Table Comparison for references

Nama: Bryan Orville Audric

NIM: 2602160750

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Author** | **Year** | **Title** | **Objective** | **Method** | **Data / Sample** | **Conclusion** | **Weakness** | **Strength** |
| Zainab Alkhalil, Chaminda Hewage, Liqaa Nawaf and Imtiaz Khan | 2021 | Phishing Attacks: A Recent Comprehensive Study and a New Anatomy | To understand evolving phishing attacks, identify motivations and techniques, propose mitigation strategies, explore impacts beyond financial crimes, emphasize security awareness training, address emerging threats, advocate for specific laws, and highlight the role of law enforcement in combating phishing. | Literature Review | Several journals from IEEE, World Wide Web Conference and Security User Risk Report | The research concludes that these threats have expanded beyond traditional data theft to include cyber terrorism, hacktivism, and nation-state attacks, necessitating multifaceted countermeasures due to the diverse nature of the attack vector. While human education is crucial in defense against phishing, complete eradication is challenging due to the sophistication of attacks. Continuous security awareness training is vital, but the development of effective anti-phishing techniques to prevent user exposure is essential. Further research is recommended to study user susceptibility, address emerging threats like social media-based phishing, and enhance law enforcement mechanisms to combat phishing effectively. | Lack of in-depth analysis on the effectiveness of specific anti-phishing techniques proposed | Comprehensive examination of various aspects of the phenomenon, including the evolution of phishing beyond traditional data theft, the identification of diverse attack vectors, and the recommendation of multifaceted countermeasures |
| ÖMER ASLAN, REFIK SAMET | 2019 | A Comprehensive Review on Malware Detection  Approaches | To offer a comprehensive review of current malware detection approaches, summarizing existing studies, explaining key methods, discussing challenges, proposing new assumptions, and providing a systematic overview for further research. | Literature Review | A wide range of academic papers, research articles, conference proceedings, and technical reports | The research on malware detection approaches concludes that while signature- and heuristic-based methods are effective for known malware, behavior-, model checking-, and cloud-based approaches perform better for unknown and complex malware. Emerging techniques like deep learning, mobile devices, and IoT-based approaches show promise but none can detect all types of malwares. The study underscores the need for further research and the combination of different detection approaches to enhance capabilities and address the evolving landscape of malware threats.for the ongoing arms race between hackers and security experts. | The absence of empirical validation or experimental evidence to substantiate the claims made in the literature review. Without original experiments or case studies, the study may lack concrete proof of the effectiveness of the discussed malware detection approaches. | A detailed overview of various malware detection methods and technologies |
| Vikky Aprelia Windarni, Anggit Ferdita Nugraha, Surya Tri Atmaja Ramadhani,Dewi Anisa Istiqomah, Fiyas Mahananing Puri, Adi Setiawan | 2023 | DETEKSI WEBSITE PHISHING MENGGUNAKAN TEKNIK FILTER PADA  MODEL MACHINE LEARNING | To detect phishing websites using machine learning with filter features and to determine the most effective machine learning models with filter method for detecting | Data Analysis | “Phishing Websites," UCI Machine Learning | The research found that the Random Forest method had the highest accuracy rate of 96.3% in detecting phishing websites. The conclusion was that the application of the Random Forest method is effective in detecting phishing websites, and future research could explore different algorithms to improve results. | The lack of comparison with other filter method | The uses of 3 different machine learning models for evaluation |
| Alyssa Anne Ubing, Syukrina Kamilia Binti Jasmi, Azween Abdullah, NZ Jhanjhi, Mahadevan Supramaniam | 2019 | Phishing Website Detection: An Improved Accuracy through Feature Selection and Ensemble | To improve the accuracy of phishing website detection using an ensemble learning approach with feature selection | Data Analysis | UCI Machine Learning Repository | The research concludes that the ensemble learning model implemented for phishing detection alongside feature selection methods exhibits high accuracy rates, reaching up to 95%, which outperforms existing technologies. By leveraging ensemble methods and combining multiple classification models, the study demonstrates a significant enhancement in detecting phishing websites. This approach showcases the potential of ensemble learning in improving cybersecurity measures and addressing online threats effectively, offering a promising solution for combating cybercrime. | Limited dataset with relatively small amount of URL used | By utilizing feature selection algorithms and ensemble methods, the study demonstrates a superior performance of the proposed model, achieving up to 95% accuracy |
| Abdelhakim Hannousse, Salima Yahiouche | 2020 | TOWARDS BENCHMARK DATASETS FOR MACHINE LEARNING  BASED WEBSITE PHISHING DETECTION: AN EXPERIMENTAL  STUDY | To propose a scheme for creating benchmark datasets that can be used for machine learning-based website phishing detection. | Data Analysis | Legitimate website from Alexa and Yandex, Phishing website from Phishtank and Openphish | This study proposes guidelines for building reproducible datasets for website phishing detection to serve as benchmarks for machine learning systems. Results show Random Forest classifiers predict phishing web pages with higher accuracy than other methods, particularly with external service-based features, despite potential network delays. Content-based and certain hyperlink features are less effective. Filter methods enhance model performance by removing unimportant features, and hybrid feature models consistently outperform single-class models. Future work includes validating findings with more benchmark datasets and exploring deep learning for phishing detection. | The further evolving tactics of phishing may invalidate this research’s finding | The use of different machine learning models, set of features from different source and different features category further supports the research’s findings |
| Arun D. Kulkarni, Leonard L. Brown, III | 2019 | Phishing websites detection using machine learning | To utilize machine learning techniques to detect phishing websites based on their URLs by comparing the performance of four classifiers (decision tree, Naïve Bayesian classifier, support vector machine, and neural network) using a dataset of 1,353 URLs categorized as legitimate, suspicious, or phishing | Data Analysis | University of California, Irvine Machine Learning Repository,  Center for Machine Learning and Intelligent Systems | The research paper successfully implemented various classifiers, such as decision trees, Naïve Bayes, Support Vector Machine, and Neural Network, to detect phishing URLs. Despite the neural network's lower accuracy due to discrete feature values, potential enhancements through additional hidden layers or units were suggested | The number of datasets is relatively small; hence the evaluation accuracy might be the result of overfitting | The use of different classifier models to get an accurate prediction |
| Muhammad Waqas Shaukat, Rashid Amin, Muhana Magboul Ali Muslam, Asma Hassan Alshehri and Jiang Xie | 2023 | A Hybrid Approach for Alluring Ads Phishing Attack Detection Using Machine Learning | To develop a more efficient solution to detect and mitigate phishing websites, given that current approaches struggle to keep up | Data Analysis | PhishTank and websites of daily use  that are reported by Alexa | The proposed layered classification model achieved high accuracy in detecting phishing websites. The model uses machine learning algorithms to analyze URL structure, text, and image features to identify phishing websites that mimic legitimate ones. This layered approach can help to mitigate the risks of phishing websites bypassing detection filters. | The number of datasets is relatively small; hence the evaluation accuracy might be the result of overfitting | The practical approach by scrapping and preprocessing not just URL but also potential Ad image with text to verify the accuracy of the proposed layered classification model |
| Ozgur Koray Sahingoz, Ebubekir Buber, Onder Demir, Banu Diri | 2019 | Machine learning based phishing detection from URLs | To develop machine learning-based systems for phishing detection by focusing on dataset construction, data processing, and utilizing various machine learning techniques to improve the accuracy and efficiency of anti-phishing systems. | Data Analysis | PhishTank, Yandex, Ebbu2017 Phishing Datase | The researcher developed a phishing detection system using seven machine learning algorithms along with NLP-based features, word vectors, and hybrid features. Their dataset of 73,575 URLs (36,400 legitimate and 37,175 phishing) is shared for research use. NLP-based features performed better than word vectors and combining both improved detection rates. They suggest using deep learning for greater efficiency despite longer training times. | Limited dataset size used for testing | The use of different source of dataset, machine learning algorithm and different category of features |
| Vahid Shahrivari, Mohammad Mahdi Darabi, Mohammad Izadi | 2020 | Phishing Detection Using Machine Learning Techniques | To compare the results of multiple machine learning methods such as Logistic Regression, Decision Tree, Random Forest, Ada-Boost, Support Vector Machine, KNN, Neural Networks, Gradient Boosting, and XGBoost for predicting phishing websites | Data Analysis | PhishTank | The conclusion of the research is that the study implemented and evaluated twelve classifiers on a dataset of legitimate and phishing websites. The results showed that ensemble classifiers, specifically Random Forest and XGBoost, performed very well in terms of computation duration and accuracy. However, it was noted that there is no guarantee that the combination of multiple classifiers will always perform better than the best individual classifier. | The scarcity of a standard feature set and the lack of a large, diverse dataset may have impacted the performance of the models | The use of 12 different machine learning methods further supports the validity of the comparison |
| Ashit Kumar Dutta | 2021 | Detecting phishing websites using machine  learning technique | To propose a framework using RNN-LSTM for classifying URLs as malicious or legitimate, with the goal of improving the accuracy of phishing URL detection. | Data Analysis | AlexaRank dataset for legitimate websites and the Phishtank dataset for phishing websites | The study demonstrates the effectiveness of the RNN-LSTM in detecting malicious URLs. The results indicate that the RNN-LSTM approach outperforms existing methods in identifying phishing websites. The study emphasizes the significance of machine learning techniques in combating phishing attacks and suggests further research to enhance detection accuracy and safeguard individual privacy. | The number of datasets is relatively small; hence the evaluation accuracy might be the result of overfitting | The use of different hyperparameter to get an accurate prediction |