Detecting Phishing Websites using Machine Learning

Ninand More[1], Dhruv Solanki[2], Devang Solanki[3], Shanil Jain[4], Praajwal Kadu[5]
Students, Department of Information Technology[2,3,4,5]
Assistant Professor, Department of Information Technology[1]
D. Y. Patil College of Engineering, Pune, Maharashtra, India

Abstract: *The criminals, who want to obtain sensitive data, first create unauthorized replicas of a real website and e-mail. The e-mail will be created using logos and slogans of a legitimate company. The nature of website creation is one of the reasons that the Internet has grown so rapidly as a communication medium. Phisher then send the "spoofed" e-mails to as many people as possible in an attempt to lure them into the scheme. When these e-mails are opened or when a link in the mail is clicked, the consumers are redirected to a spoofed website, appearing to be from the legitimate entity. We discuss the methods used for detection of phishing Web sites based on url importance properties.*

Keywords: Machine Learning, Phishing Website Detection, Phishing, URL Detection

REFERENCES

[1]. K. L. Chiew, C. L. Tan, K. Wong, K. S. C. Yong, and W. K. Tiong, ''A new hybrid ensemblefeature selection framework for machine learning-based phishing detection system,'' Inf. Sci.,vol. 484, pp. 153–166, May2019.

[2]. O. K. Sahingoz, E. Buber, O. Demir, and B. Diri, ''Machine learning based phishing detectionfrom URLs,'' Expert Syst. Appl., vol. 117, pp. 345–357, Mar. 2019.

[3]. R. M. Mohammad, F. Thabtah, and L. McCluskey, ''Tutorial and critical analysis of phishing Websites methods,''Comput. Sci. Rev., vol. 17, pp. 1–24, Aug. 2015.

[4]. K. L. Chiew, K. S. C. Yong, and C. L. Tan, ''A survey of phishing attacks: Their types, vectors and technical approaches,'' Expert Syst. Appl., vol. 106, pp. 1–20, Sep. 2018.

[5]. A. Aleroud and L. Zhou, ''Phishing environments, techniques, and countermeasures: A survey,''Comput. Secur., vol. 68, pp. 160–196, Jul. 2017.

[6]. V. Suganya, ''A review on phishing attacks and various anti phishing techniques,'' Int. J.Comput. Appl., vol. 139, no. 1, pp. 20–23, Apr. 2016.

[7]. N. Abdelhamid, A. Ayesh, and F. Thabtah, ''Phishing detection based associative classification data mining,'' Expert Syst. Appl., vol. 41, no. 13, pp. 5948–5959, Oct. 2014.

[8]. I. Qabajeh and F. Thabtah, ''An experimental study for assessing email classification attributes using feature selection methods,''in Proc. 3rd Int. Conf. Adv. Comput. Sci. Appl. Technol. (ACSAT), Dec. 2014, pp. 125–132.

[9]. R. S. Rao and A. R. Pais, ''Detection of phishing Websites using an efficient feature-based machine learning framework,'' Neural Comput. Appl., vol. 31, pp. 3851–3873, Jan. 2018.

[10]. W. Ali, ''Phishing Website detection based on supervised machine learning with wrapper features selection,'' Int. J. Adv. Comput. Sci. Appl., vol. 8, no. 9, pp. 72–78, 2017.

[11]. APWG. Phishing Activity Trends Report 3rd Quarter 2019. Accessed: Mar. 21, 2020. [Online]. Available: https://docs.apwg.org/ reports/apwg_trends_report_q3_2019.pdf Dept. of