Web Phishing Detection

Literature Survey

Team Members: -

A.P. Sai Nandhan

Yashwanth B.T.

Muhammed Rameez M

Sanjeev Karthick

Paper 1: Intelligent Web-Phishing detection and protection scheme using integrated features of images, frames, and text

Author:

Moruf A Adebowale, M Alamgir Hossain

Published:

Received 25 April 2018, Revised 29 July 2018, Accepted 29 July 2018, Available online 4 August 2018, Version of Record 17 August 2018, scienceDirect.com

A phishing attack is one of the most significant problems faced by online users because of its enormous effect on the online activities performed. In recent years, phishing attacks continue to escalate in frequency, severity and impact. Several solutions, using various methodologies, have been proposed in the literature to counter the web-phishing threats. Notwithstanding, the

existing technology cannot detect the new phishing attacks accurately due to the insufficient integration of features of the text, image and frame in the evaluation process. The use of related features of images, frames and text of legitimate and non-legitimate websites and associated artificial intelligence algorithms to develop an integrated method to address these together.

Paper 2: A Survey and Classification of Web phishing detection schemes

Author:

Gaurav Varshney, Manoj Misra, Pradeep K. Atrey

Published:

26 October 2016, Wiley Online Library.com

Phishing is a fraudulent technique that is used over the Internet to deceive users with the goal of extracting their personal information such as username, passwords, credit card, and bank account information. The key to phishing is deception. Phishing uses email spoofing as its initial medium for deceptive communication followed by spoofed websites to obtain the needed informationfrom the victims. Phishing was discovered in 1996, and today, it is one of the

most severe cybercrimes faced by the Internet users. Researchers are working on the prevention, detection, and education of phishing attacks, but to date, there is no complete and accurate solution for thwarting them.

Paper 3: Web Phishing Detection Using a Deep Learning Framework

Author:

Tony T. Luo, Ting Zhu, Wei Wang, Futai Zou

Published:

26 September 2018, Hindawi.com

Web service is one of the key communications software services for the Internet. Web phishing is one of many security threats to web services on the Internet. Web phishing aims to steal private information, such as usernames, passwords, and credit card details, by way of impersonating a legitimate entity. It will lead to information disclosure and property damage. This paper mainly focuses on applying a deep learning framework to detect phishing websites. This paper first designs two types of features for web phishing: original features and interaction features. A detection model based on Deep

Belief Networks (DBN) is then presented. The test using real IP flows from ISP (Internet Service Provider) shows that the detecting model based on DBN can achieve an approximately 90% true positive rate and 0.6% false positive rate.

Paper 4: Systemation of Knowledge: A systematic review of software-based web phishing detection

Author:

Z. Dou, I. Khalil, A. Khreishah, A. Al-Fuqaha and M. Guizani.

Publisher:

"Systematization of Knowledge (SoK): A Systematic Review of Software-Based Web Phishing Detection," in IEEE Communications Surveys & Tutorials, vol. 19, no. 4, pp. 2797-2819, Fourth quarter 2017, doi: 10.1109/COMST.2017.2752087, "ieeexplore.ieee.org"

Phishing is a form of cyberattack that leverages social engineering approaches and other sophisticated techniques to harvest personal information from users of websites. The average annual growth rate of the number of unique phishing websites detected by the Anti Phishing Working Group is 36.29% for the past six years and 97.36% for the past two years. In the wake of this rise, alleviating phishing

attacks has received a growing interest from the cyber security community. Extensive research and development have been conducted to detect phishing attempts based on their unique content, network, and URL characteristics.