Archives available at journals.mriindia.com

International Journal on Advanced Computer Engineering and Communication Technology

ISSN: 2278- 5140
Volume 14 Issue 01, 2025

UnveilThreatAI – An AI-Powered Cybersecurity Risk and Consequence Analyzer with Visual Analytics

[1]Dr.A.B.Gavali, [2]Aniket Baral, [3]Ketaki Pawar, [4]Gitanjalee Rakshe
[1 2 3 4]Department of Artificial Intelligence and Data Science
S. B. Patil College of Engineering, Indapur, Pune, India
Email: dnyane.ash@gmail.com [1], aniketbaral08966@gmail.com [2], ketakipawar 877@gmail.com[3], rakshegitanjalee@gmail.com [4]

Peer Review Information

Submission: 11 Sept 2025

Revision: 10 Oct 2025

Acceptance: 22 Oct 2025

Keywords

Cybersecurity, Privacy Risk, Artificial Intelligence, Machine Learning, Visual Analytics, Online Safety

Abstract

As social media and online content sharing have grown in popularity, people frequently share documents, photos, and links without being aware of the hidden cybersecurity risks. Current tools typically only check one kind of content at a time and don't display the potential repercussions of risky behaviour. The literature on current cybersecurity techniques is reviewed in this paper, along with their advantages and disadvantages, including the absence of multi-content analysis, poor clarification, and a lack of visual risk reporting. It also draws attention to the problem statement, which calls for a solution that can assess various kinds of content and assist users in analyzing possible outcomes.

Introduction

The increasing use of social media and online content sharing platforms has led to an unprecedented rise in cybersecurity and privacy threats. People commonly share links, documents, and photos without understanding the dangers. Phishing attempts via hyperlinks, malicious file content, and hidden metadata in photos can all result in data leaks, identity theft, or financial fraud. Existing tools are limited as they detect only one category of threat at a time. There is a need for an integrated solution that not only detects multiple risks but also demonstrates the consequences of unsafe online actions. The contributions of this paper are:

- Development of a web-based platform for multi-content risk detection.
- AI/ML-powered identification of hidden dangers like malicious content, phishing, and meta-data leaks.
- Visualization of the potential chain of consequences for each detected risk.
- The creation of practical suggestions to raise user awareness and enhance online security

The remainder of this paper is organized as follows: Section 2 presents the literature review, Section 3 identifies the research gap, Section 4 describes the proposed system, Section 5 discusses the results and discussion, Section 6 provides a comparison with existing systems, and Section 7 concludes the paper.

Literature Survey

The following table highlights current issues, potential fixes, and future research directions while summarizing important studies pertinent to the proposed UnveilThreatAI system.

Sr. No.	Paper Title	Author(s)	Year	Existing Problem Statement	Existing Problem Solution	Future Scope
1	From Past to Present: A Survey of Malicious URL Detection Techniques, Datasets and Code Repositories [1]	Ye Tian, Yanqiu Yu, Jianguo Sun, Yanbin Wang	2025	Malicious URLs pose threats like phishing, malware, and data theft. Existing surveys lack coverage of new methods (Transformers/LLMs), ignore multimodal features, and provide no unified benchmarks.	Comprehensive survey of detection methods (blacklist, heuristic, ML, DL, Transformer, GNN, LLM). Curated datasets and repositories.	Focus on multimodal detection, standardized benchmarks, and adversarial resilience.
2	Evaluating Large Language Models for Phishing Detection, SelfConsistency, Faithfulness, and Explainability [2]	Shova Kuikel, Aritran Piplai, Palvi Aggarwal	2025	Traditional phishing detectors lack explainability, consistency, and robustness.	Fine-tuned LLMs (BERT, LLaMA, Wizard) with CC-SHAP explainability .	Improve accuracyexplainability balance, integrate cognitive models.
3	Analyzing PDFs like Binaries: Adversarially Robust PDF Malware Analysis via Intermediate Representation and Language Model[3]	Side Liu, Jiang Ming, Guodong Zhou, Xinyi Liu, Jianming Fu, Guojun Peng	2025	PDF malware classifiers are vulnerable to adversarial attacks and outdated feature engineering.	Proposed PDFObj IR + Object Reference Graph + PDFObj2Vec with language models & GNN, achieving 99.9% accuracy and strong adversarial robustness.	Enhance robustness, integrate into enterprise pipelines, adapt to future obfuscation tactics.
4	EXPLICATE: Enhancing Phishing Detection through Explainable AI and LLM-Powered Interpretability [4]	Bryan Lim, Roman Huerta, Alejandro Sotelo, Anthonie Quintela, Priyanka Kumar	2025	Existing phishing detectors act as black boxes, reducing trust and accuracy.	ML classifier + LIME, SHAP, and LLM-based explanations; GUI with Chrome extension.	Extend to brand impersonation attacks, adaptive learning, real-time detection.
5	A Comprehensive Review of AI's Current Impact and Future Prospects in Cybersecurity [5]	Abdullah Al Siam, Moutaz Alazab, Albara Awajan, Nuruzzam	2024	Lack of clarity on AI's effectiveness in cybersecurity.	Review of AI in malware, phishing, intrusion detection, risk analysis.	Advance explainable AI, real-time cyber defense, address privacy issues.

		an Faruqui				
6	A State-of-the-Art Review on Phishing Website Detection Techniques [6]	Wenhao Li, Selvakumar Manickam, Yung-Wey Chong, Weilan Leng, Priyadasri Nanda	2024	Traditional blacklist methods fail against phishing websites.	Survey of ML/DLbased detection methods (feature-based, contentbased, hybrid).	Combine deep learning with explainable AI and real-time detection.
7	Phishing Website Detection Using Deep Learning Models [7]	Ume Zara, Kashif Ayub, Hikmat Ullah Khan, Ali Daud, Tariq Alsahfi, Saima Gulzar	2024	Difficulty detecting evolving phishing patterns and hidden URLs.	DL models (CNN, RNN, hybrid) for phishing website detection.	Enhance accuracy, multilingual capability, better generalization to new attacks.
8	A Survey of Malware Detection Using Deep Learning [8]	Ahmed Bensaouda, Jugal Kalita, Mahmoud Bensaouda	2024	Malware evolves rapidly; traditional methods fail. DL models lack explainability and are vulnerable to adversarial attacks.	Survey of DL techniques (CNN, DNN, RNN, GAN, transfer learning) for multiple OS platforms.	Standardized benchmark datasets, improve XAI, strengthen adversarial robustness.
9	A Survey on Immersive Cyber Situational Awareness Systems [9]	Hussain Ahmad, Faheem Ullah, Rehan Jafri	2024	Current cyber SA tools rely on 2D visualization, causing high cognitive load.	Survey of immersive technologies (VR/AR/MR) for cybersecurity.	Apply immersive SA in SOCs, AI-driven analytics, real-time defense.
10	Enhancing Steganography Detection with AI: Fine-Tuning a Deep Residual Network for Spread Spectrum Image Steganography [10]	Oleksandr Kuznetsov, Emanuele Frontoni, Kyrylo Chernov, Kateryna Kuznetsova, Ruslan Shevchuk, Mikolaj Karpinski	2024	DL-based steganalysis performs poorly on SSIS.	Fine-tuned SRNet for SSIS datasets; improved detection accuracy	Develop multi-task learning, better datasets, reduce performance tradeoffs.
11	A Convolutional Neural Network to Detect Possible Hidden Data in Spatial Domain Images[11]	Jean De La Croix Ntivuguruzwa, Tohari Ahmad	2023	ML-based steganalysis gives low accuracy and unstable training.	CNN with spatial rich filters, depthwise separable convolutions,	Extend to multimedia, improve real-time detection.

					and multiscale pooling.	
12	Explainable Artificial Intelligence (XAI) for Malware Analysis [12]	Harikha Manthena, Shaghayegh Shajarian, Jeffrey Kimmell, Mahmoud Abdelsalam, Sajad Khorsandroo, Maanak Gupta	2023	Malware detection with DL is opaque, reducing trust.	SHAP, LIME, attention, visualization-based explainability techniques.	Domain-specific XAI for cybersecurity, integrate into SOCs, improve robustness.
13	TransURL: Improving Malicious URL Detection with Multi-layer Transformer Architecture [13]	Baijian Yang, Quanyan Zhu, et al.	2022	Existing ML/DL URL classifiers fail on zeroday malicious URLs.	Multi-layer Transformer (TransURL) leveraging self-attention; outperforms CNN/RNN.	Apply to multilingual/ phishing URLs, integrate with threat intelligence.
14	Visualizing Privacy Utility Trade-Offs in Differentially Private Data Releases [14]	Priyanka Nanayakkara, Johes Bater, Xi He, Jessica Hullman, Jennie Rogers	2022	Practitioners lack intuitive tools to balance privacy vs. accuracy.	ViP visualization showing trade-offs between ϵ, accuracy, and risk.	Extend beyond healthcare, integrate AI-driven analytics, improve usability.
15	Applications of Deep Learning for Phishing Detection: A Systematic Literature Review [15]	Cagatay Catal, Gorkem Giray, Bedir ¨ Tekinerdogan, Sandeep Kumar, Suyash Shukla	2022	Blacklist methods fail for zero-day phishing; ML needs handcrafted features.	Systematic review of 43 DL models (CNN, RNN, LSTM, DNN) for phishing detection.	Explore hybrid DL+NLP models, improve zero-day phishing detection.

Research Gap

Several shortcomings in current cybersecurity solutions are discovered by the literature review few of them are listed as follows,

- **Lack of Multimodal Integration**: The majority of tools only analyze single type of content, such as documents, images, URLs. And do not provide a single method for analyzing threats to multiple content types.
- **Absence of Consequence Visualization**: The current systems don't to produces the potential chain of consequences limits users' understanding of risks.
- **Limited Explainability**: While some studies make use of explainable AI, they usually don't provide comprehensive risk reporting and instead only examine specific threats (such as phishing).
- **Scalability Issues**: Current tools struggle with real-time analysis of diverse inputs, reducing their effectiveness in dynamic environments.
- **Absence of User-Centric Design**: Few solutions provide actionable suggestions to raise user awareness and encourage safe practices, or user-friendly interfaces. By

combining the multi-content analysis, consequence visualization, and clear reporting into a single platform, UnveilThreatAI fills these gaps.

This paper presented UnveilThreatAI, AI-powered web-based platform that integrates risk assessment and consequence visualization. Unlike existing methods, it combines different types of content into a single, coherent system and provides a visual understanding of risk impact.

Problem Statement

Often, users are unaware of the hidden risks in the content they share. The comprehensive, AI-based approach that covers documents, hyperlinks, and images and illustrates the potential outcomes is not offered by current cybersecurity tools.

Conclusion

This paper presented UnveilThreatAI, AI-powered web-based platform that integrates risk assessment and consequence visualization. Unlike existing methods, it combines different types of content into a single, coherent system and provides a visual understanding of risk impact.

References

[1] Y. Tian, Y. Yu, J. Sun, and Y. Wang, “From Past to Present: A Survey of Malicious URL Detection Techniques, Datasets and Code Repositories,” 2025.

[2] S. Kuikel, A. Piplai, and P. Aggarwal, “Evaluating Large Language Models for Phishing Detection, Self-Consistency, Faithfulness, and Explainability,” 2025.

[3] S. Liu, J. Ming, G. Zhou, X. Liu, J. Fu, and G. Peng, “Analyzing PDFs like Binaries: Adversarially Robust PDF Malware Analysis via Intermediate Representation and Language Model,” 2025.

[4] B. Lim, R. Huerta, A. Sotelo, A. Quintela, and P. Kumar, “EXPLICATE: Enhancing Phishing Detection through Explainable AI and LLM-Powered Interpretability,” 2025.

[5] A. Al Siam, M. Alazab, A. Awajan, and N. Faruqui, “A Comprehensive Review of AI’s Current Impact and Future Prospects in Cybersecurity,” 2024.

[6] W. Li, S. Manickam, Y. W. Chong, W. Leng, and P. Nanda, “A State-of-the-Art Review on Phishing Website Detection Techniques,” 2024.

[7] U. Zara, K. Ayub, H. U. Khan, A. Daud, T. Alsahfi, and S. Gulzar, “Phishing Website Detection Using Deep Learning Models,” 2024.

[8] A. Bensaouda, J. Kalita, and M. Bensaouda, “A Survey of Malware Detection Using Deep Learning,” 2024.

[9] H. Ahmad, F. Ullah, and R. Jafri, “A Survey on Immersive Cyber Situational Awareness Systems,” 2024.

[10] O. Kuznetsov, E. Frontoni, K. Chernov, K. Kuznetsova, R. Shevchuk, and M. Karpinski, “Enhancing Steganography Detection with AI: Fine-Tuning a Deep Residual Network for Spread Spectrum Image Steganography,” 2024.

[11] J. D. L. C. Ntivuguruzwa and T. Ahmad, “A Convolutional Neural Network to Detect Possible Hidden Data in Spatial Domain Images,” 2023.

[12] H. Manthena, S. Shajarian, J. Kimmell, M. Abdelsalam, S. Khorsandroo, and M. Gupta, “Explainable Artificial Intelligence (XAI) for Malware Analysis,” 2023.

[13] B. Yang, Q. Zhu, et al., “TransURL: Improving Malicious URL Detection with Multi-layer Transformer Architecture,” 2022.

[14] P. Nanayakkara, J. Bater, X. He, J. Hullman, and J. Rogers, “Visualizing Privacy-Utility Trade-Offs in Differentially Private Data Releases,” 2022.

[15] C. Catal, G. Giray, B. Tekinerdogan, S. Kumar, and S. Shukla, “Applications of Deep Learning for Phishing Detection: A Systematic Literature Review,” 2022.