WEB PHISHING DETECTION USING MACHINE LEARNING

LITERATURE SURVEY:

This article surveys the literature on the detection of phishing attacks. Phishing attacks target vulnerabilities that exist in systems due to the human factor. Many cyber attacks are spread via mechanisms that exploit weaknesses found in endusers, which makes users the weakest element in the security chain.

M. Khonji, Y. Iraqi, and A. Jones are with Khalifa University, Defining the phishing problem. It is important to note that the phishing definition in the literature is not consistent, and thus a comparison of a number of definitions is presented. Categorizing anti-phishing solutions from the perspective of phishing campaign life-cycle. This presents the various anti-phishing solution categories such as detection. It is important to view the overall anti-phishing picture from a high-level perspective before diving into a particular technique, namely: phishing detection techniques (which is the scope of this survey). Presenting evaluation metrics that are commonly used in the phishing domain to evaluate the performance of phishing detection techniques. This facilitates the comparison between the various phishing detection techniques. Presenting a literature survey of anti-phishing detection techniques, which incorporates software detection techniques as well as user-awareness techniques that enhance the detection process of phishing attacks. Presenting a comparison of the various proposed phishing detection techniques in the literature.

REFERENCES:

[1] S. Sheng, M. Holbrook, P. Kumaraguru, L. F. Cranor, and J. Downs, “Who falls for phish?: a demographic analysis of phishing susceptibility and effectiveness of interventions,” in Proceedings of the 28th international conference on Human factors in computing systems, ser. CHI ’10. New York, NY, USA: ACM, 2010, pp. 373–382.

[2] B. Krebs, “HBGary Federal hacked by Anonymous,” http: //krebsonsecurity.com/2011/02/hbgary-federal-hacked-by-anonymous/, 2011, accessed December 2011.

[3] B. Schneier, “Lockheed Martin hack linked to RSA’s SecurID breach,” http://www.schneier.com/blog/archives/2011/05/lockheed martin.html, 2011, accessed December 2011.

[4] C. Whittaker, B. Ryner, and M. Nazif, “Large-scale automatic classification of phishing pages,” in NDSS ’10, 2010.

[5] X. Dong, J. Clark, and J. Jacob, “Modelling user-phishing interaction,” in Human System Interactions, 2008 Conference on, may 2008, pp. 627 –632.

[6] W. D. Yu, S. Nargundkar, and N. Tiruthani, “A phishing vulnerability analysis of web based systems,” in Proceedings of the 13th IEEE Symposium on Computers and Communications (ISCC 2008). Marrakech, Morocco: IEEE, July 2008, pp. 326–331.

[7] Anti-Phishing Working Group (APWG), “Phishing activity trends report — second half 2010,” http://apwg.org/reports/apwg report h2 2010. pdf, 2010, accessed December 2011.

[8] Anti-Phishing Working Group (APWG), “Phishing activity trends report — first half 2011,” http://apwg.org/reports/apwg trends report h1 2011.pdf, 2011, accessed December 2011.

[9] Anti-Phishing Working Group (APWG), “Phishing activity trends report — second half 2011,” http://apwg.org/reports/apwg trends report h2 2011.pdf, 2011, accessed July 2012.

[10] B. Schneier, “Details of the RSA hack,” http://www.schneier.com/blog/ archives/2011/08/details of the.html, 2011, accessed December 2011.

[11] P. Kumaraguru, Y. Rhee, A. Acquisti, L. F. Cranor, J. Hong, and E. Nunge, “Protecting people from phishing: the design and evaluation of an embedded training email system,” in Proceedings of the SIGCHI conference on Human factors in computing systems, ser. CHI ’07. New York, NY, USA: ACM, 2007, pp. 905–914.

[12] A. Alnajim and M. Munro, “An anti-phishing approach that uses training intervention for phishing websites detection,” in Proceedings of the 2009 Sixth International Conference on Information Technology: New Generations. Washington, DC, USA: IEEE Computer Society, 2009, pp. 405–410.

[13] S. Gorling, “The Myth of User Education,” Proceedings of the 16th Virus Bulletin International Conference, 2006.

[14] G. Gaffney, “The myth of the stupid user,” http://www.infodesign.com. au/articles/themythofthestupiduser, accessed March 2011.

[15] A. Stone, “Natural-language processing for intrusion detection,” Computer, vol. 40, no. 12, pp. 103 –105, dec. 2007.