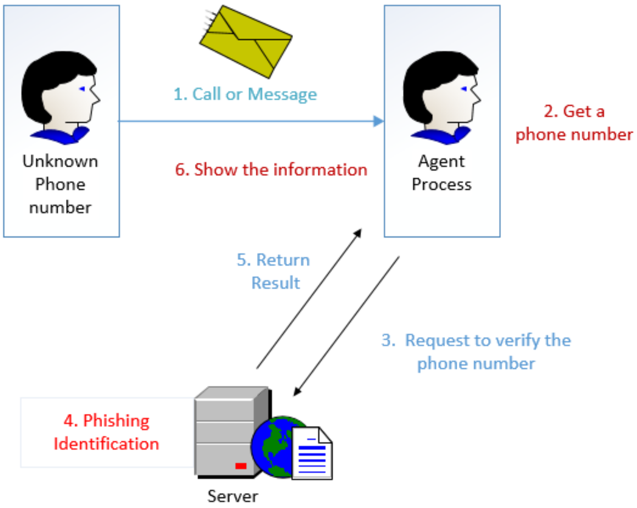

Fig.3 System Working Flow of Proposed.

users who are registered as a normal user in iCAMs system after a verification process. An agent work as a background process in mobile to get a phone number when a user receives a message or phone call. Firstly, the agent will check the phone number with the local contact list of the users. If the phone number is in the contact list then it will do nothing, otherwise it sends a request to a server with parameter is the unknown phone number. The server receives the request and process to identify the phone number is phishing or not. The result of the phishing identification will return to the agent on the mobile phone. Depend on the result and users private policy, the agent will decide to show the information on the screen or block a call. The information shows in the mobile device is the classification of the unknown phone number and the percentage of legitimate also display to the mobile UI. The client UI prototype is shown in Fig.4.

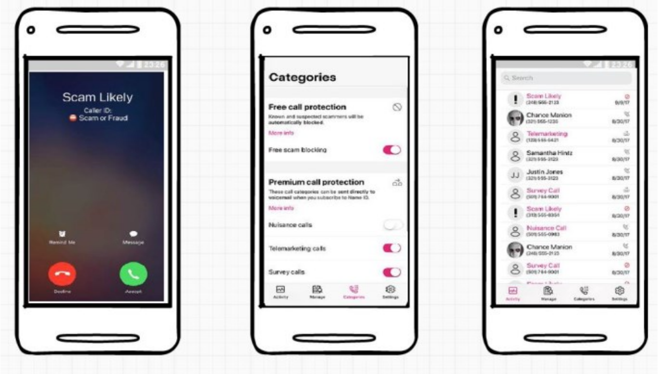

Fig.4 Client UI Prototype

The architecture of phishing identification of the proposed idea is depicted in Fig. 5. Our Architecture divided into three main modules. The first module is the Global Phone Book including three datasets e.g. Whitelist, Blacklist and Users Contact. The Whitelist indicate the legitimate phone number and the Blacklist identify the phishing number. The phishing number is classified into four categories for instance Telemarketing, Survey Call, Nuisance Call, and Scam Likely. While the User Contact dataset is the local contact of all registered users. All of these datasets are encoded to guarantee information security. It could be manually updated with the verified information and user-report or automatically updated by the machine learning module. The second module is Working Flow module. This module matches the requested phone number with Whitelist and Blacklist datasets to indicate the request is phishing or legitimate. If the requested phone number is not existing in both datasets it will send a request to a Machine Learning module. The third module is the Machine Learning module. This module will responsibility for classification and analysis the phone number combine with other information e.g. the number of references in the Users Contact dataset, calling time, and some useful information from searching on the internet. The output of the classification process is the percentage of prediction for each category. The output data is compared with the threshold to determine which category it belongs to. After that, it updates the Whitelist or Blacklist dataset and resends the request to the Working Flow. This is a continuous process of training and classification. The accuracy rate of machine learning will increase when the number of registered user increase.

The client in our system is a mobile application developed for both android and IOS systems. It provides many functions for a user such as a user verify and registration, phishing report, private policy setting, etc... The server side providing a web service is implemented by PHP language programming. The storage of the Global Phone Book is MySQL database. JSON is a protocol of communication between client and server in iCaMs. In the server-side also provide a web page for management the phonebook dataset. It provides some basic function like authors and authentication, edits and updates the global phonebook. To implement the classification and analysis process by machine learning, we use TensorFlow developed by Python 3.6 under platform Anaconda3.

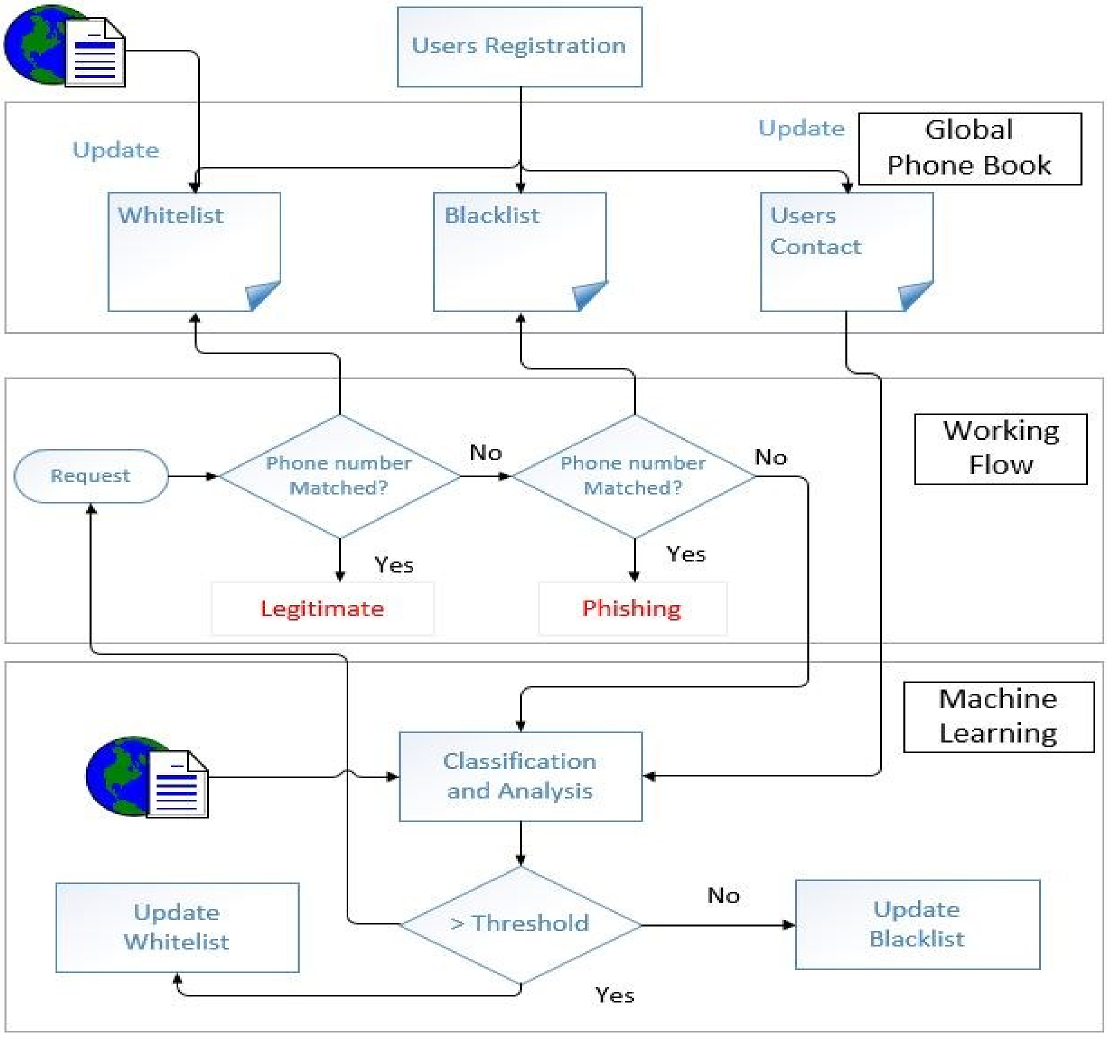

Fig.5 The Architecture of Phishing Identification

4. Conclusion

In this paper, we proposed an intelligent system that combines machine learning with a mobile application to prevent the mobile phishing attack. By using iCaMs, people are protected from voice scam and also help others to avoid it by reporting the scam likely phone number. The effectiveness and correctness of iCaMs depend on the number of users and machine learning algorithm applying to the classification process.

ACKNOWLEDGMENT

본 논문은 과학기술정보통신부 및 정보통신기획평가원의 Grand ICT 연구센터지원사업 (IITP-2019-2015-0-00742), 과학기술정보통신부 및 정보통신기획평가원의글로벌핵심인재양성지원사 업 (2019-0-01579) 과 2019 년도 정부 (과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원(No.2019-0-00421, 인공지능대학원지원)의 연구결과로 수행되었음.

References

[1]http://www.koreaherald.com/view.php?ud=20190317000115
[2]http://koreajoongangdaily.joins.com/news/article/article.aspx?aid=3059014
[3]https://docs.apwg.org/reports/apwg_trends_report_q1_2018.pdf
[4] Belal Amro, Phishing Techniques in Mobile Devices Journal of Computer and Communications, 2018.
[5] Ankit Kumar Jain and B. B. Gupta, A novel approach to protect against phishing attacks at client side using auto-updated white-list, Journal on Information Security, 2016.
[6] Madhuresh Mishra, Gaurav, Anurag Jain, A Preventive Anti-Phishing Technique using
Code word, International Journal of Computer Science and Information Technologies, 2012.
[7] https://www.t-mobile.com/