Plagiarism - Report

Originality Assessment

21%

Overall Similarity

Date: Mar 17, 2024
Matches: 907 / 4260 words
Sources: 48

Remarks: Moderate similarity detected, you better improve the document (if needed).

Verify Report:

Phish Catcher: Client-Side Defense Against Web-Spoofing Attacks Using Machine Learning

1. SUVVARI PAVAN KUMAR, 2. VUDUTALA SRIHARI, 3. KADAVATI MANOHAR, 4. PAPPALA SRINIDHI, 5. DR.K.N.S. LAKSHMI (Professor)

Department of Computer Science Engineering

Sanketika Vidya Parishad Engineering College, Visakhapatnam, India

ABSTRACT

Cybersecurity faces a significant challenge in safeguarding users' confidential information, such as passwords and PIN codes, from phishing attacks. These attacks, which employ various deceptive tactics like fake login pages, phishing emails, and click-jacking, aim to trick users into divulging sensitive data. Traditional security strategies often encounter issues of latency and accuracy in detecting these fraudulent activities. To address this challenge, we propose a client-side defense mechanism leveraging machine learning techniques to detect spoofed web pages and protect users from phishing attacks.

In this work, we introduce PhishCatcher, a Google Chrome extension developed as a proof of concept for our machine learning-based approach. PhishCatcher utilizes a random forest classifier trained on four types of web features to classify URLs as suspicious or trustworthy. To evaluate the effectiveness of our extension, we conducted experiments on real web applications, testing 400 classified phishing URLs and 400 legitimate URLs. The results demonstrate a remarkable accuracy and precision of 98.5% for detecting spoofed web pages.

Furthermore, we assessed the latency of PhishCatcher by measuring its response time over forty phished URLs. The average recorded response time was just 62.5 milliseconds, indicating minimal impact on user experience while providing robust protection against phishing attacks.

Overall, our approach offers a highly accurate and efficient solution for detecting phishing attempts, thereby enhancing user security and mitigating the risks associated with online fraud.

INTRODUCTION

In October 2022, members 21 of the National Institute for Research in Digital Science and Technology (Inria) in France fell victim to a phishing attack. They received an email in French prompting them to confirm their webmail account, with a link that directed them to a fake login page resembling the legitimate central authentication login page of Inria. This incident highlights the persistent threat of phishing attacks, which aim to deceive users into disclosing sensitive information such as passwords.

With the rapid advancement of technology, the online world has witnessed significant growth in various sectors including e-commerce, online banking, distance learning, e-health, and e-governance. As a result, billions of users have embraced this digital trend, creating personalized accounts on numerous websites to access specialized services. However, this convenience comes with risks, as users are often required to provide personal information, including usernames and passwords, to login to these accounts.

Phishing attacks exploit this vulnerability by impersonating legitimate websites and tricking users into divulging their confidential information. Attackers employ various techniques, including email phishing, trojan horses, keyloggers, and man-in-the-middle proxies, to steal valuable data 12 such as login credentials. These attacks not only pose a threat to individual privacy but also endanger national security, intellectual property, and organizational secrets.

Traditional security measures, including firewalls, digital certificates, encryption software, and two-factor authentication, have proven insufficient in combating sophisticated phishing attacks. While server-side solutions may offer protection, they often require extensive modifications to websites and are prone to oversight by developers. Therefore, client-side solutions have emerged as an alternative approach to safeguarding users without the need for server support.

Anti-phishing tools can be classified based on their detection mechanisms, including blacklists, heuristics, and machine learning. While blacklists provide high accuracy, they may miss zero-day attacks and are susceptible to spam URLs. Heuristic-based techniques offer promising results in identifying phishing sites, but their latency increases over time. In contrast, machine learning-based approaches leverage statistical properties of training data to classify URLs as legitimate or malicious.

In this paper, we propose PhishCatcher, a stateless client-side tool designed to protect against web spoofing attacks using machine learning techniques. PhishCatcher is implemented as a Google Chrome extension, employing the random forest algorithm to classify login web pages as either legitimate or spoofed. We conducted experiments to evaluate the effectiveness and accuracy of PhishCatcher on real web applications, yielding remarkable results.

The contributions of this research include the proposal and development of a client-side anti-phishing mechanism, the design and implementation of the PhishCatcher Google Chrome extension, careful selection of web features for the phishing classifier algorithm, and experimental analysis of Phish Catcher's performance.

The remainder of the paper is organized as follows: Section II provides a summary of related work in the literature, Section III discusses the research methodology, Section IV describes the design and development of the Google Chrome extension, Section V presents the testing results, Section VI evaluates the extension, and Section VII concludes the paper.

RELATED WORK

Numerous techniques and tools have been developed to mitigate the risks posed by phishing attacks. This section provides an overview of existing anti-phishing tools and frameworks, categorized into seven major schemes:

A. Visual Similarity and Page Content Investigation:

- SpoofCatch: Utilizes visual similarity to identify phishing websites based on screenshots of login

pages.

- Strategies based on visual distinction between phishing and legitimate websites using text, layout, and images.

- Content-focused methodologies employing Term Frequency-Inverse Document Frequency (TF-IDF) filter and Gestalt philosophy.

- PWDHASH++: Analyzes visual similarities between websites based on Gestalt philosophy.

B. Hybrid Approach for Phishing Detection:

- Dynamic Category Decision Algorithm (DCDA): Utilizes deep learning for phishing detection.

- Hybrid machine learning models combining multiple techniques for improved effectiveness.

- Repeated Incremental Pruning to Produce Error Reduction (RIPPER) algorithm for malicious email detection.

C. Anti-Phishing [2] Machine Learning Techniques:

- Various machine learning algorithms such as logistic regression, random forest, and support vector machines (SVM) used for phishing detection.

- Application of machine learning on noisy datasets for effective detection.

- Implementation of supervised learning techniques for the classification of malicious websites.

D. Online Training Procedures Preventing Phishing:

- Studies evaluating [1] the effectiveness of online learning strategies and phishing email filters.

- Methods for identifying phishing websites based on website source code analysis.

- Development of embedded email training schemes for educating users about phishing threats.

E. Automated Classification of Fake and Genuine Websites:

- Automated Individual White List (AIWL): Maintains a white list of known legitimate websites.

- [1] Scalable machine learning classifiers for dynamic management of website blacklists.

F. URL Analysis for Detecting Phishing:

- Lightweight URL-based phishing detection approaches leveraging supervised learning.
- Lexical evaluation of URL tokens and tokenization methods for improved prediction efficiency.

G. Significant Anti-Phishing Tools:

- Spoofguard: Browser extension displaying photographic passwords and warning users of potential scams.
- BOGUSBITER: Anti-phishing tool employing an offensive defense strategy by feeding bogus data to malicious sites.
- MadTracer: Detection tool for malvertising attacks, capturing harmful domain tracks more effectively than existing solutions.
- Prophiler: Framework 3 for reducing the number of web pages needing evaluation to identify harmful websites.
- DAISY: Lightweight identification and prevention system for defending software-defined networks against DoS attacks.

These tools and techniques demonstrate a diverse range of approaches 2 to combat phishing attacks, incorporating visual similarity analysis, machine learning algorithms, URL analysis, and online training procedures. Continued 3 research and development in this field are essential to stay ahead of evolving phishing threats and protect users' online security and privacy.

RESEARCH METHODOLOGY

In our research methodology, we conducted an extensive review of relevant literature to understand the current state-of-the-art in phishing attacks, web spoofing, machine learning, and various detection mechanisms. Subsequently, we explored several machine learning-based frameworks for detecting malicious login pages and compared them with our proposed plug-ins. Additionally, we performed 1 Document Object Model (DOM) analysis and utilized JavaScript and Python to develop a sophisticated Google Chrome extension for detecting spoofing attacks. The primary objective was to

create a browser add-on that acts as a classifier for fake and authentic login pages, providing phishing warnings to users in real-time.

A. Model Selection:

We selected the random forest classifier for our model, as it has shown superior performance in detecting phishing attempts compared to other techniques. While data mining-based methods are effective in identifying phishing attacks, implementing them directly in browsers for real-time detection presents challenges. Unlike conventional server-based approaches, our proposed method runs the classification algorithm inside the browser, offering benefits such as better privacy and independence from network latency.

B. Pre-processing:

For feature extraction, we utilized data from various sources, including 2 the UCI Machine Learning Repository, a collection of hijacked journal websites, blacklisted URLs from PhishTank, and genuine URLs from moz.com/top500.

C. Features Collection:

We faced challenges in selecting suitable 1 features due to the absence of well-fitting datasets and disagreement in the literature regarding ultimate distinguishing attributes of phished websites. Despite this, we curated a feature set based on a thorough analysis of existing strategies, focusing on address bar, abnormal, HTML and JavaScript, and domain-based features.

D. Classification and Classifier Selection:

We employed 1 a supervised learning approach for classification, with the random forest algorithm selected due to its versatility, ease of use, and superior performance. 4 The random forest algorithm creates an ensemble of decision trees, reducing overfitting and improving accuracy. Our proposed

model utilizes the random forest classifier to identify potential phishing attacks and alert users through the PhishCatcher browser extension.

In summary, our research methodology involved a comprehensive review of literature, development of a sophisticated browser extension using JavaScript, and implementation of the random forest classifier for real-time phishing detection, aiming to enhance user privacy and security during online browsing.

PLUGIN DESIGN

Browser extensions or add-ons are small software packages that can modify and enhance the browsing experience according to the user's preferences. They are typically developed using web-based programming languages such as HTML, CSS, and JavaScript. In this section, we provide an overview of the design and development of our tool PhishCatcher, a Google Chrome extension aimed at identifying and protecting against phishing attacks. The main concept behind PhishCatcher is to perform classification within the client's browser and display the results in real-time, thus improving latency and preserving user privacy.

TESTING

To evaluate the performance of PhishCatcher, we conducted testing against real web application scenarios. Instead of applying unit testing for each feature individually, we focused on aggregated analysis of all features considered for classifying legitimate and bogus URLs. However, we present screenshots of a few tested URLs captured by PhishCatcher to provide insight into its performance.

Dataset:

Legitimate and corresponding fake URLs of 90 hijacked journals.

310 blacklisted URLs from PhishTank.

310 legitimate URLs from https://moz.com/top500.

After multiple experiments, we identified seventeen prominent features, categorized as shown in

Table 2. Some of these features have been previously utilized in different tools and analyses.

Test Cases:

Test Case 1:

URL: https://www.education-online.nl/Cliquez.ici.cas.inria.fr.cas.login/login.html

Result: Phishing

Description: This test case involves a sophisticated phishing attack on Inria. PhishCatcher correctly identifies the phishing attempt, distinguishing between the genuine and 12 fake login pages. Six features contributed to this identification, including URL length, domain prefix/suffix length, favicon, request URL, anchor, and script link.

Test Case 2:

URL: http://www.ijiq.com

Result: Phishing

Description: In this scenario, a spam 2 URL of a hijacked journal is used to deceive users. PhishCatcher successfully detects the phishing attempt, alerting users when accessing the bogus URL. 5 Features such as URL length, domain prefix/suffix length, favicon, request URL, and anchor played a role in identifying the phishing URL.

Test Case 3:

URL: http://www.revistas-academicas.com

Result: Phishing

Description: Another phishing attempt involving a hijacked journal URL, where PhishCatcher 1 accurately identifies the phishing attack. The features responsible for detection are detailed in Table 4.

Test Case 4:

Authentic Web Page: http://www.ahistcon.org/revistaayer.html

Counterfeit Web Page: http://www.ayeronline.com

Description: This test case evaluates Phish Catcher's performance by comparing genuine and corresponding hijacked URLs for the same journal. PhishCatcher correctly identifies the genuine URL while also detecting the spam URL of the hijacked journal.

Test Case 5:

Description: Testing results for top-ranked legitimate websites such as Facebook, Google, Microsoft, and Apple. PhishCatcher correctly identifies these websites as safe to use.

Evaluation:

The proposed model was tested over multiple trials to assess accuracy and latency. Latency experiments were conducted, and the results were recorded in the form of a confusion matrix for further calculation of precision, recall, and accuracy of the model.

Conclusion:

In today's digital landscape, users heavily rely on online applications across various domains, including banking, e-commerce, social networking, education, and more. However, the rise of sophisticated web spoofing attacks poses significant security and privacy risks to users. While several tools exist to combat these attacks, many have shortcomings.

We have developed PhishCatcher, an optimized and user-friendly browser plug-in, to intelligently detect phishing attacks using supervised machine learning. Unlike traditional approaches, PhishCatcher runs the classification directly in the browser, addressing latency issues and enhancing tool efficiency. The plug-in's simple user interface provides clear phishing alerts and highlights corresponding phishing features, aiding user understanding.

PhishCatcher utilizes a feature set of thirty features categorized into four groups, each acting as a decision tree. A random forest classifier aggregates these decision trees' outcomes to identify fake and genuine login pages. Testing and evaluation involved a dataset of 400 malicious and 400 legitimate URLs, assessed using a confusion matrix, yielding impressive results with 98.5% precision, 98.5% recall, and 98.5% accuracy. Additionally, the plug-in exhibited low latency, averaging just 62.5

milliseconds over forty phishing URLs.

Future Work:

While PhishCatcher has demonstrated strong performance, there are opportunities for further enhancement:

Feature Expansion: Adding more automated features could improve overall performance.

Classifier Diversity: Implementing other discriminative classifiers such as Support Vector Machines (SVM) could enhance prediction accuracy, especially with larger datasets.

Evaluation Metrics: Evolving evaluation metrics using different tools for more comprehensive performance analysis would provide deeper insights.

Continued research and development in these areas will contribute to advancing phishing detection techniques, ultimately bolstering users' online security and privacy.

REFERENCES

[1] W. Khan, A. Ahmad, A. Qamar, M. Kamran, and M. Altaf, ‘‘SpoofCatch: A client-side protection tool against phishing attacks,’’ IT Prof., vol. 23, no. 2, pp. 65–74, Mar. 2021.

[2] B. Schneier, ‘‘Two-factor authentication: Too little, too late,’’ Commun. ACM, vol. 48, no. 4, p. 136, Apr. 2005.

[3] S. Garera, N. Provos, M. Chew, and A. D. Rubin, ‘‘A framework for detection and measurement of phishing attacks,’’ in Proc. ACM Workshop Recurring malcode, Nov. 2007, pp. 1–8.

[4] R. Oppliger and S. Gajek, ‘‘Effective protection against phishing and web spoofing,’’ in Proc. IFIP Int. Conf. Commun. Multimedia Secur. Cham, Switzerland: Springer, 2005, pp. 32–41.

[5] T. Pietraszek and C. V. Berghe, ‘‘Defending against injection attacks through context-sensitive string evaluation,’’ in Proc. Int. Workshop Recent Adv. Intrusion Detection. Cham, Switzerland: Springer, 2005, pp. 124–145.

[6] M. Johns, B. Braun, M. Schrank, and J. Posegga, ‘‘Reliable protection against session fixation attacks,’’ in Proc. ACM Symp. Appl. Comput., 2011, pp. 1531–1537.

[7] M. Bugliesi, S. Calzavara, R. Focardi, and W. Khan, ‘‘Automatic and robust client-side

protection for cookie-based sessions,'' in Proc. Int. Symp. Eng. Secure Softw. Syst. Cham, Switzerland: Springer, 2014, pp. 161–178.

[8] A. Herzberg and A. Gbara, ''Protecting (even naıve) web users from spoofing and phishing attacks,'' Cryptol. ePrint Arch., Dept. Comput. Sci. Eng., Univ. Connecticut, Storrs, CT, USA, Tech. Rep. 2004/155, 2004.

[9] N. Chou, R. Ledesma, Y. Teraguchi, and J. Mitchell, ''Client-side defense against web-based identity theft,'' in Proc. NDSS, 2004, 1–16.

[10] B. Hämmerli and R. Sommer, Detection of Intrusions and Malware, and Vulnerability Assessment: 4th International Conference, DIMVA 2007 Lucerne, Switzerland, July 12-13, 2007 Proceedings, vol. 4579. Cham, Switzerland: Springer, 2007.

[11] C. Yue and H. Wang, ''BogusBiter: A transparent protection against phishing attacks,'' ACM Trans. Internet Technol., vol. 10, no. 2, pp. 1–31, May 2010.

[12] W. Chu, B. B. Zhu, F. Xue, X. Guan, and Z. Cai, ''Protect sensitive sites from phishing attacks using features extractable from inaccessible phishing URLs,'' in Proc. IEEE Int. Conf. Commun. (ICC), Jun. 2013, pp. 1990–1994.

[13] Y. Zhang, J. I. Hong, and L. F. Cranor, ''Cantina: A content-based approach to detecting phishing web sites,'' in Proc. 16th Int. Conf. World Wide Web, May 2007, pp. 639–648.

[14] D. Miyamoto, H. Hazeyama, and Y. Kadobayashi, ''An evaluation of machine learning-based methods for detection of phishing sites,'' in Proc. Int. Conf. Neural Inf. Process. Cham, Switzerland: Springer, 2008, pp. 539–546.

[15] E. Medvet, E. Kirda, and C. Kruegel, ''Visual-similarity-based phishing detection,'' in Proc. 4th Int. Conf. Secur. privacy Commun. Netowrks, Sep. 2008, pp. 1–6.

[16] W. Zhang, H. Lu, B. Xu, and H. Yang, ''Web phishing detection based on page spatial layout similarity,'' Informatica, vol. 37, no. 3, pp. 1–14, 2013.

[17] J. Ni, Y. Cai, G. Tang, and Y. Xie, ''Collaborative filtering recommendation algorithm based on TF-IDF and user characteristics,'' Appl. Sci., vol. 11, no. 20, p. 9554, Oct. 2021.

[18] W. Liu, X. Deng, G. Huang, and A. Y. Fu, ''An antiphishing strategy based on visual

similarity assessment,'' IEEE Internet Comput., vol. 10, no. 2, pp. 58–65, Mar. 2006.

[19] A. Rusu and V. Govindaraju, ''Visual CAPTCHA with handwritten image analysis,'' in Proc. Int. Workshop Human Interact. Proofs. Berlin, Germany: Springer, 2005, pp. 42–52.

[20] P. Yang, G. Zhao, and P. Zeng, ''Phishing website detection based on multidimensional features driven by deep learning,'' IEEE Access, vol. 7, pp. 15196–15209, 2019.

[21] P. Sornsuwit and S. Jaiyen, ''A new hybrid machine learning for cybersecurity threat detection based on adaptive boosting,'' Appl. Artif. Intell., vol. 33, no. 5, pp. 462–482, Apr. 2019.

[22] S. Kaur and S. Sharma, ''Detection of phishing websites using the hybrid approach,'' Int. J. Advance Res. Eng. Technol., vol. 3, no. 8, pp. 54–57, 2015.

[23] W. W. Cohen, ''Fast effective rule induction,'' in Machine Learning Proceedings. Amsterdam, The Netherlands: Elsevier, 1995, pp. 115–123.

[24] V. Muppavarapu, A. Rajendran, and S. K. Vasudevan, ''Phishing detection using RDF and random forests,'' Int. Arab J. Inf. Technol., vol. 15, no. 5, pp. 817–824, 2018.

[25] V. K. Nadar, B. Patel, V. Devmane, and U. Bhave, ''Detection of phishing websites using machine learning approach,'' in Proc. 2nd Global Conf. Advancement Technol. (GCAT). Rajasthan, Jaipur, India: Amity University, Oct. 2021, pp. 1–8.

[26] J. Mao, W. Tian, P. Li, T. Wei, and Z. Liang, ''Phishing-alarm: Robust and efficient phishing detection via page component similarity,'' IEEE Access, vol. 5, pp. 17020–17030, 2017.

[27] N. C. R. L. Y. Teraguchi and J. C. Mitchell, ''Client-side defense against web-based identity theft,'' Dept. Comput. Sci., Stanford Univ., Stanford, CA, USA, 2004. [Online]. Available: http://crypto.stanford. edu/SpoofGuard/webspoof.pdf

[28] 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.

[29] A. Sharma and D. Upadhyay, ''VDBSCAN clustering with map-reduce technique,'' in Recent Findings in Intelligent Computing Techniques. Singapore: Springer, 2018, pp. 305–314.

[30] A. K. Jain and B. B. Gupta, ''Comparative analysis of features based machine learning approaches for phishing detection,'' in Proc. 3rd Int. Conf. Comput. Sustain. Global Develop. (INDIACom), Mar. 2016, pp. 2125–2130.

[31] P. Rao, J. Gyani, and G. Narsimha, ‘‘Fake profiles identification in online social networks using machine learning and NLP,’’ Int. J. Appl. Eng. Res., vol. 13, no. 6, pp. 973–4562, 2018.

[32] G. Xiang, J. Hong, C. P. Rose, and L. Cranor, ‘‘CANTINA+: A featurerich machine learning framework for detecting phishing web sites,’’ ACM Trans. Inf. Syst. Secur., vol. 14, no. 2, pp. 1–28, Sep. 2011.

[33] V. S. Lakshmi and M. S. Vijaya, ‘‘Efficient prediction of phishing websites using supervised learning algorithms,’’ Proc. Eng., vol. 30, pp. 798–805, 2012.

[34] D. Sahoo, C. Liu, and S. C. H. Hoi, ‘‘Malicious URL detection using machine learning: A survey,’’ 2017, arXiv:1701.07179.

[35] E. Kremic and A. Subasi, ‘‘Performance of random forest and SVM in face recognition,’’ Int. Arab J. Inf. Technol., vol. 13, no. 2, pp. 287–293, 2016.

[36] K. Yu, L. Tan, S. Mumtaz, S. Al-Rubaye, A. Al-Dulaimi, A. K. Bashir, and F. A. Khan, ‘‘Securing critical infrastructures: Deep-learning-based threat detection in IIoT,’’ IEEE Commun. Mag., vol. 59, no. 10, pp. 76–82, Oct. 2021.

[37] P. Chen, L. Desmet, and C. Huygens, ‘‘A study on advanced persistent threats,’’ in Communications and Multimedia Security. Aveiro, Portugal: Springer, Sep. 2014, pp. 63–72.

[38] E. Sisinni, A. Saifullah, S. Han, U. Jennehag, and M. Gidlund, ‘‘Industrial Internet of Things: Challenges, opportunities, and directions,’’ IEEE Trans. Ind. Informat., vol. 14, no. 11, pp. 4724–4734, Nov. 2018.

[39] S. Alaparthi and M. Mishra, ‘‘Bidirectional encoder representations from transformers (BERT): A sentiment analysis Odyssey,’’ 2020, arXiv:2007.01127.

[40] P. A. Barraclough, M. A. Hossain, M. A. Tahir, G. Sexton, and N. Aslam, ‘‘Intelligent phishing detection and protection scheme for online transactions,’’ Exp. Syst. Appl., vol. 40, no. 11, pp. 4697–4706, Sep. 2013.

[41] S. Van Acker, D. Hausknecht, and A. Sabelfeld, ‘‘Measuring login webpage security,’’ in Proc. Symp. Appl. Comput., Apr. 2017, pp. 1753–1760.

[42] J. Ma, L. K. Saul, S. Savage, and G. M. Voelker, ‘‘Identifying suspicious URLs: An application of large-scale online learning,’’ in Proc. 26th Annu. Int. Conf. Mach. Learn., Jun. 2009,

pp. 681–688.

[43] I. Fette, N. Sadeh, and A. Tomasic, ‘‘Learning to detect phishing emails,’’ in Proc. 16th Int. Conf. World Wide Web, May 2007, pp. 649–656.

[44] M. G. Alkhozae and O. A. Batarfi, ‘‘Phishing websites detection based on phishing characteristics in the webpage source code,’’ Int. J. Inf. Commun. Technol. Res., vol. 1, no. 6, pp. 1–9, 2011.

[45] 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 Proc. SIGCHI Conf. Human Factors Comput. Syst., Apr. 2007, pp. 905–914.

[46] Y. Cao, W. Han, and Y. Le, ‘‘Anti-phishing based on automated individual white-list,’’ in Proc. 4th ACM workshop Digit. identity Manage., Oct. 2008, pp. 51–60.

[47] C. Whittaker, B. Ryner, and M. Nazif, ‘‘Large-scale automatic classification of phishing pages,’’ in Proc. Netw. Distrib. Syst. Secur. Symp. (NDSS), San Diego, CA, USA, Feb./Mar. 2010.

[48] M. Zouina and B. Outtaj, ‘‘A novel lightweight URL phishing detection system using SVM and the similarity index,’’ Human-Centric Comput. Inf. Sci., vol. 7, no. 1, p. 17, Dec. 2017.

[49] J. Ma, L. K. Saul, S. Savage, and G. M. Voelker, ‘‘Beyond blacklists: Learning to detect malicious web sites from suspicious URLs,’’ in Proc. 15th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, Jun. 2009, pp. 1245–1254.

[50] M. Khonji, Y. Iraqi, and A. Jones, ‘‘Lexical URL analysis for discriminating phishing and legitimate websites,’’ in Proc. 8th Annu. Collaboration, Electron. Messaging, Anti-Abuse Spam Conf., Sep. 2011, pp. 109–115.

[51] M. Khonji and Y. Iraqi, ‘‘Enhancing phishing e-mail classifiers: A lexical URL analysis approach,’’ Int. J. Inf. Secur. Res., vol. 2, nos. 1–2, p. 40, 2012.

[52] V. P. Reddy, V. Radha, and M. Jindal, ‘‘Client-side protection from phishing attack,’’ Int. J. Adv. Eng. Sci. Technol., vol. 3, no. 1, pp. 39–45, 2011.

[53] Z. Li, K. Zhang, Y. Xie, F. Yu, and X. Wang, ‘‘Knowing your enemy: Understanding and detecting malicious web advertising,’’ in Proc. ACM Conf. Comput. Commun. Secur., Oct. 2012, pp. 674–686.

[54] Y. Mansour, S. Muthukrishnan, and N. Nisan, ‘‘Doubleclick AD exchange auction,’’ 2012,

arXiv:1204.0535.

[55] S. Bell and P. Komisarczuk, ‘‘An analysis of phishing blacklists: Google safe browsing, OpenPhish, and PhishTank,’’ in Proc. Australas. Comput. Sci. Week Multiconference, Feb. 2020, pp. 1–11.

[56] D. Canali, M. Cova, G. Vigna, and C. Kruegel, ‘‘Prophiler: A fast filter for the large-scale detection of malicious web pages,’’ in Proc. 20th Int. Conf. World wide web, Mar. 2011, pp. 197–206.

[57] S. Ford. Wepawet. (2009). [Online]. Available: http://wepawet.cs. ucsb.edu/index.php

[58] M. Imran, M. H. Durad, F. A. Khan, and H. Abbas, ‘‘DAISY: A detection and mitigation system against denial-of-service attacks in software-defined networks,’’ IEEE Syst. J., vol. 14, no. 2, pp. 1933–1944, Jun. 2020.

[59] Q. Yan, F. R. Yu, Q. Gong, and J. Li, ‘‘Software-defined networking (SDN) and distributed denial of service (DDoS) attacks in cloud computing environments: A survey, some research issues, and challenges,’’ IEEE Commun. Surveys Tuts., vol. 18, no. 1, pp. 602–622, 1st Quart., 2016.

[60] A. Subasi, E. Molah, F. Almkallawi, and T. J. Chaudhery, ‘‘Intelligent phishing website detection using random forest classifier,’’ in Proc. Int. Conf. Electr. Comput. Technol. Appl. (ICECTA), Nov. 2017, pp. 1–5.

[61] M. Bugliesi, S. Calzavara, R. Focardi, and W. Khan, ‘‘CookiExt: Patching the browser against session hijacking attacks,’’ J. Comput. Secur., vol. 23, no. 4, pp. 509–537, Sep. 2015.

[62] R. M. Mohammad, F. Thabtah, and L. McCluskey, ‘‘Phishing websites features,’’ School Comput. Eng., Univ. Huddersfield, West Yorkshire, U.K., Tech. Rep., 2015. [Online]. Available: http://eprints.hud.ac.uk/id/ eprint/24330/6/MohammadPhishing14July2015.pdf

[63] D. Dua and C. Graff. (2017). UCI Machine Learning Repository. [Online]. Available: http://archive.ics.uci.edu/ml

[64] M. Jalalian and M. Dadkhah, ‘‘The full story of 90 hijacked journals from August 2011 to June 2015,’’ Geographica Pannonica, vol. 19, no. 2, pp. 73–87, 2015.

[65] F. A. Khan and A. Gumaei, ‘‘A comparative study of machine learning classifiers for network intrusion detection,’’ in Artificial Intelligence and Security. New York, NY, USA: Springer, Jun. 2019, pp. 75–86.

[66] N. Moustafa and J. Slay, ‘‘The 16 significant features of the UNSW-NB15 and the KDD99 data sets for network intrusion detection systems,’’ in Proc. 4th Int. Workshop Building Anal. Datasets Gathering Exper. Returns Secur. (BADGERS), Nov. 2015, pp. 25–31.

[67] Y. T. Ho, C. Wu, M. Yang, T. Chen, and Y. Chang, ‘‘Replanting 32 your forest: NVM-friendly bagging strategy for random forest,’’ in Proc. IEEE Non-Volatile Memory Syst. Appl. Symp. (NVMSA), Aug. 2019, pp. 1–6.

[68] G. Sonowal and K. S. Kuppusamy, ‘‘PhiDMA—A 24 phishing detection model with multi-filter approach,’’ J. King Saud Univ. Comput. Inf. Sci., vol. 32, no. 1, pp. 99–112, Jan. 2020.

[69] Y. Zhang, S. Egelman, L. Cranor, and J. Hong, ‘‘Phinding phish: Evaluating anti-phishing tools,’’ Carnegie Mellon Univ., 2018, doi: 10.1184/R1/6470321.v1.

[70] A. K. Jain and B. B. Gupta, ‘‘A 1 machine learning based approach for phishing detection using hyperlinks information,’’ J. Ambient Intell. Humanized Comput., vol. 10, no. 5, pp. 2015–2028, May 2019.

Sources

1	https://www.nature.com/articles/s41598-022-10841-5 INTERNET	4%
2	www.ncbi.nlm.nih.gov/pmc/articles/PMC7581503/ INTERNET	2%
3	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9963556 INTERNET	2%
4	bing.com/videos INTERNET	2%
5	https://link.springer.com/chapter/10.1007/978-3-030-33432-1_12 INTERNET	1%
6	ieeexplore.ieee.org/document/7289347/ INTERNET	1%
7	link.springer.com/chapter/10.1007/978-3-030-00557-3_46 INTERNET	1%
8	ieeexplore.ieee.org/document/10155142/ INTERNET	1%
9	www.researchgate.net/profile/Zawar-H-Khan-P-Eng/public... INTERNET	<1%
10	https://dblp.org/pid/237/6328 INTERNET	<1%
11	https://books.google.com/books/about/Detection_of_Intrusions_an... INTERNET	<1%
12	www.zoho.com/workplace/articles/navigating-phishing-emails.html INTERNET	<1%
13	https://scholar.google.com/citations?user=hje68Y4AAAAJ INTERNET	<1%
14	www.sciencedirect.com/science/article/pii/S187705092030... INTERNET	<1%

15	https://jis-eurasipjournals.springeropen.com/articles/10.1186/s13635... INTERNET	<1%
16	ieeexplore.ieee.org/document/7809531 INTERNET	<1%
17	https://scholar.google.com/citations?user=FIpwvjsAAAAJ INTERNET	<1%
18	https://dl.acm.org/doi/10.1145/2699026.2699115 INTERNET	<1%
19	https://www.semanticscholar.org/paper/DAISY:-A-Detection-and-Mitig... INTERNET	<1%
20	https://link.springer.com/chapter/10.1007/978-3-030-24265-7_7 INTERNET	<1%
21	https://digital-skills-jobs.europa.eu/en/organisations/national... INTERNET	<1%
22	https://www.semanticscholar.org/paper/Efficient-prediction-of-phis... INTERNET	<1%
23	https://dblp.org/rec/journals/hcis/ZouinaO17 INTERNET	<1%
24	https://dblp.org/pid/225/4271 INTERNET	<1%
25	https://www.researchgate.net/publication/279232562_The_full_story_of... INTERNET	<1%
26	https://arxiv.org/pdf/2105.07769.pdf INTERNET	<1%
27	https://www.semanticscholar.org/paper/Protect-sensitive-sites-from... INTERNET	<1%
28	https://dl.acm.org/doi/10.1109/COMST.2016.2519912 INTERNET	<1%
29	https://dblp.org/rec/journals/iajit/MuppavarapuRV18 INTERNET	<1%

30	https://www.researchgate.net/publication/32... INTERNET <1%
31	https://www.ripublication.com/ijaer18/ijaerv13n6_133.pdf INTERNET <1%
32	https://scholar.google.com/citations?user=xj7c79YAAAAJ INTERNET <1%
33	www.sciencedirect.com/science/article/pii/S1877050915013411 INTERNET <1%
34	https://link.springer.com/chapter/10.1007/978-3-642-02490-0_66 INTERNET <1%
35	https://ieeexplore.ieee.org/abstract/document/7724641 INTERNET <1%
36	https://dl.acm.org/doi/abs/10.1145/2695664.2695709 INTERNET <1%
37	link.springer.com/chapter/10.1007/978-3-662-44885-4_5 INTERNET <1%
38	https://arxiv.org/abs/2007.01127 INTERNET <1%
39	https://dl.acm.org/doi/10.5555/3489212.3489234 INTERNET <1%
40	https://ieeexplore.ieee.org/document/10155142 INTERNET <1%
41	www.sciencedirect.com/science/article/pii/S0957417413001255 INTERNET <1%
42	https://ieeexplore.ieee.org/document/6148476 INTERNET <1%
43	https://www.researchgate.net/publication/221521469_Effective... INTERNET <1%
44	https://ieeexplore.ieee.org/document/9391742 INTERNET <1%

45	https://dl.acm.org/doi/10.1007/11663812_7 INTERNET <1%
46	https://ieeexplore.ieee.org/document/1607989 INTERNET <1%
47	https://scholar.google.com/citations?user=llys3iMAAAAJ INTERNET <1%
48	https://www.researchgate.net/publication/371540338_Exploring_the... INTERNET <1%

EXCLUDE CUSTOM MATCHES	OFF
EXCLUDE QUOTES	OFF
EXCLUDE BIBLIOGRAPHY	OFF