Hindawi
Security and Communication Networks
Volume 2018, Article ID 7251750, 11 pages
https://doi.org/10.1155/2018/7251750

Research Article

Leverage Website Favicon to Detect Phishing Websites

Kang Leng Chiew, Jeffrey Soon-Fatt Choo, San Nah Sze, and Kelvin S. C. Yong

Faculty of Computer Science and Information Technology, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak, Malaysia

Correspondence should be addressed to Kang Leng Chiew; klchiew@unimas.my

Received 23 October 2017; Revised 10 January 2018; Accepted 30 January 2018; Published 6 March 2018

Academic Editor: Luca Caviglione

Phishing attack is a cybercrime that can lead to severe financial losses for Internet users and entrepreneurs. Typically, phishers are fond of using fuzzy techniques during the creation of a website. They confuse the victim by imitating the appearance and content of a legitimate website. In addition, many websites are vulnerable to phishing attacks, including financial institutions, social networks, e-commerce, and airline websites. This paper is an extension of our previous work that leverages the favicon with Google image search to reveal the identity of a website. Our identity retrieval technique involves an effective mathematical model that can be used to assist in retrieving the right identity from the many entries of the search results. In this paper, we introduced an enhanced version of the favicon-based phishing attack detection with the introduction of the Domain Name Amplification feature and incorporation of addition features. Additional features are very useful when the website being examined does not have a favicon. We have collected a total of 5,000 phishing websites from PhishTank and 5,000 legitimate websites from Alexa to verify the effectiveness of the proposed method. From the experimental results, we achieved a 96.93% true positive rate with only a 4.13% false positive rate.

1. Introduction

Phishing attacks can be defined as an act of deceiving victims via e-mail or a website to gain their trust to disclose their personal and financial information. With the advancement of information technology, many business agencies (e.g., banks, tourism, hotels, and airlines) can incorporate e-commerce, electronic payments, and social networking technologies into their businesses to increase sales. But this creates opportunities for phishers to gain illegal profits by disguising a wide range of services offered by financial institutions, social networking, and e-commerce websites. The Antiphishing Working Group (APWG) reported a total of 128,378 unique phishing websites detected in the second quarter of 2014 phishing activity trends report [1]. The report showed evidence that phishing activities are on the rise, which revealed that the existing antiphishing solutions were unable to resist phishing attacks efficiently.

The most common way to create a phishing website is through content replication of popular websites such as PayPal, eBay, Facebook, and Twitter. Phishing websites can be produced quickly and require little effort. This is because the phisher can simply clone the website with some modifications in the input tag to collect personal information. Furthermore, this process can be shortened by using a phishing kit [2] available on the black market. Inadvertently, advances in information technology also help phishers to develop high-profile phishing techniques to avoid phishing detectors. Figure 1 shows an example of a phishing website masquerading as PayPal. There are two flaws identified in the address bar (as shown by the red line box in Figure 1):

(i) The domain name is completely different from the genuine PayPal website.

(ii) It obfuscates the URL with HTTPS as part of the URL.

Although there are many solutions proposed to detect phishing websites, these solutions have some shortcomings. First, existing textual-based antiphishing solutions depend on the textual content of a webpage to classify the legitimacy of a website. Therefore, these solutions are incompetent to classify image-based phishing websites. A phisher can replace the textual contents with images to evade phishing detectors. Second, some phishers create phishing websites that are visually similar (e.g., webpage layout) to the legitimate website to phish potential victims. They preserve iconic images