IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 1, January 2024

Detection of Phishing Website using Machine Learning

Vaishnavi Bhoyar[1], Komal Dharak[2], Dipali Gawali[3]
Department of Computer Science[1,2,3]
Nutan Maharashtra Institute of Engineering and Technology, Pune, India

Abstract: *Phishing is a widespread cybercrime where scammers trick people into sharing personal or confidential information by pretending to be a legitimate website. Despite various machine learning methods developed for detecting phishing websites using features from web samples, not much attention has been given to picking the right features efficiently. This study aims to figure out the crucial features necessary for effective phishing detection, improving the accuracy and efficiency of machine learning systems. By identifying and prioritizing these features, our research contributes to creating simpler methods that keep users safe from phishing threats. We focus on pinpointing the key characteristics that consistently set phishing websites apart from legitimate ones, enhancing the precision and reliability of phishing detection algorithms through careful analysis and experimentation. This work aligns with broader goals of strengthening cybersecurity measures, protecting individuals and organizations from falling victim to online deception, and giving users more robust tools for secure online navigation.To evaluate the feature selection in developing ageneralizable phishing detection, these classifiers are trained by a separate out-of-sample data set of 14,000 website samples. The maximum F-measure gained feature selection is 95% using Random Forest classification. Also, there are 9 universal features selected over all the three data sets. The F-measure value using this universal feature set is approximately 93% which is a comparable result in contrast to performance. Since the universal feature set contains no features from third-part services, this finding implies that with no inquiry from external sources, we can gain a fast phishing detection which is also robust toward zero-day attacks.*

Keywords: Phishing Detection, Feature Selection, Generalizable phishing detection

I. INTRODUCTION

Phishing is a significant security concern employing clever tricks to lure individuals into clicking harmful links and divulging sensitive information. Despite their relatively uncomplicated technical nature, these attacks prove remarkably effective due to attackers' creation of deceptive websites mirroring legitimate ones, making detection challenging.

With technological advancements, phishing attacks have become more sophisticated, posing a growing challenge for ordinary users to distinguish fake emails or websites. Our Android project, centered around machine learning, addresses this issue. Users can log in, employ machine learning algorithms to assess URL legitimacy, and, upon confirming phishing attempts, submit complaints. Additionally, an officer login allows for reviewing and addressing complaints, taking necessary actions against the identified phishing URLs and their sources.

By integrating machine learning into our Android application, we aim to provide a user-friendly tool that enhances online security, enabling individuals to identify and report phishing attempts. This collaborative effort seeks to reduce the prevalence and success of phishing attacks, benefiting both individuals and organizations.

II. CONCLUSION

Its successfully demonstrated the potential of machine learning in detecting phishing websites. However, ongoing efforts are necessary to refine the model, adapt to evolving threats, and integrate the solution into existing cybersecurity frameworks for comprehensive protection against phishing attacks.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 1, January 2024

III. ACKNOWLEDGMENT

We Hereby declare that this is our own work and we're still in the process of enforcing this idea. This is the idea we've presented but still needs to be worked upon.

REFERENCES

[1] R. Basnet, S. Mukkamala, and A. H. Sung, "Detection of phishing attack: A machine learning approach," in Soft Computing Applications in Industry. Springer, 2008, pp. 373–383.

[2] R. S. Rao and A. R. Pais, "Detection of phishing websites using anefficient feature-based machine learning framework," Neural Computingand Applications, pp. 1–23, 2018.

[3] T. Peng, I. Harris, and Y. Sawa, "Detecting phishing attacks using naturallanguage processing and machine learning," in Semantic Computing(ICSC), 2018 IEEE 12th International Conference on. IEEE, 2018,pp. 300–301.

[4] R. Sadeghi, T. Banerjee, and W. Romine, "Early hospital mortalityprediction using vital signals," Smart Health, vol. 9-10, pp. 265–274,2018.

[5] R. Sadeghi andJ. Hamidzadeh, " Automatic support vector data description, " Soft Computing,vol. 22,no. 1,pp. 147 – 158, 2018.

[6] M. Zabihimayvan, R. Sadeghi, H. N. Rude, and D. Doran, "A softcomputing approach for benign and malicious web robot detection,"Expert Systems with Applications, vol. 87, pp. 129–140, 2017.

[7] J. Hamidzadeh, M. Zabihimayvan, and R. Sadeghi, "Detection of website visitors based on fuzzy rough sets," Soft Computing, vol. 22, no. 7,pp. 2175–2188, 2018.