International Journal of Engineering Research and Science & Technology

IJERST



ISSN : 2319-5991

www.ijerst.com

Email: editor@ijerst.com or editor.ijerst@gmail.com

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com

Vol. 21, Issue 2, 2025

REAL TIME SPEAR PHISHING MAIL DETECTION USING AI & ML

Chimmili Suvarnakumari, **21PD1A0517**, Department of CSE, West Godavari Institute Of Science and Engineering, Andhra Pradesh, India.

Taneti Sumitha, 21PD1A0582, Department of CSE, West Godavari Institute Of Science and Engineering, Andhra Pradesh, India.

Bolla Jnanendra Kumar, 21PD1A0512, Department of CSE, West Godavari Institute Of Science and Engineering, Andhra Pradesh, India.

Vantaku Hemanth Kumar, 22PD5A0527, Department of CSE, West Godavari Institute Of Science and Engineering, Andhra Pradesh, India.

Prathipadu Muralikrishna, 20PC1A0517, Department of CSE, West Godavari Institute Of Science and Engineering, Andhra Pradesh, India

Dr. M.Aravind Kumar, Professor, Department of ECE, West Godavari Institute Of Science and Engineering, Andhra Pradesh, India.

Dr. M.Aravind Kumar, Professor, Department of ECE, West Godavari Institute Of Science and Engineering, Andhra Pradesh, India

ABSTRACT

Spear phishing continues to pose a serious cybersecurity risk, leveraging highly tailored and deceptive emails to exploit individuals and organizations. Traditional email security systems often fall short in identifying these advanced threats, emphasizing the need for intelligent, AI-powered solutions. This study proposes a real-time detection framework combining Random Forest (RF), Long Short-Term Memory (LSTM), and Support Vector Machines (SVM). RF aids in feature selection and classification, LSTM captures the sequential flow of email content, and SVM identifies subtle anomalies in text. By analyzing metadata, linguistic features, and embedded links, the model effectively detects malicious emails with high accuracy and reduced false positives. While the approach significantly strengthens email security, it also acknowledges ongoing challenges like adaptive phishing tactics and adversarial threats. Overall, this research highlights the promise of AI in combating spear phishing and lays the groundwork for next-generation email security solutions.

INTRODUCTION

Spear phishing is a highly targeted cyberattack that exploits human trust by using carefully crafted emails designed to impersonate legitimate sources, making it difficult for traditional email filters to detect. To address this challenge, Artificial Intelligence (AI) and Machine Learning (ML) techniques such as Random Forest (RF), Long Short-Term Memory (LSTM), and Support Vector Machines (SVM) are employed for real-time detection. RF aids in effective feature selection and classification, LSTM captures sequential patterns in email content, and SVM identifies subtle textual anomalies that signal phishing attempts. By analyzing metadata, linguistic cues, and behavioral indicators, these models significantly enhance detection accuracy, reduce false positives, and strengthen overall cybersecurity resilience.

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 2, 2025

LITERATURE SURVEY

Spear phishing is a sophisticated cybersecurity threat that uses social engineering to deceive individuals into disclosing sensitive information. Traditional methods like signature-based and heuristic detection often fail against these highly personalized attacks. This survey reviews research on spear phishing detection, highlighting the increasing use of Artificial Intelligence (AI) and Machine Learning (ML) to address these challenges. Early systems, such as Fette et al.'s "Phishing Email Classifier" (2007), relied on signature-based methods, identifying phishing attempts through known patterns and suspicious features. However, these approaches struggle with zero-day attacks and evolving tactics, emphasizing the need for more adaptive and intelligent solutions.

EXISTING SYSTEM

Various phishing email detection methods leverage Natural Language Processing (NLP) and Machine Learning (ML) techniques, focusing on feature extraction from email attributes like URLs, sender information, and text content. Support Vector Machines (SVM) are often used for classification due to their ability to handle high-dimensional data, while models like Decision Trees, Random Forest, and Neural Networks are also explored. NLP techniques such as keyword and sentiment analysis identify deceptive language, and deep learning models like Long Short-Term Memory (LSTM) networks improve detection by analyzing text sequences. Together, these methods automate classification and enhance phishing threat identification.

PROPOSED SYSTEM

The proposed spear phishing detection system integrates Random Forest (RF), Long Short-Term Memory (LSTM), and Support Vector Machines (SVM) to enhance real-time email threat detection. Emails are preprocessed to extract features from metadata, subject lines, and content, with RF used for feature selection, prioritizing attributes like sender authenticity and suspicious keywords. LSTM analyzes email text to detect contextual anomalies, while SVM classifies emails as legitimate or malicious based on these features. This combined approach improves classification accuracy, reduces false positives, and adapts to new threats through continuous updates, offering a robust AI-driven defense against spear phishing and strengthening email security.

HARDWARE REQUIREMENTS

1. Processor : Intel i5
2. Speed : 1:1 GHz (min)
3. Ram : 8 GB RAM
4. Hard Disk : 512 GB

SOFTWARE REQUIREMENTS

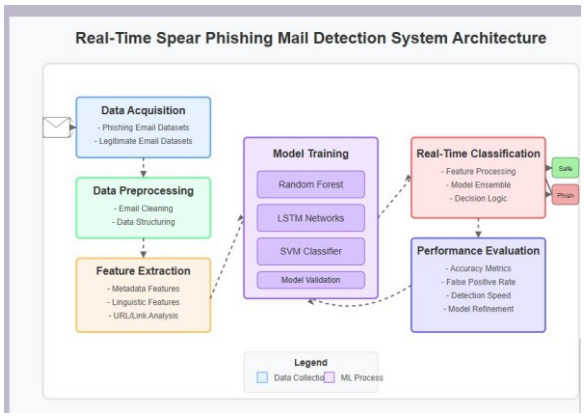
1. Operating System :Windows Family
2. Version Python : 3.9.1
3. Programming language: Python
4. Development IDE : Jupyter Notebook

SYSTEM ARCHITECTURE

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com

Vol. 21, Issue 2, 2025



A real-time spear phishing email detection system using AI and ML can be implemented by integrating Natural Language Processing (NLP) and machine learning algorithms into an email filtering pipeline. The system continuously monitors incoming emails and uses NLP to extract and analyze contextual features such as sender authenticity, subject relevance, body content, attachment types, and URLs. These features are fed into a trained ML model—such as Random Forest, XGBoost, or a deep learning-based classifier like LSTM or BERT—which has been trained on a labeled dataset of phishing and legitimate emails. The model predicts whether the email is a spear phishing attempt in real time, and if flagged, the system can alert the user or quarantine the message to prevent user interaction, ensuring proactive protection against targeted phishing attacks.

CONCLUSION

The proposed methodology presents a real-time spear phishing detection system utilizing Random Forest, LSTM, and SVM to enhance detection accuracy. The multi-stage process, spanning feature extraction, classification, and evaluation, ensures strong protection against phishing attacks. This hybrid AI approach boosts security with adaptive learning and real-time analysis capabilities. Future research will aim to enhance resilience against adversarial attacks and improve computational efficiency for large-scale email security implementation

REFERENCES

1. J. Lee, Y. Lee, D. Lee, H. Kwon, and D. Shin, "Spear Phishing Using Machine Learning," in Proceedings of the International Conference on Advances in Computer Science and Engineering (ICACSE), 2023, pp. 531-540. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-031-37940-6_43
2. X. Zhang and Y. Yuan, "Spear-Phishing Detection Method Based on Few- Shot Learning," in Advances in Cyber Security: Principles, Techniques, and Applications, 2023, pp. 245-256. [Online]. Available: https://link.springer.com/chapter/10.1007/978-981-99-7872-4_20
3. I. Fayaz and Z. A. Nagar, "Email Phishing Detection Using Machine Learning," International Journal of Emerging Technologies and Innovative Research, vol. 10, no. 3, pp. 316-322, 2023. [Online]. Available: https://www.jetir.org/papers/JETIR2410316.pdf
4. S. Sabale, G. Nadar, S. Naikawadi, S. Patil, and R. Patel, "Phishing Detection using Machine Learning Technique: Random Forest,"

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com

Vol. 21, Issue 2, 2025

International Journal of Research Publication Reviews, vol. 3, no. 2, pp.12241226,2022.[Online].Available: https://ijrpr.com/uploads/V3ISSUE2/ijrpr2782-phishing-detection-using-machine-learning-technique.pdf

5. M. Gopalsamy, "Identification and Classification of Phishing Emails Based on Machine Learning Techniques to Improve Cybersecurity," International Journal of Cyber-Security and Digital Forensics, vol. 12, no. 1,pp.45-56,2023.[Online].Available: https://www.researchgate.net/publication/387523133_Identification_And_Classification_Of_Phishing_Emails_Based_on_Machine_Learning_Techniques_To_Improvise_Cyber_security