**Oral**

**A prototype of Comparison of the performance of the learning algorithms for verification phishing uniform resource locator (URLs) using machine learning.**

L.M. Kephas, G.T. Nhinda, N. Suresh

Information Technology, School of computing, University of Namibia

Email: [Kephas.loide.m@gmail.com](mailto:Kephas.loide.m@gmail.com); Tel: +264 81 433 9296

Abstract

Phishing is an online criminal act that occurs when a malicious webpage mimics as a legitimate webpage to acquire sensitive information from the user [1]. Detecting phishing websites is one of the crucial problems facing the internet community because of its high impact on the day-to-day online transactions performed. Therefore, the objective of this study was to evaluate the performance of three learning algorithms for verification of phishing URLs using three machine learning algorithms which were Naïve Bayesian, Decision tree, and Logistic regression to determine which of the three algorithms is best for determining phishing URLs. Personal extreme programming (PXP) development methodology is used for this prototype. This methodology is designed to be applied by software engineers individually and is iterative and applying its practices allows the developer to be more flexible and responsive to changes [2]. The experiment was performed using 7030 URLs, which were divide into two samples: training and testing, 80% for training and 20% for testing and the observed result showed that decision tree provided the best accuracy of 90% as compared to the other two.

**Keywords**- Phishing URL, Legitimate, Machine learning, Algorithms, verification, lexical features.

References

[1] S. Marchal, J. François, R. State, and T. Engel, “PhishStorm: Detecting Phishing With Streaming Analytics,” *IEEE Trans. Netw. Serv. Manag.*, vol. 11, no. 4, pp. 458–471, Dec. 2014.

[2] Y. Dzhurov, I. Krasteva, and S. Ilieva, “Personal Extreme Programming – An Agile Process for Autonomous Developers,” p. 8.

Uniform resource locator (URL) is the address of a resource on the Internet. A URL indicates the location of a resource as well as the protocol used to access it