International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 05 Issue: 03 | March – 2026 DOI:10.55041/ISJEM05952
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

AI-Powered Phishing Detection and Mid Stalker Analyzer

K. Muralidhar [1], Jeedimadla Pragna [2], Gopal Vishvesh [3], Utnoor Roopika [4]

Assistant Professor[1], Department of Computer Engineering,
Methodist College of Engineering and Technology, Abids, Hyderabad,
Telangana, 500001, India.

Students[234], Department of Computer Engineering,
Methodist College of Engineering and Technology, Abids, Hyderabad, Telangana, 500001, India.

ABSTRACT - With the increasing dependence on digital platforms for communication, financial transactions, and social interactions, cybersecurity threats have grown significantly in recent years. Among these threats, phishing attacks and online stalking have emerged as major concerns due to their ability to exploit user behavior and compromise sensitive information. Phishing attacks deceive users into revealing confidential data through fake websites, emails, or messages, while stalker activities involve monitoring and analyzing user interactions for malicious purposes.

Traditional cybersecurity solutions rely on static rule-based mechanisms, which are often ineffective against modern, evolving attack techniques. These systems fail to adapt to dynamic patterns and lack real-time detection capabilities, resulting in increased vulnerability for users.

This research proposes an intelligent system that integrates phishing detection and mid-stalker analysis using artificial intelligence techniques**.** The system utilizes machine learning algorithms, natural language processing, and behavioral pattern analysis to identify malicious URLs, detect suspicious email content, and monitor abnormal user activity. By combining content-based and behavior-based detection, the system provides a comprehensive security solution.

INDEX TERMS: Phishing Detection, Cybersecurity, Machine Learning, NLP, Behavioral Analysis, Anomaly Detection, AI Security

INTRODUCTION

The rapid growth of internet technologies has transformed the way people communicate, work, and access services. Online platforms have become essential for daily activities such as banking, shopping, education, and social networking. However, this increased reliance on digital systems has also led to a rise in cyber threats.

Phishing attacks are one of the most common forms of cybercrime, where attackers impersonate trusted entities to trick users into sharing sensitive information. These attacks are often carried out through emails, fake websites, or malicious links. At the same time, online stalking has become a serious issue, where attackers monitor user behavior, track online activities, and collect personal data without consent. These threats can result in financial loss, identity theft, and serious privacy violations.

The Need / Motivation :

Existing security systems are not capable of handling advanced cyber threats effectively. Most systems rely on predefined rules and signatures, which makes them ineffective against new and unknown attacks. Additionally, users often lack awareness and fail to identify phishing attempts or suspicious activities.

There is a need for an intelligent and automated system that can:

- Detect phishing attacks in real time
- Analyze user behavior patterns
- Identify suspicious activities

Existing Work:
Several studies have focused on detecting phishing attacks using machine learning models such as Decision Trees, Random Forest, and Support Vector Machines. These systems mainly analyze URL features, domain information, and webpage content to classify threats. In addition, Natural Language Processing techniques are used to examine email text and identify suspicious patterns. Some research also explores behavioral monitoring to detect unusual user activities. However, most existing solutions address phishing or behavior analysis separately rather than combining both.

Drawbacks:
Existing systems often rely on predefined rules or historical data, making them less effective against new and evolving phishing attacks. Many models struggle with high false positives and may incorrectly classify legitimate content as malicious. Another limitation is the lack of integration between phishing detection and user behavior analysis. Privacy concerns also arise when tracking user activity for anomaly detection. Additionally, complex models may require high computational resources, limiting real-time performance.

Proposed Approach:
The proposed system presents an AI-driven solution that integrates phishing detection and mid-stalker analysis into a single framework. It uses machine learning algorithms to examine various features such as URL structure, domain details, and email content to identify potential phishing attacks. By applying data preprocessing and feature extraction techniques, the system improves the accuracy of classification. This approach enables the model to detect both known and previously unseen threats more effectively.

In addition to phishing detection, the system incorporates a behavioral analysis module that monitors user activities to identify unusual patterns. It tracks interactions such as login behavior, browsing habits, and access frequency to detect possible stalking behavior. Natural Language Processing techniques are also used to analyze textual data and improve detection of suspicious messages. The system generates real-time alerts to notify users, ensuring quick response and enhanced security.

I. NEED OF THE STUDY

In today's digital environment, cybersecurity has become a critical issue due to the increasing number of cyberattacks. Phishing and stalking activities are growing rapidly, targeting individuals and organizations alike. Traditional detection systems are limited in their ability to detect these threats effectively.

One major challenge is that phishing attacks continuously evolve, making it difficult for rule-based systems to identify them. Similarly, stalking behavior often goes unnoticed because existing systems do not analyze user interactions over time.

Additionally, many users are unaware of cybersecurity risks, making them more vulnerable to attacks. There is also a lack of systems that can provide continuous monitoring and real-time detection.

To address these challenges, an AI-based system is required that can analyze both content and behavior. The proposed system fulfills this need by providing an automated and intelligent solution for detecting phishing and stalking activities.

II. FUNDAMENTALS OF CYBERSECURITY AND AI

Basics of Phishing and Stalker Behavior

Phishing is a cyberattack technique used to obtain sensitive information by impersonating legitimate sources. Attackers create fake emails, websites, or messages that appear authentic, tricking users into revealing confidential data.

Stalker behavior refers to the continuous monitoring of user activities. This may include tracking browsing patterns, analyzing interactions, or collecting personal information. Such behavior can lead to privacy breaches and security risks.

Basics of Artificial Intelligence

Artificial Intelligence (AI) enables machines to perform tasks that require human intelligence. Machine learning, a subset of AI, allows systems to learn from data and improve their performance over time.

In cybersecurity, AI is used for:

- Detecting anomalies
- Classifying malicious content

III.LITERATURE REVIEW

S.No	Title (Year)	Author(s)	Methodology	Outcome	Limitations
1	Machine Learning Based Phishing Detection (2024)	R. Sharma, P. Verma	Implemented Random Forest and Support Vector Machine to classify phishing URLs using feature extraction techniques	Achieved high accuracy in detecting known phishing websites	Struggled to detect newly emerging and zero-day phishing attacks
2	Email Phishing Detection using NLP (2023)	S. Reddy, K. Naidu	Applied Natural Language Processing techniques to analyze email content and identify suspicious patterns	Improved detection of phishing emails based on textual analysis	Required large training datasets and high processing time
3	Deep Learning for URL Classification (2022)	A. Gupta, M. Singh	Used deep neural networks to classify URLs as legitimate or malicious based on learned patterns	Enhanced detection accuracy compared to traditional methods	High computational cost and complex model training
4	Behavior-Based Stalker Detection System (2023)	V. Kumar, S. Iyer	Monitored user activity patterns and applied anomaly detection to identify suspicious behavior	Successfully detected abnormal user interactions indicating stalking	Raised concerns regarding user privacy and data security
5	Hybrid Phishing Detection Model (2021)	N. Patel, R. Joshi	Combined rule-based methods with machine learning algorithms for phishing detection	Improved detection rate by integrating multiple techniques	Lacked adaptability to rapidly changing phishing strategies
6	Real-Time Cyber Threat Detection using ML (2022)	K. Mehta, D. Shah	Developed a real-time system using classification algorithms to monitor and detect cyber threats	Provided faster detection and response to phishing attempts	Generated false positives in certain scenarios
7	User Behavior Anomaly Detection System (2023)	P. Kulkarni, S. Deshpande	Used anomaly detection techniques to analyze user behavior and detect irregular patterns	Identified suspicious activities effectively in controlled environments	Required continuous monitoring and large volumes of data
8	AI-Based Email Security	T. Banerjee, A.	Utilized AI and NLP models to detect phishing emails and malicious	Improved accuracy in identifying advanced phishing emails	Limited generalization for unseen and evolving

	Framework (2024)	Chatterjee	attachments		threats
9	Feature-Based Phishing Detection Model (2022)	H. Mishra, R. Tiwari	Extracted URL features such as domain age and length to classify phishing websites	Achieved moderate detection performance	Dependent on feature quality and selection process
10	Integrated AI Cybersecurity Framework (2023)	G. Nair, S. Pillai	Combined machine learning, NLP, and behavioral analysis for comprehensive threat detection	Provided a multi-layered approach to cybersecurity	Required high computational resources and system complexity

IV. PROPOSED SYSTEM

The proposed system integrates phishing detection and stalker analysis into a unified framework. It is designed to analyze both content and behavior to detect threats effectively.The system consists of multiple stages, including input processing, data preprocessing, detection, and alert generation. Initially, the system accepts input in the form of URLs, emails, or user activity data. The data is then processed and analyzed using machine learning models. The phishing detection module identifies malicious content, while the behavioral analysis module monitors user activity patterns. The system continuously tracks user interactions and identifies abnormal behavior. Based on the analysis, it generates alerts to notify users about potential threats.

V. METHODOLOGY

The methodology of the proposed system is divided into several stages to ensure accurate detection.

Data collection is performed using phishing datasets and user activity logs. These datasets provide information about malicious URLs, email content, and user behavior. In the preprocessing stage, data is cleaned and normalized. Text data is tokenized, and important features are extracted. Feature extraction includes identifying URL characteristics, keywords, and behavioral patterns.

The system uses machine learning models such as Random Forest, Decision Trees, and Logistic Regression for classification. NLP techniques

are applied to analyze email content and detect phishing patterns. Behavioral analysis is performed using anomaly detection techniques to identify unusual user activities. The final output is generated by classifying data into safe, suspicious, or malicious categories.

VI. IMPLEMENTATION

The system is implemented using Python due to its simplicity and extensive support for machine learning libraries. The development process involves using libraries such as Scikit-learn, Pandas, NumPy, and NLTK. Flask is used to create a web-based interface for user interaction.

The implementation process includes data preprocessing, model training, and testing. The trained model is used to analyze input data and generate predictions.

The system works in real time, allowing users to input URLs or emails and receive immediate results. Alerts are generated to notify users about potential threats.

VII. CHALLENGES

Several challenges were encountered during the development of the system:

- Rapid evolution of phishing techniques
- Limited availability of high-quality datasets
- Privacy concerns in behavioural analysis

- False positives and false negatives
- Computational complexity of machine learning models

VIII. RESULTS AND DISCUSSION

The proposed system was evaluated based on its accuracy and performance. The results indicate that the system performs well in detecting phishing attacks and identifying suspicious behaviour.

Machine learning models achieved high accuracy in classification tasks. Behavioural analysis improved the detection of unusual activities. The combination of both approaches provided better results compared to traditional systems.

The system demonstrated real-time performance and generated consistent predictions, making it suitable for practical applications.

IX. FUTURE SCOPE

Future improvements can include:

- Integration with browser extensions
- Development of mobile applications
- Use of deep learning models
- Real-time monitoring systems
- Enhanced privacy and security measures

X. CONCLUSION

This research presents an AI-powered system for detecting phishing attacks and analyzing stalker behavior. The integration of machine learning and behavioral analysis provides a comprehensive approach to cybersecurity.

The system improves detection accuracy, enhances user safety, and provides real-time protection. With further advancements, it can be widely deployed to strengthen digital security systems.

XI. ACKNOWLEDGEMENT

We would like to express our sincere gratitude to our project guide and institution for their support and guidance. We also acknowledge the use of datasets and tools that contributed to this research.

XII. REFERENCES

[1]S. Saha and S. K. Saha, “Machine learning-based analysis of phishing URLs,” Proceedings of ICCCA, India, 2016.

[2]A.Singh and S. Gupta, “Heuristic and ML-driven phishing detection framework,” International Journal of Computer Applications, 2019.

[3]R. Mondal and S. De, “Detection of phishing through lexical URL patterns using SVM,” ICISC, Coimbatore, 2020.

[4]W. Xiao, X. Li, Y. Zhang, et al., “Deep learning enhancements for phishing website classification,” IEEE Access, vol. 7, 2019.

[5]H. Liang and W. Zou, “BERT-based evaluation of malicious URLs,” IEEE BigDataSecurity, 2021.

[6]S. Marchal, J. Francois, and N. Asokan, “Advanced strategies for detecting phishing websites,” IEEE Access, 2016.

[7]S. Patil and A. Kapse, “Network forensic inspection for anomaly detection in Indian cyber systems,” ICICCT, 2017.

[8]M. Kumar and S. Singh, “ML-enabled intrusion detection for enterprise communication networks,” ICCES, 2021.

[9]S. Haider, M. Khan, and A. Raza, “CNN–LSTM hybrid architecture for cyberattack identification,” Future Generation Computer Systems, vol. 108, 2020.

[10] CERT-In, “Cyber Security Trends and Incidents in India,” Ministry of Electronics and IT (MeitY), Government of India, 2022.

[11] National Crime Records Bureau (NCRB), “Cyber Crime in India – Annual Statistical Report,” Government of India, 2021.

[12] A. Tavallaee, E. Bagheri, W. Lu, and A. Ghorbani, “Evaluation of the NSL-KDD dataset for intrusion detection research,” IEEE CISDA, 2009.

[13] L. Breiman, “Random Forests for classification and anomaly detection,” Machine Learning Journal, 2001.

[14] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “BERT: Bidirectional transformer representations for language understanding,” in Proc. NAACL, 2019.

[15] S. Gupta, R. Sharma, and P. Kumar, “Flow-level anomaly categorization using ML techniques,” CICN, Bhopal, 2018.

[16] R. Thomas and P. S. Pillai, "AI-enabled cyber-threat prediction model for BFSI networks," ESCI, Pune, 2020.