

Paper ID: MISS21_87

Email Phishing Using Random Forest

Saugata Choudhury*, Arijita Bhowmik

Techno College of Engineering Agartala, Maheshkhala, Tripura, India

*Corresponding author

Abstract

Background: Phishing is one of the many (and lucrative) scams practiced today. In criminal law, fraud is defined as will full deception committed solely for the purpose of harming personal gain or reputation [1]. Over the years, phishing has become a widespread problem, resulting in countless victims, some of whom are unaware that they have been harmed. The sole purpose of phishing is to obtain confidential information from the victim. There is still no consensus on how best to detect phishing [2].

Objective: In this article, we will look at the problem of using machine learning algorithms such as Support Vector Machine and Random Forest. With this approach, you can use a variety of URL attributes such as "/" character, ASN number, middle character path, and so on to find out if a URL is good or bad. The URL (Uniform Resource Locator), the length of the URL, the popularity of the website, or the content of the page itself [3]. Here we explore and report on how to use Random Forest Machine learning algorithms to classify phishing attacks with the main goal of developing an improved phishing email classifier with better prediction accuracy and fewer features [1].

Methodology: Random Forest is one of the most popular machine learning algorithms used to solve many problems. A guided set of rules used to resolve type and regression problems. It consists of parallel decision trees that take input and generate specific classes.

Result and discussion: There are two main phases in machine learning: the learning phase and the testing phase. The predictive accuracy of the classifier depends only on the information obtained in the learning process. Low GIs result in low accuracy of the estimate, but high GIs lead to high accuracy of the classifier.

Conclusion and future work: In this paper we had mentioned a manner to perceive whether or not a given internet site URL is phishing or not. For this purpose, we used famous device gaining knowledge of random forest. In the future, we will discover a higher manner to discover a phishing internet site through the use of superior capabilities of the URL.

References

- [1] School of Mathematics, Statistics and Computer Science, University of KwaZulu-Natal, Private Bag Box X54001, Durban 4000, South Africa.
- [2] International Journal of Network Security & Its Applications (IJNSA) Vol.8, No.4, July 2016.
- [3] Journal of Critical Reviews ISSN- 2394-5125 vol 7, issue 18, 2020

