Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 125 (2018) 617–623

Procedia Computer Science

www.elsevier.com/locate/procedia

6th International Conference on Smart Computing and Communications, ICSCC 2017, 7-8 December 2017, Kurukshetra, India

Rule-Based Framework for Detection of Smishing Messages in Mobile Environment

Ankit Kumar Jain*, B. B. Gupta

Department of Computer Engineering, National Institute of Technology Kurukshetra-136119, Haryana (India) ankitjain@nitkkr.ac.in

Abstract

Smishing is a cyber-security attack, which utilizes Short Message Service (SMS) to steal personal credentials of mobile users. The trust level of users on their smart devices has attracted attackers for performing various mobile security attacks like Smishing. In this paper, we implement the rule-based data mining classification approach in the detection of smishing messages. The proposed approach identified nine rules which can efficiently filter smishing SMS from the genuine one. Further, our approach applies rule-based classification algorithms to train these outstanding rules. Since the SMS text messages are very short and generally written in Lingo language, we have used text normalization to convert them into standard form to obtain better rules. The performance of the proposed approach is evaluated, and it achieved more than 99% true negative rate. Furthermore, the proposed approach is very efficient for the detection of the zero hour attack too.

Peer-review under responsibility of the scientific committee of the 6th International Conference on Smart Computing and Communications.

Keywords: Smishing; Mobile Phishing; Data mining; Short messaging service; Machine learning

1. Introduction

These days mobile security is a major concern because attackers have diverted their mind from personal computers to smartphones because with the increase in technology. Moreover, people are more attracted towards smartphones as it is a small and multi-functioning device, Mobile devices are more popular these days as compared to laptops because of their small screen size, lower production cost, and portability.

* Corresponding author.
E-mail address: ankit.jain2407@gmail.com

1877-0509
Peer-review under responsibility of the scientific committee of the 6th International Conference on Smart Computing and Communications
10.1016/j.procs.2017.12.079

Smishing word is constructed by combining two words that are SMS and Phishing [1]. Mobile phishing is an emerging threat in which malicious person sends an SMS message to the user, and that SMS contains links to malicious applications and Webpages. The phishers will not only get money but also acquire information about contact numbers, mobile device versions, photos, etc. According to a study, 44% of users are not aware of the security solution available for mobile devices [2].

It was estimated that out of 10, seven people do not take any action against unwanted messages [3]. Attackers have now shifted their focus to mobile users due to several reasons. First is - extensive use of smartphones, second is - increase in dependency of users on smartphone applications for performing various tasks. Third, the user believes that with two-factor authentication method, only trusted messages will be delivered to their devices [5]. Multiple reports evidently indicate that Smishing attacks have dramatically increased over the last few years. In 2016, a customer of UK bank Santander lost 22,700 Pound in an SMS phishing scam [4]. According to Dimensional Enterprise Mobile security Survey report and it shows that smishing attack stands at the second position in all kind of mobile devices attacks [14].

There is two type of defense methods are used to detect fake mobile SMS. The first method is the blacklist based technique that stops the incoming SMS from the fake sources [17]. However, blacklist-based techniques do not cover all the fake sources, as a criminal can purchase any mobile number to send the bogus SMS. The second type of solution is based on the machine learning algorithm where various features are extracted and compute from the SMS to take appropriate decision. The advantage of the machine learning based technique is that it can detect the fake message coming from any source. Data mining methods help in the feature extraction and finding the relation between them [16]. These approaches identifying hidden knowledge from datasets in terms of rules and make the decision based on extracted rules. Human easily understands these rules and their rules are written in the form of IF-condition THEN action.

In this paper, we employed the rule-based data mining classification approach in the prediction of smishing SMS. We study the various characteristics of text messages in depth and then found nine rules which can efficiently filter smishing SMS to the legitimate one. We then use rule-based classification algorithm namely Decision Tree, RIPPER and PRISM to apply these rules. In this, we have also identified the minimal effective feature set in the detection of smishing messages. Moreover, we recognise the best rule-based classification algorithm in the classification of smishing messages. The performance of the proposed approach is evaluated, and it achieved more than 99% of true negative rate and 92% true positive rate.

The rest of the paper is organized as follows. Section 2 presents the related work. Section 3 presents the overview of the rule-based approach. Section 4 presents the experimental evaluation. Finally, Section 5 concludes the paper and present the future work.

2. Related Work

This section discusses the various existing mobile phishing detection techniques. The existing mobile phishing detection techniques divide into following categories.

2.1. User Education Based Schemes

The educational based solutions emphasis on educating the mobile users about the characteristics of phishing message through training, workshop and awareness programs so that they correctly identify the phishing attack [8]. However, the phishing attack becomes successful due to human flaws and ignorance. This conceptual knowledge may help the users in avoiding phishing attacks.

2.2. Technical solutions to mitigate mobile phishing attack

The technical solutions are also cost-effective and easy to implement (driven and download) as compare to educational based solutions. In this, Amrutkar et al. [9] proposed a mechanism named KAYO, which differentiates between the malicious and genuine mobile webpages. It detects mobile malicious pages by measuring 44 mobile