International Journal of Engineering & Technology, 7 (4.31) (2018) 1-6

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Identifying the Most Effective Feature Category in Machine Learning-based Phishing Website Detection

Choon Lin Tan, Kang Leng Chiew[1]*, Nadianatra Musa[2], Dayang Hanani Abang Ibrahim[3]

Faculty of Computer Science and Information Technology, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak, Malaysia
**Corresponding author E-mail: klchiew@unimas.my*

Abstract

This paper proposes an improved approach to categorise phishing features into precise categories. Existing features are surveyed from the current phishing detection works and grouped according to the improved categorisation approach. The performances of various feature sets are evaluated using the C4.5 classifier, whereby the content URL obfuscation category is found to perform the best, achieving an accuracy of 95.97%. Additional benchmarking is conducted to compare the performance of the winning feature set against other feature sets utilised in existing phishing detection techniques. Results suggest that the winning feature set is indeed an effective feature category which has contributed significantly to the performance of existing machine learning-based phishing detection systems.

Keywords: *Classification; Feature Categorisation; Machine Learning; Phishing Detection; Web Security*

1. Introduction

Phishing is a cyber-threat that utilises counterfeit websites to steal sensitive user information such as account login credentials, credit card numbers, etc. Victims are usually led to the phishing websites by clicking on a URL in a fraudulent email that claims to originate from reputable institutions. At the phishing website, victims will then be presented with familiar visual cues (e.g., logo, colour, design, etc.) to convince them to submit their personal information.

Over the years, the severity of phishing attacks has not seen any significant decline despite mitigation efforts such as increasing public's awareness and deploying technical security solutions. As of June 2017, the Anti-Phishing Working Group (APWG) has reported that a number of unique phishing websites remained high at 50,720 [1]. In another report by RSA, it is estimated that global organisations suffered $9 billion loss due to phishing incidents in 2016 [2]. As a result, users are hesitant to fully utilise online banking and e-commerce services.

The blacklist-based detection system is among the most widely deployed anti-phishing solutions in conventional browsers such as Google Chrome and Mozilla Firefox. The blacklist-based detection system queries a central database of known phishing URLs and issues a warning when the user navigates to a known phishing website. However, recent studies have shown that blacklist-based solutions are unable to capture newly launched phishing websites [3], [4].

Another established form of the anti-phishing solution is the machine learning-based detection system. This technique is considered as state-of-the-art due to its' ability to recognise even new phishing websites. Machine learning-based detection systems rely on classifiers which function as decision systems to detect phishing based on features harvested from a variety of sources such as webpage URL, HTML contents, third party services, etc.

The selection of effective features is crucial in developing high-performance machine learning-based phishing detection systems.

Many existing researchers adopt features that appeared commonly in prior phishing detection studies. They tend to consider the existing common features as good features, even without established experimental results to support such belief. On the other hand, some researchers may propose additional or new features to enhance their phishing detection system [5]–[9].

In addition, the performance of features is rarely assessed by category-based benchmarking. As a result, anti-phishing researchers may labour in vain when focusing on certain feature categories that are less effective, thus failing to attain the desired phishing detection performance. Hence, establishing proper category-based benchmarking is important so that security experts can concentrate their efforts on a superior feature category that has more potential to improve the phishing detection rate. Moreover, capitalising on superior feature category facilitates rapid development of efficient yet effective anti-phishing applications.

Thus, in this paper, we surveyed the features employed in existing phishing detection works and proposed an improved categorisation approach to classify them. Through the experiments, we provide benchmarking results to compare the performance of different feature categories. Additionally, we conducted distribution analysis on the features' values to provide more insight as to why certain feature categories are better in differentiating between phishing and legitimate websites. In summary, the main contributions of this paper are highlighted as follows:

a) Proposing a new categorisation approach to group phishing features into more precise categories.
b) Conducting benchmarking to assess and compare the performance of each feature category.
c) Identifying a small set of superior features from the winning category that achieves a high phishing detection rate while effectively minimising computational processing power.

The remainder of this paper is organised as follows: Section 2 introduces related studies on feature evaluation from existing machine learning-based phishing detection techniques. Section 3 lists the improved phishing feature categories. Section 4 describes the experimental setup, the results and findings. Finally, Section 5