Expert Systems with Applications 40 (2013) 4697–4706

Contents lists available at SciVerse ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Intelligent phishing detection and protection scheme for online transactions

P.A. Barraclough [a], M.A. Hossain [a,*], M.A. Tahir [b], G. Sexton [a], N. Aslam [a]

[a] Computational Intelligence Group, University of Northumbria at Newcastle, Newcastle Upon Tyne NE1, United Kingdom
[b] College of Computing and Information Sciences, Al-Imam Mohammad Ibn Saud Islamic University, Riyadh, 11432, Saudi Arabia

ARTICLE INFO

Keywords:
Phishing
Neuro-Fuzzy scheme
Legitimate site rules
Online transaction

ABSTRACT

Phishing is an instance of social engineering techniques used to deceive users into giving their sensitive information using an illegitimate website that looks and feels exactly like the target organization website. Most phishing detection approaches utilizes Uniform Resource Locator (URL) blacklists or phishing website features combined with machine learning techniques to combat phishing. Despite the existing approaches that utilize URL blacklists, they cannot generalize well with new phishing attacks due to human weakness in verifying blacklists, while the existing feature-based methods suffer high false positive rates and insufficient phishing features. As a result, this leads to an inadequacy in the online transactions. To solve this problem robustly, the proposed study introduces new inputs (Legitimate site rules, User-behavior profile, PhishTank, User-specific sites, Pop-Ups from emails) which were not considered previously in a single protection platform. The idea is to utilize a Neuro-Fuzzy Scheme with 5 inputs to detect phishing sites with high accuracy in real-time. In this study, 2-Fold cross-validation is applied for training and testing the proposed model. A total of 288 features with 5 inputs were used and has so far achieved the best performance as compared to all previously reported results in the field.

1. Introduction

Phishing is a major problem nowadays causing losses of finance, particularly in online transactions (Financial Fraud Action UK & Credit clearing Company, 2012). Phishing definition varies from literature to literature. Jacobson and Meyers defined phishing as an act to fraudulently acquire user's sensitive information (personal identity number, passcode, password, credit/debit card number) through illegitimate website that looks exactly like the target website (Jakobsson & Myers, 2007). According to the UKCards Association's Press Release report, an increase in phishing attacks in online transaction has caused losses of £21.6 million between January and June 2012, which is a growth of 28% from June 2011 (Financial Fraud Action UK & Credit clearing Company, 2012). This significant increase is caused by a huge number of phishing websites created by criminals as a means of deceiving users into providing their credentials for financial benefit (Carter, 2012). Phishing techniques are improved regularly and are getting more sophisticated causing tremendous losses annually.

Despite various anti-phishing approaches developed to combat the problem, these approaches suffer high false positive rates. As a result, there is still a lack of accuracy and real-time solutions causing inadequacy in online transaction (Xiang, Hong, Rose, & Cranor, 2011). Some of these approaches employs feature-based using machine learning algorithms (Aburrous, Hossain, Dahal & Thabtah, 2010; Martin, Anutthamaa, Sathyavathy, Marie Francois, & Venkatesan, 2011; Xiang et al., 2011; Liu, Giu, & Wenyin, 2010; Sanglerdsinlapachai, & Rungsawang 2010; Xiang & Hong, 2009). Others are content-based approaches with lexical Uniform Resource Locator (URL) (Le, Markopoulou, & Faloutsos, 2011; Zhang et al., 2012). Some approaches uses heuristics (Zhang, Hong, & Crano, 2007; Afroz & Greenstadt, 2009; Shahria & Zulkerinen, 2010), while other approaches employs visual similarity (Chen, Dick, & Miller, 2010) and others utilizes URL blacklists (Xiang et al., 2011; Sheng et al., 2009; Spiezle, 2007).

The existing blacklists, which are largely used in industries, cannot generalize well to new phishing attacks (Cranor, Eglman, Hong, & Zhang, 2006). Also Sheng et al. revealed that the accuracy for protection offered by blacklist is not greater than 40% and are slow in response to new phishing attacks as updates take longer (Sheng, 2009). It is a case in which 83% of launched phishing websites takes 12 h to appear in a blacklist. Moreover, no comprehensive features that are wholly representative of phishing strategies have been proposed.

To address the problem robustly, it is important to build a state-of-the-art model using Neuro-Fuzzy scheme with five inputs. Neuro-Fuzzy is a Fuzzy Logic and a Neural Network. The point for using Neuro-Fuzzy is that, it has a universal approximations with ability to use Fuzzy IF...THEN rules. Neural Network performs well when

* Corresponding author.
E-mail addresses: phoebe.barraclough@northumbria.ac.uk (P.A. Barraclough), alamgir.hossain@northumbria.ac.uk (M.A. Hossain), mtahir@ccis.imamu.edu.sa (M.A. Tahir), g.sexton@northumbria.ac.uk (G. Sexton), nauman.aslam@northumbria.ac.uk (N. Aslam).

0957-4174/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.eswa.2013.02.009

dealing with raw data, while Fuzzy Logic deals with reasoning on a higher level, using linguistic information from domain experts (Negnevitsky, 2002). Five inputs are tables where features are stored which include: Legitimate site rules, User-behavior profile, PhishTank, User-specific sites and Pop-Ups from Emails. From these, 288 features are extracted to be used as training and testing data. The advantage of five inputs is that they are wholly representatives of phishing techniques and strategies. Further, training and testing experiments were performed using a 2-Fold cross-validation method based on Adaptive Neuro-Fuzzy Inference System (ANFIS) to measure the system accuracy and robustness. Cross-validation is a testing method and also signifies a group of methods, while in this case it is used to address over-fitting problems (Taher, 2010). Adaptive Neuro-Fuzzy Inference System is a hybrid intelligent system which has the ability for reasoning and learning. The experimental results shows that Neuro-Fuzzy with five inputs has the best performance compared to all previously reported approaches.

The main contributions in this study are the five inputs as they are important elements. This study is significant because the system will restore user's confidence in online transactions.

In Section 1.1, the objectives are presented followed by the review of literature and related work. Section 3 describes the proposed Neuro-fuzzy approach with five inputs. Learning rules and Adaptive Neuro-Fuzzy Inference System are also described in this section. Section 4 covers feature extraction and analysis. The experimental procedure including training and testing is covered in Section 5 together with results and discussion. Contribution to knowledge is also described in Section 5. Section 6 concludes this paper and outline future work. The aim is to design and develop an intelligent phishing detection and protection model for online transactions based on Neuro-Fuzzy and five inputs.

1.1. Specific objectives

- To identify and extract phishing features based on five inputs.
- To develop a Neuro-Fuzzy model using advanced techniques.
- To train and validate the Fuzzy Inference model in real-time environment.
- To provide a comparative study to demonstrate the merit of the proposed approach.

The advantage is to make users more secure and build their confidence in online transactions.

2. Related work

Phishing attacks are becoming more and more sophisticated in techniques daily, especially for online transaction. Approaches have been developed to tackle phishing attacks. These are classified in three groups, which include feature-based, content-based, heuristics-based and blacklists-based approaches. Existing feature-based approaches are as follows:

An intelligent phishing detection was developed by Aburrous et al. (2010). Their approach was based on Fuzzy data-mining algorithms with 27 features and six criteria. The approach achieved 83.7% accuracy. However, this approach has inadequate features. Incorporating additional comprehensive features could overcome the limitation. A similar framework for Predicting Phishing Websites was proposed by Martin et al. (2011). In their approach, Neural Network was used for training and testing in order to predict their system performance. They explored Anti-Phishing Working Group and PhishTank to extract phishing website features. They discovered that phishing websites lived only for 2.25 days before taken down. Though interesting, formal results has not been presented, making it hard to review their performance. Equally, CANTINA+ was proposed by Xiang et al. (2011). Their method used fifteen features based on Hypertext Mark-up Language Document Object Model, Search Engines, a machine learning algorithms and PhishTank. They obtained 92% accuracy, however the approach suffered high false positives. Adding more effective phishing features based on a machine learning technique may be the solution. Similarly, Sanglerdsinlapachai, and Rungsawang (2010) proposed features using machine learning web-based phishing detection. The approach was based on the domain top-page similarity to test whether a page is phishing or legitimate. However, the approach incurred 19.50% error rates. Adding more relevant features would improve the accuracy.

Equally, Liu et al. (2010), created an Automatic Detection for phishing targets from phishing web pages. Their approach was based on density-based spatial clustering of applications with noise to cluster suspicious webpages from legitimate webpages. Their experiment result obtained 91.44% accuracy with false rate of 3.4%. The performance was good, but could expose users to a high risk. A combination of Fuzzy Logic and support vector machine could be used to reduce the errors.

Xiang and Hong, (2009) proposed a linear classifier. Their approach utilized Hyper-Text Mark-up language, Domain Object Model with 10 features to identify phishing sites. They achieved 89% accuracy. However, the coverage of features is limited to deal with phishing techniques. More effective features may be used to solve the problem.

Another work utilizes content-based approach. For instance, PhishDef was proposed by Le et al. (2011). The approach was based on a selection of lexical URL features resistant to obfuscation techniques. An automatic and hand-selected evaluation of classification accuracy using lexical features was also performed. The experiments achieved 95% accuracy. However focusing only on uniform resource locator features is a high risk. Using algorithms and site contents could be a solution. Moreover, Zhang et al. (2012) proposed a novel description model to detect phishing. Their approach was based on phishing domain ontology, using 3 principle of phishing descriptive model along with statistical algorithm. They obtained 97% accuracy. However, false rates still exists.

Other studies employed heuristic-based approach. CANTINA was introduced by Zhang et al. (2007). The method was based on calculating TF-IDF score for each word in the page. In experiment, the pure TF-IDF approach detected 89% phishing sites, but suffered 11% false positives. One possible solution is to make legitimate page indexed. Similarly, Afroz and Greenstadt (2009) proposed a PhishZoo. Their approach was based on fuzzy hashing, using profiles of trusted websites appearances to identify phishing sites. This approach offers a reduced effect on phishing site appearance, allowing users to recognize phishing sites, but there is still a lack of generalization to new phishing due to human interventions. Using an automatic process could reduce human intervention. Also, Phish Tester was designed by Shahriar and Zulkemine (2010) to automate testing processes. The approach was based on an application behavior model, using five heuristic coverage criteria to identify inconsistence that leads to a conclusion whether a given website is legitimate or phishing. However, they attained 3% error rates. Using machine learning technique could solve the problem.

Another method, Chen et al. (2010) took a holistic view of the visual similarity between websites and applied compression algorithms on the sites as indivisible elements to detect phishing. One issue with their method was that it could not handle attacks well.

In addition, blacklist-based approaches also exist. Anti-phishing toolbars that utilizes blacklists include Xiang et al., Microsoft Internet explorer 8 to analyses page properties to distinguish phishing sites (Xiang et al., 2011; Spiezle, 2007). The investigation by Sheng