# An effective security alert mechanism for real-time phishing tweet detection on Twitter

## ABSTRACT

Phishing is a form of social engineering crime uses to deceive victims by directing them to a fraudulent website where their private and confidential information are collected for further illegal actions. Phishing attacks have now targeted users at Online Social Networks (OSN)s such as Twitter, Facebook, Myspace, etc. which traditionally, targeting email users. Twitter has become so prevalent to phishers to spread phishing attacks nowadays due to its vast information dissemination and difficult to be detected unlike email. As such, the effectiveness of security alert to prompt Twitter users for the tweet containing phishing Uniform Resource Locator (URL) in real-time is crucial. Many solutions have been proposed but their effectiveness are inadequate and doubtful. In this paper, we propose an effective security alert mechanism making use of a classification model derived from a supervised machine learning technique of Random Forest (RF) and the identified 11 best classification features yielded 94.75% accuracy higher than 94.56% yielded by other researchers who used more than 11 features trained on the same dataset collected from Twitter. To determine its effectiveness, we used 200 phishing URLs collected from Twitter and PhishTank respectively. From our experiment, we are able to justify that such proposed security alert mechanism managed to prompt 97.50% effectively the security alert to Twitter users in real-time.

**Keyword:** Phishing; Twitter; Classification model; Random Forest (RF); Security alert mechanism