

Detecting Phishing Website using Machine Learning

Prof. K. S. Mulani¹, Vaibhav Shewale², Mansi Salve³, Niranjan Kale⁴, Komal Shinde⁵

Professor, Department of Information Technology¹

Students, Department of Information Technology^{2,3,4,5}

Sinhgad Institute of Technology, Lonavala, Maharashtra, India

Abstract: *Phishing attacks continue to pose a major threat for computer system defenders, often forming the first step in a multi-stage attack. There have been great strides made in phishing detection; however, some phishing emails appear to pass through filters by making simple structural and semantic changes to the messages. We tackle this problem through the use of a machine learning classifier operating on a large corpus of phishing and legitimate emails. We design SAFEPC (Semi-Automated Feature generation for Phish Classification), a system to extract features, elevating some to higher level features, that are meant to defeat common phishing email detection strategies. To evaluate SAFE-PC, we collect a large corpus of phishing emails from the central IT organization at a tier-1 university. The execution of SAFE-PC on the dataset exposes hitherto unknown insights on phishing campaigns directed at university users. SAFEPC detects more than 70% a state-of-the-art email filtering tool. It also outperforms Spam Assassin, a commonly used email filtering tool. We also developed an online version of SAFE-PC, that can be incrementally retrained with new samples. Its detection performance improves with time as new samples are collected, while the time to retrain the classifier stays constant.*

Keywords: Detecting Phishing Website, Website management, Safety tips, Safety Requirement.

REFERENCES

- [1]. "WC-PAD: Web Crawling based Phishing Attack Detection" Nathezhtha.T, Sangeetha.D, Vaidehi.V
- [2]. "Detection of Phishing Attacks with Machine Learning Techniques in Cognitive Security Architecture" Ivan Ortiz-Garces, Roberto O. Andrade, and Maria Cazares
- [3]. "A Methodical Overview on Phishing Detection along with an Organized Way to Construct an Anti-Phishing Framework" Srushti Patil, Sudhir Dhage
- [4]. "A survey of the QR code phishing: the current attacks and countermeasures" Kelvin S. C. Yong, Kang Leng Chiew and Choon Lin Tan
- [5]. "Fuzzy Rough Set Feature Selection to Enhance Phishing Attack Detection" Mahdieh Zabihimayvan and Derek Doran

BIOGRAPHY

- Vaibhav Shewale- an Undergraduate Scholar pursuing Bachelors of Engineering in Information Technology from Sinhgad Institute of Technology. He is working under the guidance of Prof. K. S. Mulani
- Mansi Salve- an Undergraduate Scholar pursuing Bachelors of Engineering in Information Technology from Sinhgad Institute of Technology. She is working under the guidance of Prof. K. S. Mulani
- Niranjan Kale - An Undergraduate Scholar pursuing Bachelors of Engineering in Information Technology from Sinhgad Institute of Technology. He is working under the guidance of Prof. K. S. Mulani
- Komal Shinde - An Undergraduate Scholar pursuing Bachelors of Engineering in Information Technology from Sinhgad Institute of Technology. She is working under the guidance of Prof. K. S. Mulani