A Phishing Email Detection Approach Using Machine Learning Techniques

Authors: Kenneth Fon Mbah, Arash Habibi Lashkari, Ali A. Ghorbani

Abstract: Phishing e-mails are a security issue that not only annoys online users, but has also resulted in significant financial losses for businesses. Phishing advertisements and pornographic e-mails are difficult to detect as attackers have been becoming increasingly intelligent and professional. Attackers track users and adjust their attacks based on users' attractions and hot topics that can be extracted from community news and journals. This research focuses on deceptive Phishing attacks and their variants such as attacks through advertisements and pornographic e-mails. We propose a framework called Phishing Alerting System (PHAS) to accurately classify e-mails as Phishing, advertisements or as pornographic. PHAS has the ability to detect and alert users for all types of deceptive e-mails to help users in decision making. A well-known email dataset has been used for these experiments and based on previously extracted features, 93.11% detection accuracy is obtainable by using J48 and KNN machine learning techniques. Our proposed framework achieved approximately the same accuracy as the benchmark while using this dataset.

Keywords: phishing e-mail, phishing detection, anti phishing, alarm system, machine learning