INTERNATIONAL JOURNAL FOR LEGAL RESEARCH AND ANALYSIS

Open Access, Refereed Journal Multi Disciplinary
Peer Reviewed

www.ijlra.com

ISSN: 2582-6433

DISCLAIMER

ISSN: 2582-6433

ABOUT US

The *International Journal for Legal Research & Analysis (IJLRA)* (ISSN: 2582-6433) is a peer-reviewed, academic, online journal published on a monthly basis. The journal aims to provide a comprehensive and interactive platform for the publication of original and high-quality legal research.

IJLRA publishes Short Articles, Long Articles, Research Papers, Case Comments, Book Reviews, Essays, and interdisciplinary studies in the field of law and allied disciplines. The journal seeks to promote critical analysis and informed discourse on contemporary legal, social, and policy issues.

The primary objective of IJLRA is to enhance academic engagement and scholarly dialogue among law students, researchers, academicians, legal professionals, and members of the Bar and Bench. The journal endeavours to establish itself as a credible and widely cited academic publication through the publication of original, well-researched, and analytically sound contributions.

IJLRA welcomes submissions from all branches of law, provided the work is original, unpublished, and submitted in accordance with the prescribed submission guidelines. All manuscripts are subject to a rigorous peer-review process to ensure academic quality, originality, and relevance.

Through its publications, the *International Journal for Legal Research & Analysis* aspires to contribute meaningfully to legal scholarship and the development of law as an instrument of justice and social progress.

ISSN: 2582-6433

PUBLICATION ETHICS, COPYRIGHT & AUTHOR RESPONSIBILITY STATEMENT

The *International Journal for Legal Research and Analysis (IJLRA)* is committed to upholding the highest standards of publication ethics and academic integrity. All manuscripts submitted to the journal must be original, unpublished, and free from plagiarism, data fabrication, falsification, or any form of unethical research or publication practice. Authors are solely responsible for the accuracy, originality, legality, and ethical compliance of their work and must ensure that all sources are properly cited and that necessary permissions for any third-party copyrighted material have been duly obtained prior to submission. Copyright in all published articles vests with IJLRA, unless otherwise expressly stated, and authors grant the journal the irrevocable right to publish, reproduce, distribute, and archive their work in print and electronic formats. The views and opinions expressed in the articles are those of the authors alone and do not reflect the views of the Editors, Editorial Board, Reviewers, or Publisher. IJLRA shall not be liable for any loss, damage, claim, or legal consequence arising from the use, reliance upon, or interpretation of the content published. By submitting a manuscript, the author(s) agree to fully indemnify and hold harmless the journal, its Editor-in-Chief, Editors, Editorial Board, Reviewers, Advisors, Publisher, and Management against any claims, liabilities, or legal proceedings arising out of plagiarism, copyright infringement, defamation, breach of confidentiality, or violation of third-party rights. The journal reserves the absolute right to reject, withdraw, retract, or remove any manuscript or published article in case of ethical or legal violations, without incurring any liability.

ISSN: 2582-6433

PHISHING THROUGH SOCIAL MEDIA: A STUDY ON THREATS AND PROTECTIVE STRATEGIES WITH CYBERSEURITY

AUTHORED BY - NARAIN SIVAKUMAR[1]

Abstract

Social media platforms have become highly appealing targets by attackers and phishing attacks have become one of the most pervasive and dynamic forms of cybersecurity threats in the digital age. The research paper discusses the complexity of the problem of phishing on social media platforms, explains the mechanisms of vulnerability of the users to the attacks psychologically and technically, and analyzes the existing protective methods and the cybersecurity measures. The research establishes that the problem of phishing in social media platforms is growing because of the accessibility, manipulation with the psychological factor, and the use of artificial intelligence and social engineering tools more popular than ever before. Conventional detection systems such as blacklist-based systems and signature detection have not been found to be effective against emerging attack vectors. Some of the most significant results of the research are as follows: (1) the use of machine learning models, especially, XGBoost and Random Forest, can already reach up to 99.7 percent detection rates when trained on balanced samples; (2) human factor and user awareness are crucial elements of phishing protection, and training helps to increase detection rates; (3) AI-based phishing schemes are a new form of threat that should be matched with adaptive detection strategies; and (4) a unified strategy that combines technology, education, and policy is an effective way towards sustainable cybersecurity resilience. The research that should be done in the future is on predictive analytics, behavioral monitoring systems, multilingual detection models, and zero-day phishing. The paper will help in the comprehension of the threats posed by phishing in the social media environment and offer practical recommendations that can help individual users and organizations to improve their cybersecurity stance against phishing attacks that are increasingly becoming sophisticated.

Keywords: Phishing Attacks, Social Media Security, Cybersecurity Threats, Machine Learning Detection, User Awareness

[1] Student at christ (deemed to be university), pune, lavasa campus, 30 valor ct, 412112,

ISSN: 2582-6433

Introduction

Phishing has developed out of the simple mass-mailing campaigns into a highly sophisticated, targeted attack-vector that focuses on the psychological susceptibility of humans, and applies social engineering to interfere with sensitive data. The spread of social media platforms (a figure of more than 4 billion active users across the world) has provided an unprecedented chance to the attacker to access and control victims on a large scale[2]. In contrast to the classic phishing by email, social media phishing functions on the trusting environment, where people are already inclined to engage with the information posted by their friends or followers, brands, and other influencers[3]. The current paper explores the growing risk of phishing through social media platforms, the reasons as to why users are exposed to such risks despite the advancement in technology and provides an assessment of new-protective measures against these threats in the current context of cybersecurity threats.

The relevance of this study is that phishing is not only a technical issue but a synthesis of psychological manipulation, social engineering, and platform design weakness and lack of user literacy on cyber threats. Conventional methods of cybersecurity that concentrate on technical protection measures have failed continuously to stop phishing attacks as historical instances of major breaches of critical infrastructures of organizations such as Twitter, Wipro, and Anthem Inc. show. The advent of AI-driven phishing attacks, deepfakes, and voice-assisted social engineering further complicates an already difficult threat environment[4].

The research questions that will be used in this research are three-fold, which are (1) to analyze the mechanisms and psychological tricks of phishing attacks on social media; (2) to examine the capacity of the current cybersecurity and detection technologies; and (3) to suggest comprehensive protective measures based on technology, education, and policy frameworks of individual users and organization. The research will answer two important research questions: (a) What makes the users vulnerable to the phishing on social media platforms and (b) What are the best cybersecurity mechanisms that can protect against and identify such attacks?

The Evolution and Nature of Phishing Attacks

[2] Fiona Carroll, John Ayooluwa Adejobi and Reza Montasari, "How Good Are We at Detecting a Phishing Attack? Investigating the Evolving Phishing Attack Email and Why It Continues to Successfully Deceive Society," SN Computer Science 3:170 (2022).
[3] Rex Martini, Phishing with Attention: A Study on AI-Enhanced Phishing Leveraging Retrieval-Augmented Generation (RAG) 2025, Master's Thesis, Karlstad University.
[4] Edwin Donald Frauenstein and Stephen Flowerday, "Susceptibility to Phishing on Social Network Sites: A Personality Information Processing Model," Computers and Security 94:101862 (2020).

ISSN: 2582-6433

Historical Context and Evolution

Phishing attacks have evolved greatly since the 1990s when they were first introduced. The research literature shows that the development of phishing results in several stages, the first stage is the early mass-mailing campaigns (financial institutions), then spear-phishing against individuals and organizations, and the modern phase of AI-enhanced and multi-channel attacks. In their study, Carroll et al. (2022) evaluated the perception of phishing emails in various periods (1998-2020) and proved that phishing is developing at a faster pace than the awareness of users, and attackers are continuously improving their tools to avoid detection by technology and people. Social media phishing as a shift of email-based phishing is a fundamental change in an attack strategy. Attackers can take several benefits on social media platforms, including access to personal data (which allows them to make targeted attacks), the machinery of establishing false credibility by impersonating accounts, and the amplification of malicious information by algorithms[5].

The transition from email-based phishing to social media phishing represents a fundamental shift in attack strategy. Social media platforms provide attackers with multiple advantages: access to personal information (enabling targeted attacks), mechanisms for building false credibility through account impersonation, and algorithmic amplification of malicious content[6]. The NASSCOM v. Ajay Sood (2005) case in the Delhi High Court marked India's first judicial recognition of phishing as an illegal offense, highlighting the evolving legal recognition of these threats[7].

Types of Phishing Attacks on Social Media

Modern phishing tricks can be seen in various forms in social networks. Deceptive phishing consists of generic fraud messages that are aimed at duping users to provide sensitive information.Spear-phishing focuses on particular users based on personal data obtained on social media accounts. Clone phishing is the act of impersonating authentic social media pages or accounts. Whaling appeals to high-value individuals including executives and organization decision-makers[8]. New types are QR code phishing (quishing), deepfake video and voice

[5] Fiona Carroll, John Ayooluwa Adejobi and Reza Montasari, "How Good Are We at Detecting a Phishing Attack? Investigating the Evolving Phishing Attack Email and Why It Continues to Successfully Deceive Society," SN Computer Science 3:170 (2022).
[6] Edwin Donald Frauenstein and Stephen Flowerday, "Susceptibility to Phishing on Social Network Sites: A Personality Information Processing Model," Computers and Security 94:101862 (2020).
[7] National Association of Software and Service Companies (NASSCOM) v. Ajay Sood & Others (2005), Delhi High Court, 119 (2005) DLT 596.
[8] Vaishnavi Bhavsar, Aditya Kadlak and Shabnam Sharma, "Study on Phishing Attacks," International Journal of Computer Applications 182(33):27–29 (2018).

ISSN: 2582-6433

phishing, and AI-generated lures based on retrieval-augmented generation (RAG) to generate very contextual and persuasive scam content. The psychological processes involved in the successful phishing attacks in the social media are urgency and fear (creating time pressure to avoid careful deliberation), authority use (impersonating an organization or people one can trust), social proof (utilizing peer networks and recommendations)[9], and reciprocity (providing incentives in exchange of information). Such mental tricks are being further enhanced by AI systems that are able to produce personalized and contextually relevant fraudulent messages[10].

Vulnerabilities in User Behavior and Social Media Platforms

Human Factors in Phishing Susceptibility

Studies have always indicated that the user behavior and decision-making process is one of the most critical weaknesses in phishing defense. Frauenstein and Flowerday (2020) investigated the role of personality traits in social network sites phishing vulnerability, which is a Big Five personality model[11]. Their analysis of 215 students of South African universities established that personality traits have a very strong influence on susceptibility, but this is largely mediated by heuristic (not systematic) information processing. The users who use fast and intuitive judgments instead of critical analysis are much more likely to fall victim to a phishing attack. Toukabri (2024) discovered that knowledge and previous familiarity with cyber incidents largely affect online behaviors. Learners who had the background in IT or had earlier exposure to phishing attempts were significantly more cautious and detected phishing attempts more effectively[12]. On the other hand, little awareness and low cyber literacy are factors that lead to risky behaviors over the internet and users do not know the signs of maliciousness in communication. The study by Al-Naimi (2024) with eye-tracking results reveals neurobiological evidence that phishing detection is nearly equally in the scope of human thinking and attention attention patterns, and technological indicators, and that there are particular gaze patterns of successful and failed phishing detection[13].

[9] Yousef Khalid Al-Hamar, An Enhancement on Targeted Phishing Attacks in the State of Qatar 2019, Ph.D. Thesis, Liverpool John Moores University.
[10] Rex Martini, Phishing with Attention: A Study on AI-Enhanced Phishing Leveraging Retrieval-Augmented Generation (RAG) 2025, Master's Thesis, Karlstad University.
[11] Fiona Carroll, John Ayooluwa Adejobi and Reza Montasari, "How Good Are We at Detecting a Phishing Attack? Investigating the Evolving Phishing Attack Email and Why It Continues to Successfully Deceive Society," SN Computer Science 3:170 (2022).
[12] Edwin Donald Frauenstein and Stephen Flowerday, "Susceptibility to Phishing on Social Network Sites: A Personality Information Processing Model," Computers and Security 94:101862 (2020).
[13] Shaikha Jamal M. S. Al-Naimi, PhishTracker: Exploiting Eye-Tracking to Analyze User Behavior Under Phishing Attacks 2024, Master's Thesis, Hamad Bin Khalifa University.

ISSN: 2582-6433

Toukabri (2024) found that awareness and prior experience with cyber incidents substantially influence online behavior. Students with IT backgrounds or previous exposure to phishing attempts demonstrated markedly greater caution and better detection capabilities. Conversely, low awareness and limited cyber literacy contribute to risky online behavior, with users failing to recognize malicious indicators in communications. The eye-tracking study by Al-Naimi (2024) provides neurobiological evidence that phishing detection depends as much on human cognition and attention patterns as on technological measures, identifying specific gaze patterns associated with successful and failed phishing detection[14].

Platform Design and Information Architecture

Phishing attacks are supported by social media platforms that are designed in ways that are unintentional. The culture of fast content and content density provides settings in which users possess fewer time and cognitive resources to critically assess the authenticity of content. Platform algorithmic amplification is able to spread malicious content across platforms before detection systems can react. The social media platforms of third-party application integration and authentication mechanisms are more attack surfaces[15].

Although trust-based design is aimed at improving the user experience, it can be used by attackers. With connections and verified accounts, users hope that the content they will view will be credible, and they will become less vigilant. The social aspect of such platforms implies that the accounts that have been compromised can be used to spread phishing information in existing circles of trust, which can be harder to detect[16].

Technical Detection and Prevention Mechanisms

Machine Learning and Artificial Intelligence Approaches

Recent research demonstrates remarkable progress in machine learning-based phishing detection. Akhtar et al. (2025) tested multiple machine learning algorithms (XGBoost,

[14] Omar Toukabri, Analysis of Social Media and Phishing Awareness among Diverse Higher Education Students: A Health Belief Model Perspective 2024, Master's Thesis, Örebro University School of Business.
[15] Yousef Khalid Al-Hamar, An Enhancement on Targeted Phishing Attacks in the State of Qatar 2019, Ph.D. Thesis, Liverpool John Moores University.
[16] Zainab Alkhalil, Chaminda Hewage, Liqaa Nawaf and Imtiaz Khan, "Phishing Attacks: A Recent Comprehensive Study and a New Anatomy," Frontiers in Computer Science 3:563060 (2021).

ISSN: 2582-6433

LightGBM, Random Forest, KNN, SVM, CNN) on phishing datasets, finding that XGBoost achieved 99.7% detection accuracy. Rajashekar (2024) identified Random Forest as the most reliable model for phishing website detection, emphasizing the critical importance of feature selection and dataset balancing[17]. These tree-based ensemble methods outperform traditional rule-based approaches by capturing complex patterns in phishing indicators.

Deep learning approaches, particularly transformer-based models, show promise for handling text obfuscation techniques. Niko (2025) demonstrated that transformer-based models like DeBERTa-v3 and ByT5 maintain consistent accuracy even when phishing emails employ leetspeak, Unicode tricks, and invisible symbols techniques that significantly degrade traditional text classification models. Domain-based feature extraction presents an alternative approach: Shirazi (2018) achieved 97% classification accuracy using only seven domain name-related features, avoiding reliance on third-party data sources and privacy-sensitive content analysis[18].

Multi-Modal Detection Systems

Contemporary phishing detection increasingly employs multi-modal approaches combining URL analysis, content examination, behavioral monitoring, and network-level indicators. Nagunwa (2022) developed machine learning models for zero-day phishing detection by combining URL, webpage content, and hosting network features, achieving reliable detection of previously unknown phishing sites[19]. Hakala (2024) demonstrated that CNN-based analysis of URL text alone can provide fast, language-independent detection suitable for real-world deployment[20].

AI-enhanced detection systems incorporating explainable AI (XAI) techniques improve trust and operational effectiveness. Akhtar et al. (2025) found that incorporating LIME (Local Interpretable Model-agnostic Explanations) into machine learning pipelines made detection models more interpretable and trustworthy for practitioners, enhancing real-world adoption. This addresses a critical gap between research laboratory performance and practical

[17] Akash Rajashekar, Phishing Website Detection Using Machine Learning Approaches 2024, Master of Science thesis, California State University, Northridge.
[18] Hossein Shirazi, Unbiased Phishing Detection Using Domain Name Based Features 2018, Master's Thesis, Colorado State University.
[19] Thomas P. Nagunwa, Fast Detection of Zero-Day Phishing Websites Using Machine Learning 2022, Doctoral Thesis, Birmingham City University.
[20] Suvi Tuulia Hakala, Validating Malicious URLs in Phishing Campaigns Using CNNs 2024, Master's Thesis, University of Helsinki, Finland.

ISSN: 2582-6433

deployment effectiveness[21].

User Awareness and Educational Interventions

The Critical Role of User Literacy

Despite technological advances, user awareness remains central to phishing defense. Nadeem et al. (2023) concluded that integrating technical solutions with behavioral strategies is essential for achieving global cybersecurity resilience[22]. Laur (2018) identified a significant research gap: while much phishing literature focuses on victims and technical tools, limited research examines how security professionals design and implement anti-phishing education, and even less addresses how phishing research translates into real-world organizational practices[23]

Baral (2021) developed a game-based anti-phishing education approach, demonstrating that teaching users about phishing is more effective when integrated with assessments of user self-efficacy and cognitive learning styles[24]. The gaming methodology increased engagement and practical applicability compared to traditional awareness training. Bahl (2023) examined reporting behavior in phishing incidents, finding that factors including self-efficacy, ease of reporting mechanisms, social norms, and personal prior experiences significantly influence whether users report suspicious communications. Organizations with simple reporting processes, meaningful feedback, and targeted training demonstrated substantially higher reporting rates and improved incident response capabilities[25].

Behavioral and Organizational Approaches

Al-Hamar (2019) demonstrated that combining intelligent detection technology with targeted user awareness training substantially reduces successful phishing attacks in organizational contexts. His ECSPAD system outperformed commercial alternatives (such as TrendMicro) at detecting targeted phishing attacks, and interactive training significantly increased users ability

[21] Hafiz Muhammad Usman Akhtar, Mustafa Hameed, Sidra Hameed, Muhammad Nauman, Nadeem Akhtar, Muhammad Zeshan Tareen, "Mitigating Cyber Threats: Machine Learning and Explainable AI for Phishing Detection," VFAST Transactions on Software Engineering 13(2):1–12 (2025).
[22] Muhammad Nadeem, Syeda Wajiha Zahra, Muhammad Nouman Abbasi, Ali Arshad, Saman Riaz and Waqas Ahmed, "Phishing Attack, Its Detections and Prevention Techniques," International Journal of Wireless Security and Networks 13–25 (2023)
[23] Brandon Laur, Information Security's Perception of Phishing Tactics 2018, Royal Roads University, Victoria, Canada.
[24] Gitanjali Baral, Building Confidence not to be Phished: Conceptualising Users' Self-Efficacy in Phishing Threat Avoidance Behaviour 2021, Master's Thesis, University of New South Wales, Australia.
[25] Robin Bahl, Catch the Phish: A Study on Decision-Making and Reporting Behavior for Phishing Attacks 2023, Master's Thesis, Delft University of Technology.

ISSN: 2582-6433

to recognize threats. However, effectiveness varies across organizational contexts, with cultural factors influencing receptiveness to awareness programs[26].

Wisani and Masilane (2024) developed a cybersecurity awareness strategy specifically targeting rural communities, identifying that inadequate training, limited resources, and weak policies create unique vulnerabilities in underserved populations. This research highlights the necessity of context-adapted awareness strategies rather than generic organizational training approaches[27]

Emerging Threats and AI-Enhanced Phishing

AI-Generated and Enhanced Attacks

Perhaps the most important new threat to phishing defense is the weaponization of artificial intelligence. Martini (2025) explored AI-assisted phishing attacks under retrieval-augmented generation (RAG) pipelines showing that open-source language models are capable of generating highly contextualized, spear-phishing emails of real-world accuracy that can be close to human-quality attacks.This piece of work emphasizes the unsophisticatedness of the phishing arms race: defense mechanisms that have been built based on standard datasets are not very resistant to attacks generated by AI[28].

According to Gabriel and Precious (2025), several changing attack vectors were identified: AI-based phishing baits, deepfake video and voice scams, QR code phishing, and phishing-as-a-service business models that make phishing available to everyone. These new threats will be cross channel (email, SMS, social media, voice) and present a detection and prevention challenge that goes beyond the single channel defense system. Deepfake and AI-enhanced campaigns have caused financial and data losses in different sectors of the industry in large amounts[29].

Adversarial Evasion and Detection Arms Race

With increasing detection systems attackers have come up with several evasion mechanisms. Frohm Krischél and Fredriksson Karvelas (2024) investigated using large language models

[26] Yousef Khalid Al-Hamar, An Enhancement on Targeted Phishing Attacks in the State of Qatar 2019, Ph.D. Thesis, Liverpool John Moores University.
[27] Pholosho Wisani Masilane, Cybersecurity Awareness Strategy for Rural Communities: A Case Study of the Mopani District in the Limpopo Province 2024, University of Venda.
[28] Aravindan Ragunathan, Certain Investigation on Web Application Security Phishing Detection and Phishing Target Discovery 2016, Research Paper.
[29] Gabriel, Ige & Precious, Adebayo. "Phishing in 2025: How Attacks Are Evolving," ResearchGate (2025).

ISSN: 2582-6433

(ChatGPT, Claude) for phishing detection, finding that prompt engineering significantly influences model effectiveness[30]. However, the closed-source nature of commercial LLMs and potential vulnerabilities to adversarial attacks represent ongoing challenges. Zhen Xue (2024) achieved high detection accuracy (high precision, recall, and ROC AUC) testing 2.7 million URLs with XGBoost and other tree-based methods, but deployment remains limited to English-language URLs, potentially limiting detection of multilingual phishing campaigns[31].

Multi-Layered Defense Framework and Recommendations

Integrated Technological, Educational, and Policy Approaches

Current research consensus emphasizes that no single technology or approach adequately addresses the phishing threat; rather, a coordinated, multi-layered framework is necessary. Alkhalil et al. (2021) proposed a comprehensive phishing attack lifecycle model identifying distinct phases (planning, preparation, attack, data theft), and recommended that defense strategies address vulnerabilities at each phase through integrated technical and non-technical means[32].

The proposed framework includes:

Technology Layer: Advanced machine learning models trained on balanced, current datasets; multi-modal detection combining URL, content, and behavioral analysis; zero-day detection capabilities; and explainable AI systems facilitating practitioner trust and appropriate action.

User Awareness Layer: Contextually adapted education programs considering organizational culture, user demographics, and prior experience; gamified learning approaches enhancing engagement and retention; simple, efficient reporting mechanisms with meaningful feedback; and periodic simulated phishing campaigns measuring and improving detection capabilities.

Organizational Layer: Clear policies and response procedures; integration of cybersecurity into organizational culture; cross-sector collaboration and threat intelligence sharing; and continuous updating of detection systems and training to address evolving threats[33].

[30] Karl Frohm Krischél and Sebastian Fredriksson Karvelas, Go Phish - Detecting Phishing Attacks with Prompt Engineered Large Language Models as Classification Tools 2025, Bachelor Thesis, Stockholm University.
[31] Zhen Xue, Enhancing Phishing Detection Through Machine Learning 2024, Master of Applied Technologies thesis, Unitec Institute of Technology.
[32] Zainab Alkhalil, Chaminda Hewage, Liqaa Nawaf and Imtiaz Khan, "Phishing Attacks: A Recent Comprehensive Study and a New Anatomy," Frontiers in Computer Science 3:563060 (2021).
[33] Robin Bahl, Catch the Phish: A Study on Decision-Making and Reporting Behavior for Phishing Attacks

ISSN: 2582-6433

Platform-Level and Legal Interventions

Social media platforms bear responsibility for reducing phishing affordances within their systems. Recommended platform interventions include enhanced authentication mechanisms (multi-factor authentication, biometric verification); improved account recovery procedures preventing unauthorized access; transparent account verification badges distinguishing legitimate accounts; and rapid response to phishing report submissions.

Legal and regulatory frameworks must evolve to address contemporary phishing threats. The 2005 NASSCOM case in India established phishing as actionable under existing statutes, but comprehensive cybersecurity legislation including the Information Technology Act amendments is necessary to facilitate prosecution and establish clear legal standards for organizational liability[34].

Conclusion

Social media phishing is a multifaceted cybersecurity issue that needs combined efforts in the areas of technology, user behavior, organizational culture, and legal regulations. According to existing research, although machine learning solutions have reached impressive detection rates (up to 99.7 percent in controlled settings), their performance in the real world is limited by the implementation issues, the constantly changing tactics of attackers, and the weaknesses related to human behavior. The design nature of social media sites in terms of trust, fastness of consumption, and network effects unwillingly provides conditions that support phishing attacks.

The core idea about the human factors that should be considered in terms of effective phishing defense implementation is that the human aspect is central, and the personality, past experiences, and cognitive processing styles can impact the susceptibility. Detecting and reporting behaviors are greatly enhanced with the help of proper user awareness and education taking into account the organizational context and differences among individuals. The introduction of AI-improved phishing attacks is something that poses a growing risk demanding constant improvement in the detection schemes where transformer-based models can have the ability to deal with text obfuscation and adversarial evasion schemes.

The research directions to be followed in the future are: (1) the creation of adaptive and real-

2023, Master's Thesis, Delft University of Technology.

[34] National Association of Software and Service Companies (NASSCOM) v. Ajay Sood & Others (2005), Delhi High Court, 119 (2005) DLT 596.

ISSN: 2582-6433

time detection systems that can detect zero-day phishing campaigns; (2) development of multilingual detection systems that go beyond English-based attacks; (3) research into behavioral monitoring systems that can provide continuous protection, not just point-in-time detection; (4) more predictive analytics to identify threats and (5) comparative research of awareness and training interventions in differing organizational settings.

Firms and individuals should apply the Zero Trust philosophy which presupposes that phishing will take place and defensive tactics should be introduced respectively. This involves the introduction of intelligent detection tools, simple yet effective user reporting tools, user awareness, and training and the creation of organizational cultures that make cybersecurity a collective responsibility. Platforms need to re-architecture systems with less phishing affordances without compromising user experience or privacy.

The need to have a multi-layered defense is becoming more evident as phishing attacks keep adapting through the use of AI and deep fakes, as well as new mediums of communication. It is only with a long-term partnership between researchers, cybersecurity professionals, platform services, policymakers and informed users that organizational and personal cybersecurity resilience will be attained in a context of ongoing and advanced phishing attacks.