Journal of Computing and Data Technology

https://journal.krrishpub.com/index.php/jcdt/index

RESEARCH ARTICLE

An Integrated Framework for Phishing Threat Mitigation via Simulation-Driven Analysis, Email Header Forensics, and URL Intelligence

Sabitha Banu [a,*], R Divya [a], Deva Dharshini TT [a], Bhoovana Sri [a], Mehdi Gheisari [b,c,d,e], Saman Khammar [f], Mustafa Ghaderzadeh [g], Saeed Lotfi [h]

[a] *Department of Computer Science with Cybersecurity, PSGR Krishnammal College for Women, Coimbatore, Tamil Nadu,- 641004, India.*
[b] *Institute of Artificial Intelligence, Shaoxing University, 508 West Huancheng Road, Yuecheng District, Zhejiang, 312000, China.*
[c] *Dept. of Computer Science and Engineering, Saveetha School of Engineering, Savitha Institute of Medical and Technical Science, Tamil Nadu, India.*
[d] *Department of Computer Engineering , Islamic Azad University, Shiraz Branch, Shiraz, Iran.*
[e] *Department of R&D , Shenzhen BYD Co LTD, No. 3009, BYD Road, Pingshan, Shenzhen, 518118, China.*
[f] *Faculty of Electrical and Computer Engineering, University of Sistan and Baluchestan, Zahedan, Iran.*
[g] *School of Nursing and Health Sciences of Boukan, Urmia University of Medical Sciences, Urmia, Iran.*
[h] *Department of Computer Engineering, K. N. Toosi University of Technology, Tehran, Iran.*

Abstract

Phishing attacks continue to represent a dominant vector for cyber intrusions, exploiting human vulnerabilities and systemic weaknesses in email communication infrastructures. This study presents an integrated and simulation-driven cybersecurity framework that combines phishing campaign emulation, email header forensics, and URL intelligence analysis to enable proactive threat detection and mitigation. The proposed approach leverages open-source platforms to systematically replicate real-world phishing scenarios, thereby facilitating controlled experimentation and behavioral analysis. Email header inspection is employed to extract and validate metadata attributes such as sender authenticity, routing paths, and authentication protocols, such as Sender Policy Framework (SPF), DomainKeys Identified Mail (DKIM), and Domain-based Message Authentication, Reporting, and Conformance (DMARC) are used to verify the authenticity of email senders and detect spoofing attempts. Concurrently, URL scrutiny mechanisms incorporate blacklist verification, domain reputation assessment, and structural pattern analysis to detect malicious or obfuscated links. The framework is designed to enhance situational awareness by correlating insights derived from simulation outputs, header anomalies, and URL threat indicators. Experimental evaluation demonstrates improved detection accuracy and response efficiency when compared to isolated analysis techniques. Furthermore, the integration of feedback-driven awareness mechanisms strengthens organizational resilience against evolving phishing tactics. The proposed methodology not only provides a comprehensive analytical perspective on phishing attack vectors but also contributes a scalable and cost-effective solution for real-world deployment in enterprise security ecosystems.

Keywords: Phishing attacks, Email header analysis, URL scrutiny, Cybersecurity, Threat mitigation.

Article information:
ISSN: 3107-9466 (Online)
Published by: **Krrish Scientific Publications Pvt. Ltd.**
DOI: https://doi.org/10.71426/jcdt.v2.i1.pp121-130
Received: *25 Feb. 2026* | Revised: *25 Apr. 2026* | Accepted: *02 May 2026* | Published: *07 May 2026*

*Corresponding author
Email addresses: sabithabanu@psgrkcw.ac.in (Sabitha Banu), mehdi.gheisari61@gmail.com, MEHDI.gheisari@yandex.com (Mehdi Gheisari), samankhammar72@gmail.com (Saman Khammar), mustafa.ghaderzadeh@gmail.com (Mustafa Ghaderzadeh), slotfi72@gmail.com (Saeed Lotfi).

1. Introduction

Phishing attacks have emerged as one of the most persistent and damaging cybersecurity threats, targeting individuals and organizations through deceptive communication channels such as emails, malicious URLs, and spoofed websites. These attacks exploit human cognitive vulnerabilities and technical loopholes to gain unauthorized access to sensitive information, financial assets, and enterprise systems [3], [23], [33], [36]. As cyber adversaries continuously evolve

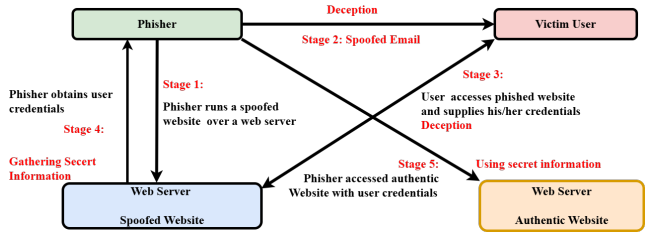

Figure 1: General architecture of phishing attack lifecycle.

their strategies, there is an increasing need for proactive and multi-layered defense mechanisms capable of detecting and mitigating phishing threats in real time.

The scale and sophistication of phishing attacks have increased significantly in recent years. Reports indicate a substantial surge in phishing incidents, particularly in the financial sector, accompanied by exponential growth in credential harvesting campaigns [1]. High-value industries such as finance and insurance remain primary targets, accounting for a significant proportion of global phishing activity [2], [3]. Modern phishing campaigns increasingly leverage artificial intelligence, automation, and multi-channel delivery mechanisms, extending beyond email into messaging platforms and collaborative environments [4], [32], [33], [3]. Furthermore, geographically diverse threat actors continue to exploit global digital infrastructures, making phishing a widespread and persistent challenge [25].

Traditional phishing detection approaches often rely on isolated mechanisms such as spam filtering, blacklist verification, or manual analysis, which are insufficient against modern adaptive threats [10]. Recent studies emphasize the need for integrated frameworks combining multiple detection strategies to improve accuracy and resilience [2], [30], [31]. However, existing solutions frequently lack coordination between simulation, forensic analysis, and threat intelligence mechanisms.

To address these limitations, this research adopts a three-pronged approach:

1. Phishing simulation: Controlled simulation of phishing campaigns to evaluate user susceptibility and organizational readiness.
2. Email header analysis: Forensic examination of email metadata to detect spoofing, authentication failures, and routing anomalies.
3. URL scrutiny: Comprehensive analysis of embedded links using reputation systems and threat intelligence platforms.

A general architecture of phishing attack lifecycle is shown in Figure 1. By integrating these complementary approaches, the proposed framework aims to provide a holistic and scalable solution for phishing threat detection and mitigation. The study contributes toward improving organizational cybersecurity posture by combining behavioral analysis, technical forensics, and intelligence-driven detection mechanisms.

2. Literature review

2.1. Phishing simulation and user awareness

Phishing simulation has been widely adopted as an effective method to assess user awareness and organizational vulnerability. Controlled phishing campaigns enable the identification of human-centric weaknesses and the evaluation of training effectiveness [23]. Automated simulation frameworks further enhance scalability and allow large-scale behavioral analysis across organizational units [8]. However, limitations remain in terms of realism and adaptability to evolving phishing strategies.

2.2. Email header analysis

Email header analysis plays a critical role in identifying spoofed identities and tracing the origin of phishing emails. Techniques based on metadata inspection, routing path verification, and authentication validation (SPF, DKIM, DMARC) have demonstrated effectiveness in detecting anomalies [4], [34]. Integration with open-source intelligence tools enables real-time detection of suspicious sources and enhances threat visibility [15]. Nevertheless, standalone header analysis may generate false positives and may not fully capture sophisticated social engineering attacks [17].

2.3. URL-based phishing detection

URL analysis has become a central component of phishing detection frameworks. Approaches based on lexical features, domain reputation, and machine learning techniques have shown significant improvements in identifying malicious URLs [6]. Combining URL analysis with other indicators, such as email metadata, further enhances detection accuracy and reduces false negatives [16]. However, attackers often evade detection through dynamic domain generation, URL obfuscation, and fast-flux hosting techniques [35].

2.4. DNS-based and AI-driven techniques

DNS-based detection mechanisms leverage domain registration data and blacklist systems to identify phishing infrastructure. These techniques are effective in blocking known threats but face challenges in detecting rapidly changing domains [11], [13]. Artificial intelligence and machine learning approaches have been introduced to improve real-time detection by analyzing patterns in email content and URL structures [23], [24]. Despite improved accuracy, these methods require continuous model updates and substantial computational resources.

2.5. Awareness and defense mechanisms

User awareness training and authentication protocols remain essential components of phishing defense strategies. Studies highlight the effectiveness of training programs in reducing susceptibility, although outcomes vary based on organizational culture and engagement levels [25], [26],[5]. Authentication mechanisms such as SPF, DKIM, and DMARC significantly enhance email security but require proper configuration and enforcement [27]–[29]. Forensic approaches provide detailed insights into phishing incidents but are often unsuitable for real-time deployment [18], [20], [21].

2.6. Emerging technologies

Recent research explores advanced detection mechanisms, including URL reputation systems, malware analysis, blockchain-based security frameworks, and cloud-based protection models [7], [14], [15] [12]. Machine learning and natural language processing techniques have also been applied to improve email classification and phishing detection accuracy [22], [33], [34]. While promising, these approaches face challenges related to scalability, integration complexity, and evolving threat landscapes.

2.7. Limitations of existing approaches

Despite significant advancements, existing phishing detection techniques exhibit several limitations:

1. Lack of practical simulation-based evaluation of user behavior [9].
2. Dependence on manual analysis for header and URL inspection [17].
3. Limited integration of threat intelligence sources [19].
4. Inability to adapt to rapidly evolving phishing techniques [2], [30]–[32].

These limitations highlight the need for integrated frameworks that combine simulation, forensic analysis, and intelligence-driven detection mechanisms.

3. Methodology

This section deals with the methodology of the proposed framework and a short description of the cybersecurity tools used for the execution of work. Figure 2 illustrates the flowchart of the phishing simulation environment setup and verification process. Table 1 provides a comparative analysis and justification of selected cybersecurity tools.

3.1. Tools description

3.1.1. GoPhish- phishing campaign simulation

GoPhish is an open-source phishing simulation tool designed to test an organization's resilience against phishing attacks [2], [3]. It enables security teams to send simulated phishing emails to employees, tracks user interactions such as link clicks and credential submissions, and provides detailed reports on phishing campaign performance and user awareness levels.

3.1.2. RubikPhish-automated phishing awareness testing

RubikPhish is a phishing simulation tool that helps organizations test employee awareness through automated phishing exercises [8]. It generates realistic phishing emails to assess user response and provides real-time feedback to improve cybersecurity training programs.

3.1.3. MXToolBox-email header analysis

MXToolBox is widely used for analyzing email headers to detect phishing attempts and email spoofing [4], [5]. It identifies anomalies in email headers and verifies SPF, DKIM, and DMARC authentication records.

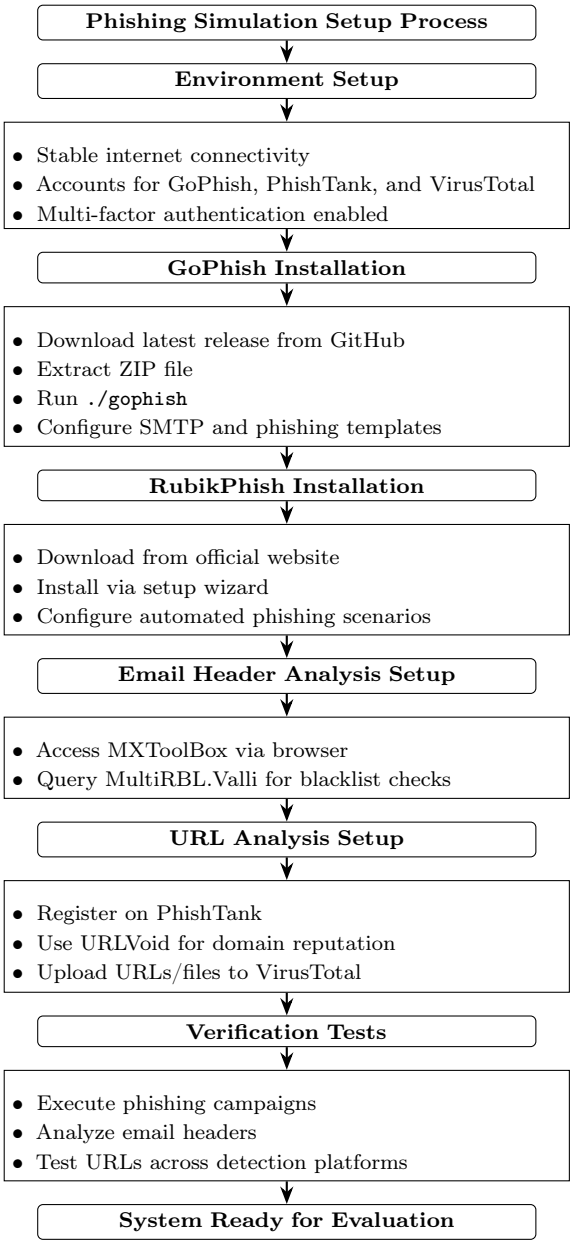

Figure 2: Flowchart of the phishing simulation environment setup and verification process.

3.1.4. MultiRBL.Valli- email blacklist lookup

MultiRBL.Valli checks email headers against known blacklists [15]. It helps detect email servers associated with phishing campaigns and identifies domains that have been flagged as suspicious or malicious. The output uses different colors to indicate the status of the queried IP or domain:It helps detect email servers associated with phishing campaigns and identifies domains that have been flagged as suspicious or malicious. The output uses different colors to indicate the status of the queried IP or domain:

- Green indicates the IP or domain is not listed in the checked blacklist.
- Red indicates the IP or domain is listed in the respective blacklist.
- Yellow/Orange indicates the blacklist has timeouts or

issues, meaning the result might be unreliable.The blacklist has timeouts or issues, meaning the result might be unreliable.
- Gray indicates the tool could not retrieve data from the specific blacklist, possibly due to connection problems.

3.1.5. PhishTank- phishing URL database

PhishTank is an open database of reported phishing sites used to identify malicious URLs [26]. It provides continuously updated of verified phishing sites and helps organizations block access to known phishing domains.

3.1.6. URLVoid- URL reputation analysis

URLVoid analyzes website reputations to determine associations with phishing attacks [6], [7]. It cross-checks URLs with multiple security engines and blacklists and provides risk assessments for suspicious links.

3.1.7. VirusTotal - URL and attachment security scanning

VirusTotal is a comprehensive security analysis platform that scans URLs, email attachments, and files for malware and phishing threats [16]. It uses multiple antivirus engines to detect known phishing links and helps identify and prevent malware infections from phishing emails.

The phishing simulation environment was set up using GoPhish and RubikPhish for executing realistic phishing attack simulations. The environment was designed to be lightweight and customizable for conducting different types of phishing campaigns, while also providing a secure environment for testing organizational phishing awareness and response strategies. This work uses GoPhish, RubikPhish, MXToolBox, MultiRBL.Valli, URLVoid, PhishTank, and VirusTotal. These tools were selected for their open-source nature, ease of integration, and relevance in real-world threat detection environments. A short description of the tools are discussed in the following subsections.

3.1.8. PhishTank — phishing URL database

PhishTank is an open database of reported phishing sites used to identify malicious URLs [26]. It provides continuously updated verified phishing records.

3.1.9. URLVoid — URL reputation analysis

URLVoid analyzes website reputations to determine associations with phishing attacks [6]. It cross-checks URLs across multiple security engines and blacklists to provide risk assessments.

3.1.10. VirusTotal — URL and attachment security scanning

VirusTotal scans URLs, email attachments, and files for malware and phishing threats [16]. It utilizes multiple antivirus engines to detect malicious links and infected files.

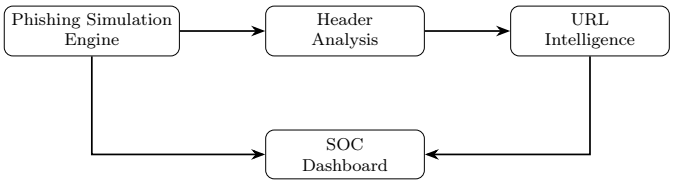

Figure 3: Integrated phishing detection workflow within the SOC environment

3.2. Framework of Security Operations Center(SOC)-based phishing detection

The proposed SOC-based phishing detection framework integrates phishing simulation outputs, email header forensics, and URL intelligence into a unified decision-support pipeline. The objective is to compute a composite phishing risk score for each email event and forward high-risk cases to the SOC dashboard for alerting and response. The overall framework, shown in Figure 3, enables centralized monitoring, cross-layer correlation, and timely mitigation of phishing attempts.

Let an incoming email instance be represented as

$$E_i = \{H_i, U_i, B_i\} \tag{1}$$

In (1), H_i denotes header-derived features, U_i denotes URL-derived features, and B_i denotes user-behavior or simulation-derived indicators associated with the ith event.

3.2.1. Modelling of header risk

The header risk score is computed from authentication and routing anomalies as written as (2).

$$R_H(i) = w_{spf}x_{spf} + w_{dkim}x_{dkim} + w_{dmarc}x_{dmarc} + w_{ip}x_{ip} + w_{dom}x_{dom} \tag{2}$$

In (2):
- $x_{spf} \in \{0, 1\}$ indicates SPF failure,
- $x_{dkim} \in \{0, 1\}$ indicates DKIM failure,
- $x_{dmarc} \in \{0, 1\}$ indicates DMARC failure,
- $x_{ip} \in \{0, 1\}$ indicates suspicious sender IP mismatch,
- $x_{dom} \in \{0, 1\}$ indicates suspicious or recently registered sender domain.

The weights satisfy the (3).

$$w_{spf} + w_{dkim} + w_{dmarc} + w_{ip} + w_{dom} = 1. \tag{3}$$

3.2.2. Modelling of URL risk

The URL risk score is defined as (4):

$$R_U(i) = w_{rep}x_{rep} + w_{age}x_{age} + w_{redir}x_{redir} + w_{short}x_{short} + w_{https}x_{https} \tag{4}$$

In (4):
- x_{rep} indicates poor reputation or blacklist presence,
- x_{age} indicates newly registered domain,
- x_{redir} indicates suspicious redirection chain,
- x_{short} indicates URL shortening or obfuscation,
- x_{https} indicates deceptive HTTPS usage.

Similarly (5),

$$w_{rep} + w_{age} + w_{redir} + w_{short} + w_{https} = 1. \tag{5}$$

3.2.3. Behavioral and simulation risk modelling

The phishing simulation component captures behavioral susceptibility through user interaction features. The behavioral risk score is given by (6).

$$R_B(i) = w_{click}x_{click} + w_{cred}x_{cred} + w_{dept}x_{dept} \quad (6)$$

In (6):

- $x_{click} \in \{0, 1\}$ indicates whether the user clicked the phishing link,
- $x_{cred} \in \{0, 1\}$ indicates whether credentials were submitted,
- $x_{dept} \in [0, 1]$ represents normalized departmental susceptibility.

The associated weights satisfy (7),

$$w_{click} + w_{cred} + w_{dept} = 1. \quad (7)$$

3.2.4. Composite phishing risk score

The final phishing risk score for email instance E_i is obtained by fusing the three risk components written as (8).

$$R_T(i) = \alpha R_H(i) + \beta R_U(i) + \gamma R_B(i) \quad (8)$$

subject to (9),

$$\alpha + \beta + \gamma = 1, \qquad \alpha, \beta, \gamma \in [0, 1]. \quad (9)$$

A decision threshold τ is then used for alert generation is given by (10).

$$D(i) = \begin{cases} 1, & \text{if } R_T(i) \geq \tau \\ 0, & \text{if } R_T(i) < \tau \end{cases} \quad (10)$$

In (10), $D(i) = 1$ indicates that the event is flagged as phishing and forwarded to the SOC dashboard.

3.2.5. Correlation strength

To quantify cross-layer consistency between header anomalies and malicious URLs, the conditional correlation strength can be expressed as (11).

$$C_{HU} = P(U_m \mid H_a) \quad (11)$$

In (11), U_m denotes malicious URL presence and H_a denotes header anomaly. A high value of C_{HU} suggests that suspicious headers strongly coincide with malicious URL activity.

Likewise, the association between authentication failure and phishing likelihood may be expressed as (12).

$$C_{AU} = P(U_m \mid x_{spf} = 1, x_{dkim} = 1, x_{dmarc} = 1). \quad (12)$$

3.2.6. Prioritization of SOC alert

For practical incident handling, the total risk score may be categorized into three alert levels are given by (13).

$$L(i) = \begin{cases} \text{Low}, & 0 \leq R_T(i) < \tau_1 \\ \text{Medium}, & \tau_1 \leq R_T(i) < \tau_2 \\ \text{High}, & R_T(i) \geq \tau_2 \end{cases} \quad (13)$$

In (13), τ_1 and τ_2 are lower and upper operational thresholds selected by the SOC team.

The complete detection and prioritization procedure is summarized in Algorithm 1.

Algorithm 1 SOC-based phishing detection.

Require: Incoming email event E_i
Ensure: Risk score and alert status

```
Extract header features H_i
Extract URL features U_i
Extract behavioral features B_i
Compute header risk score R_H(i)
Compute URL risk score R_U(i)
Compute behavioral risk score R_B(i)
Compute total risk score:
    R_T(i) = αR_H(i) + βR_U(i) + γR_B(i)
if R_T(i) ≥ τ then
    Flag E_i as phishing
    Forward alert to SOC dashboard
else
    Mark E_i as low-risk
end if
return R_T(i) and alert status
```

4. Experimental procedure

To systematically investigate phishing attack behavior and evaluate detection mechanisms, a controlled experimental framework was developed integrating phishing simulation, email header forensics, and URL intelligence analysis. The experimental workflow emulates real-world adversarial scenarios while ensuring safe and ethical execution within an isolated organizational testbed.

The overall methodology follows a three-stage pipeline: (i) phishing campaign simulation, (ii) email header analysis, and (iii) URL threat intelligence evaluation. This structured workflow enables end-to-end assessment of phishing attacks from user interaction to infrastructure-level detection.

4.1. Simulation of Phishing campaign

The phishing simulation phase was designed to replicate real-world adversarial strategies to assess user susceptibility and organizational preparedness. Two complementary tools, GoPhish and RubikPhish, were employed to provide both manual control and automated large-scale testing capabilities.

4.1.1. Phishing Email generation using GoPhish

GoPhish was utilized as the primary platform for crafting and deploying phishing campaigns. The configuration process involved SMTP server setup, template customization, and real-time tracking of user interactions.

- Configured sender identities, subject lines, and email templates to closely mimic legitimate organizational communication patterns.
- Embedded phishing URLs generated through controlled infrastructure.
- Executed campaigns targeting multiple user groups across departments.
- Monitored key interaction metrics including email opens, link clicks, and credential submissions.

Table 1: Comparative analysis and justification of selected cybersecurity tools.

Tool category	Tools compared	Key features	Justification for selection
Phishing simulation	GoPhish vs. PhishMe	Open-source platform with customizable phishing campaigns and SMTP integration	Provides flexibility, extensibility, and cost-effective deployment compared to proprietary PhishMe
Phishing automation	RubikPhish vs. SET toolkit	Automated phishing workflows with integrated user awareness feedback mechanisms	Enhances user awareness and training; SET Toolkit is primarily designed for penetration testing
Email header analysis	MXToolBox vs. EmailHarvester	SPF, DKIM, DMARC validation, DNS diagnostics, and blacklist lookup tools	Offers comprehensive email authentication and diagnostic capabilities beyond basic email extraction
Blacklist verification	MultiRBL.Valli vs. Spamhaus	Aggregated DNSBL queries across multiple blacklist databases	Enables broader detection coverage through consolidation of multiple blacklist sources
URL Reputation analysis	URLVoid vs. Google safe browsing	Multi-engine scanning, domain intelligence, and blacklist aggregation	Provides richer multi-source threat intelligence compared to single-engine Safe Browsing
Phishing intelligence platforms	PhishTank vs. OpenPhish	Community-driven phishing database with URL verification and reporting mechanisms	Supports collaborative validation and continuous updates, improving detection reliability
Malware and URL analysis	VirusTotal vs. Hybrid analysis	Multi-engine antivirus scanning versus behavioral sandbox-based analysis	Enables rapid detection using multiple engines, while hybrid analysis provides deeper behavioral insights

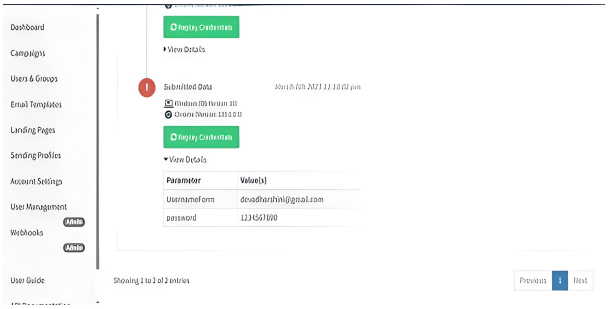

Figure 4: Phishing email template generated using GoPhish.

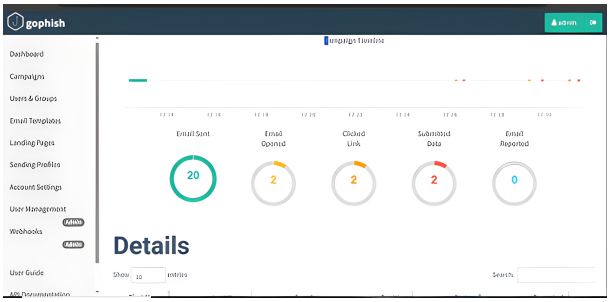

Figure 5: GoPhish campaign monitoring dashboard.

Figure 4 illustrates a representative phishing email template designed to exploit authority and urgency cues. The campaign monitoring interface, shown in Figure 5, provides real-time analytics of user interactions.

4.1.2. Automated Phishing Simulation using RubikPhish

RubikPhish complemented GoPhish by enabling automated phishing scenario generation and scalable awareness testing. This tool was particularly effective for evaluating behavioral responses across diverse user groups.

- Designed automated phishing scenarios tailored to department specific roles.

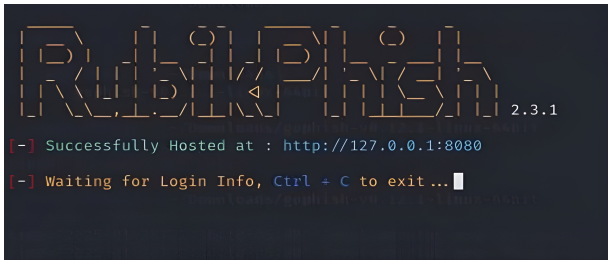

Figure 6: Automated phishing testing using RubikPhish.

- Generated attack templates based on real-world phishing vectors.
- Integrated feedback mechanisms to improve user awareness post-simulation.
- Produced detailed reports capturing behavioral trends and susceptibility patterns.

The automation capability significantly enhanced experimental coverage and reproducibility. Figure 6 demonstrates the automated phishing testing environment.

The outputs of the phishing campaigns, including captured emails and user interaction logs, were subsequently subjected to forensic analysis through email header inspection.

4.2. Email header analysis

Email header forensics was conducted to identify structural and authentication anomalies associated with phishing attacks. This stage focuses on validating sender authenticity and detecting inconsistencies in message routing.

4.2.1. Header analysis using MXToolBox

MXToolBox was employed to perform detailed inspection of email headers.

- Extracted complete header metadata from suspected phishing emails.

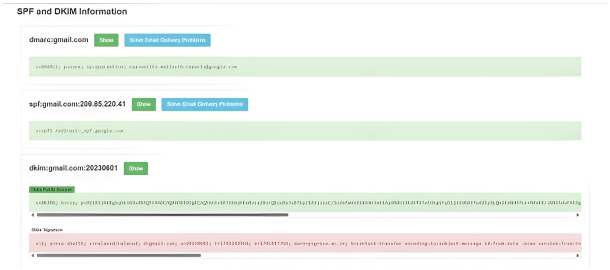

Figure 7: DKIM and domain key verification.

- Analyzed routing paths, sender IP addresses, and mail transfer agents.
- Verified authentication mechanisms including SPF, DKIM, and DMARC.
- Identified inconsistencies such as spoofed domains and anomalous routing behavior.

The verification process revealed multiple indicators of compromise. Figure 7 illustrates the DKIM and domain key validation results.

4.2.2. Blacklist verification using MultiRBL

To further validate email source legitimacy, MultiRBL was used for blacklist-based threat intelligence.

- Extracted sender IP addresses and domains from email headers.
- Queried multiple DNS-based blacklists.
- Analyzed blacklist hits and severity classifications.
- Correlated blacklist presence with phishing likelihood.

This step enabled identification of known malicious infrastructure and improved detection confidence.

Following header-level validation, embedded URLs were analyzed to detect adversarial infrastructure patterns and evasion techniques.

4.3. URL Analysis

The URL analysis phase focused on detecting malicious links embedded within phishing emails through reputation analysis and multi-engine scanning.

4.3.1. Phishing URL verification using PhishTank

PhishTank was utilized as a community-driven repository for validating phishing URLs.

- Extracted URLs from phishing emails and campaigns.
- Cross-referenced URLs with the PhishTank database.
- Identified known phishing links and reported new malicious URLs.
- Evaluated the coverage of community-driven threat intelligence.

A significant proportion of URLs matched known phishing entries, highlighting the effectiveness of collaborative intelligence.

4.3.2. URL reputation analysis using URLVoid

URLVoid was employed for comprehensive domain and reputation analysis.

- Evaluated domain age, hosting infrastructure, and registrar information.

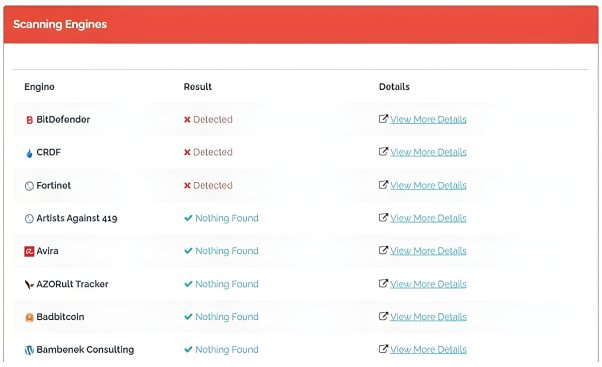

Figure 8: URL reputation and threat indicators.

- Assessed blacklist presence and multi-engine threat indicators.
- Identified suspicious patterns such as newly registered domains and anomalous hosting behavior.

Figure 8 shows representative URL reputation analysis and threat indicators.

4.3.3. Malware and URL scanning using VirusTotal

VirusTotal was used to perform multi-engine scanning of URLs and attachments.

- Submitted URLs and files to multiple antivirus engines.
- Examined detection signatures and behavioral indicators.
- Identified malware payloads and credential harvesting mechanisms.

This multi-layer validation significantly improved detection reliability and reduced false positives, providing a robust foundation for subsequent correlation analysis.

5. Results and analysis

5.1. Results of Phishing simulation

The conducted phishing simulation campaigns provided quantitative insights into organizational susceptibility across different functional departments. The overall click-through rate was observed to be 15%, while the credential submission rate reached 10%, indicating a non-trivial level of user vulnerability to socially engineered attacks.

Department-wise analysis (Table 2) reveals that customer service (32%) and sales (28%) exhibited the highest click rates, reflecting increased exposure to external communications and customer interactions. In contrast, finance (8%) and IT (12%) demonstrated comparatively lower susceptibility, likely due to stronger security awareness and stricter operational controls.

The effectiveness of phishing templates was also evaluated, where IT notification-based emails achieved the highest success rate (37%), emphasizing the exploitation of authority and urgency in phishing design.

Figure 9 further illustrates the comparative click-rate distribution across departments, clearly highlighting the disparity in user behavior and risk exposure.

Table 2: Department-wise phishing susceptibility analysis.

Department	Click rate (%)	Credential submission (%)	Risk level
Finance	8	4	Low
IT	12	6	Low
Sales	28	18	High
Customer service	32	21	Very high
HR	18	11	Medium
Operations	16	9	Medium

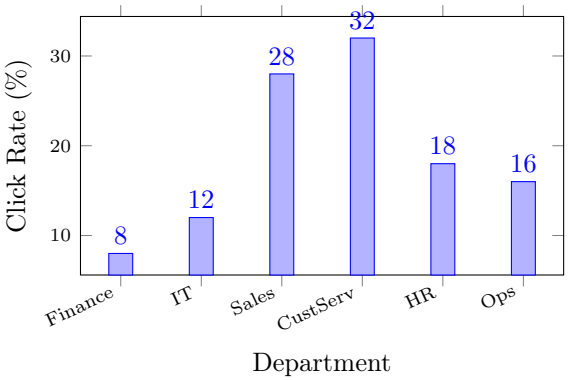

Figure 9: Department-wise phishing susceptibility.

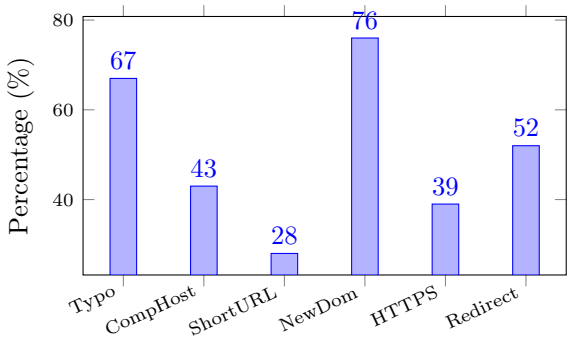

Figure 10: URL-based phishing indicators.

Table 3: Metrics of email authentication failure.

Metric	Percentage (%)
SPF Failure	81
DKIM Failure	73
DMARC Failure	89
Forged sender address	68
IP Mismatch	42

5.2. Findings of Email header analysis

Email header forensics revealed multiple anomalies strongly associated with phishing attacks. As summarized in Table 3, DMARC failures (89%), SPF failures (81%), and DKIM failures (73%) were the most prominent indicators, confirming the widespread use of spoofed or unauthenticated email infrastructures.

Additionally, 68% of analyzed emails contained forged sender addresses, while 42% exhibited IP location mismatches, indicating inconsistencies between claimed and actual sender origins. The presence of newly registered domains (37%) further reinforces the temporal characteristics of phishing campaigns.

These findings highlight that authentication protocol failures serve as reliable indicators for early phishing detection and can be effectively leveraged in automated filtering systems.

5.3. Results of URL analysis

The URL inspection phase revealed distinct adversarial infrastructure patterns commonly employed in phishing campaigns. As depicted in Figure 10, newly registered domains (14 days) accounted for 76% of malicious URLs, making them the most dominant indicator.

Typosquatting attacks (67%) were also prevalent, exploiting user trust through visually similar domain names. Furthermore, redirect chains (52%) and compromised hosting services (43%) indicate the use of multi-stage obfuscation techniques to evade detection.

The misuse of HTTPS (39%) demonstrates that attackers increasingly leverage secure protocols to gain user trust, while URL obfuscation (28%) techniques further complicate detection mechanisms.

5.4. Correlation analysis

Cross-domain correlation analysis reveals strong interdependencies between phishing indicators. Table 4 summarizes these relationships.

Notably, emails with forged headers were found to increase the likelihood of malicious URLs by approximately 3.2×. The presence of newly registered domains significantly amplifies phishing probability by 5.7×, indicating the importance of domain age as a predictive feature.

Moreover, simultaneous failure of SPF, DKIM, and DMARC resulted in malicious URL presence in 91% of cases, highlighting a strong combined effect of authentication anomalies.

These findings confirm that phishing attacks are inherently multi-layered, often combining spoofed headers with obfuscated URL structures to maximize success rates while evading detection systems.

5.5. Key contributions

- Unified phishing detection framework integrating simulation, header, and URL analysis
- Empirical organizational vulnerability assessment
- Strong correlation modeling (91% authentication failure linkage)

Table 4: Correlation between phishing indicators.

Indicator relationship	Strength
Forged header → Malicious URL	3.2×
New domain → Phishing content	5.7×
Auth failure → Malicious URL	91%

- Real-time SOC deployable pipeline

6. Conclusion

This study demonstrates that isolated phishing detection techniques are insufficient against modern cyber threats. A multi-layered analytical approach integrating simulation, email forensics, and URL intelligence significantly enhances detection accuracy and resilience. The strong correlation between authentication failures and malicious payload delivery provides a robust foundation for automated threat detection systems. Future work will explore AI-driven behavioral analytics and multi-channel phishing detection to address the evolving threat landscape.

Declarations and ethical statements

Conflict of interest: The author declare that there is no conflict of interest.
Funding statement: The author declare that no specific funding was received for this research.
Artificial Intelligence usage statement: During the preparation of this manuscript, the author utilized Grammarly tool solely for language refinement and grammatical corrections. The author carefully reviewed and revised the generated content and take full responsibility for the accuracy, integrity, and originality of the final manuscript.
Availability of data and materials: The data and/or materials that support the findings of this study are available from the corresponding author upon reasonable request.

CRediT authorship contribution statement

Sabitha Banu: Conceptualization, Investigation, Writing – review & editing, Validation. **R Divya**: Data curation. **Deva Dharshini TT**: Data curation & Visualization. **Bhoovana Sri**: Data curation & Visualization. **Mehdi Gheisari**: Conceptualization, Writing – review & editing, Validation. **Saman Khammar**: Analysis & Validation. **Mustafa Ghaderzadeh**: Analysis & Validation. **Saeed Lotfi**: Analysis & Validation.

Publisher's note

Krrish Scientific Publications Pvt. Ltd. and the ***Journal of Computing and Data Technology*** remain neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

[1] Tanti R. Study of Phishing Attack and their Prevention Techniques. *International Journal of Scientific Research in Engineering and Management*. 2024 Oct;8(10):1-8. Available from: https://doi.org/10.55041/IJSREM38042

[2] kumar Soma A. Hybrid RNN-GRU-LSTM Model for Accurate Detection of DDoS Attacks on IDS Dataset. *Journal of Modern Technology*. 2025 May 14;2(01):283-91. Available from: https://doi.org/10.71426/jmt.v2.i1.pp283-291

[3] Bossetta M. The Weaponization of Social Media: Spear Phishing and Cyberattacks on Democracy. *Journal of international affairs*. 2018 Jan 1;71(1.5):97-106. Available from: https://www.jstor.org/stable/26508123

[4] Khamis SA, Foozy CF, Aziz MF, Rahim N. Header Based Email Spam Detection Framework Using Support Vector Machine (SVM) Technique. In *International conference on soft computing and data mining*. 2019 Dec 5 (pp. 57-65). Cham: Springer International Publishing. Available from: https://doi.org/10.1007/978-3-030-36056-6_6

[5] Asif AU, Shirazi H, Ray I. Machine learning-based phishing detection using URL features: A comprehensive review. In *International Symposium on Stabilizing, Safety, and Security of Distributed Systems*. 2023 Sep 30 (pp. 481-497). Cham: Springer Nature Switzerland. Available from: https://doi.org/10.1007/978-3-031-44274-2_36

[6] Chanyour T, El Kasmi Alaoui S, Kaddari A, Hmimz Y, Chiba Z. Blockchain and Reputation Based Secure Service Provision in Edge-Cloud Environments. In *International Conference on Artificial Intelligence and Smart Environment*. 2023 Nov 23 (pp. 15-20). Cham: Springer Nature Switzerland. Available from: https://doi.org/10.1007/978-3-031-48573-2_3

[7] Kumaraguru P, Rhee Y, Acquisti A, Cranor LF, Hong J, Nunge E. Protecting people from phishing: the design and evaluation of an embedded training email system. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 2007 Apr 29 (pp. 905-914). Available from: https://dl.acm.org/doi/abs/10.1145/1240624.1240760

[8] Lallie HS, Shepherd LA, Nurse JR, Erola A, Epiphaniou G, Maple C, Bellekens X. Cybersecurity in the age of COVID-19: A timeline and analysis of cyber-crime and cyber-attacks during the pandemic. *Computers & security*. 2021 Jun 1;105:102248. Available from: https://doi.org/10.1016/j.cose.2021.102248

[9] Kumar N, Goel V, Ranjan R, Altuwairiqi M, Alyami H, Asakipaam SA. A Blockchain-Oriented Framework for Cloud-Assisted System to Countermeasure Phishing for Establishing Secure Smart City. *Security and Communication Networks*. 2023;2023(1):8168075. Available from: https://doi.org/10.1155/2023/8168075

[10] Yadav A, Kumar A, Singh V. Open-source intelligence: a comprehensive review of the current state, applications and future perspectives in cyber security. *Artificial Intelligence Review*. 2023 Nov;56(11):12407-38. Available from: https://doi.org/10.1007/s10462-023-10454-y

[11] Halim MI, Hasan MZ, Kabir MH, Hasan MN, Jaki H, Ahmad, Hasan H. Enhancing Phishing Detection: A Machine Learning Approach to Predicting Malicious Emails, URLs, and SMS Messages. *Applied Computational Intelligence and Soft Computing*. 2025;2025(1):6633979. Available from: https://doi.org/10.1155/acis/6633979

[12] Ramanathan V, Wechsler H. Phishing detection and impersonated entity discovery using Conditional Random Field and Latent Dirichlet Allocation. *Computers & Security*. 2013 May 1;34:123-39. Available from: https://doi.org/10.1016/j.cose.2012.12.002

[13] Gupta I, Singh N, Singh AK. Layer-based privacy and security architecture for cloud data sharing. *Journal of Communications Software and Systems*. 2019 Jun 1;15(2):173-85. Available from: https://doi.org/10.24138/jcomss.v15i2.617

[14] Kritika E. A comprehensive literature review on phishing URL detection using deep learning techniques. *Journal of Cyber*

Security Technology. 2025 Oct 2;9(4):315-43. Available from: https://doi.org/10.1080/23742917.2024.2378552

[15] Jain A. Enhancing forensic analysis of digital evidence using machine learning: Techniques, applications, and challenges. *Int. J. Innov. Res. Multidiscip. Perspect. Stud.(IJIRMPS)*. 2024 Sep;18:1-8. Available from: https://www.researchgate.net/profile/Pankaj-Malik-4/publication/383870594

[16] Paul E, Callistus O, Somtobe O, Esther T, Somto K, Clement O, Ejimofor I. Cybersecurity strategies for safeguarding customer's data and preventing financial fraud in the United States financial sectors. *International Journal on Soft Computing*. 2023 Aug;14(3):01-16. Available from: https://doi.org/10.5121/ijsc.2023.14301

[17] Shaukat K, Luo S, Varadharajan V, Hameed IA, Xu M. A survey on machine learning techniques for cyber security in the last decade. *IEEE Access*. 2020 Dec 2;8:222310-54. Available from: https://doi.org/10.1109/ACCESS.2020.3041951

[18] Chatchalermpun S, Daengsi T. Improving cybersecurity awareness using phishing attack simulation. In *IOP conference series: materials science and engineering* 2021 Feb 1 (Vol. 1088, No. 1, p. 012015). IOP Publishing. Available from: 10.1088/1757-899X/1088/1/012015

[19] Jartelius M. The 2020 data breach investigations report–a CSO's perspective. *Network Security*. 2020 Jul;2020(7):9-12. Available from: https://doi.org/10.1016/S1353-4858(20)30079-9

[20] Downs JS, Holbrook M, Cranor LF. Behavioral response to phishing risk. In *Proceedings of the anti-phishing working groups 2nd annual eCrime researchers summit*. 2007 Oct 4 (pp. 37-44). Available from: https://doi.org/10.1145/1299015.1299019

[21] Kheruddin MS, Zuber MA, Radzai MM. Phishing attacks: Unraveling tactics, threats, and defenses in the cybersecurity landscape. *Authorea Preprints*. 2024 Jan 15. Available from: https://www.authorea.com/doi/full/10.22541/au.170534654.48067877

[22] Cheng TH, Lin YD, Lai YC, Lin PC. Evasion techniques: Sneaking through your intrusion detection/prevention systems. *IEEE Communications Surveys & Tutorials*. 2011 Oct 13;14(4):1011-20. Available from: https://doi.org/10.1109/SURV.2011.092311.00082

[23] Penaganti R. Graph Neural Network-Based Framework for Real-Time Financial Fraud Detection in Digital Payment Ecosystems. *Journal of Computing and Data Technology*. 2025 Nov. 21 ;1(2):91-97.Available from: https://doi.org/10.71426/jcdt.v1.i2.pp91-97

[24] Lamina OA, Ayuba WA, Adebiyi OE, Michael GE, Samuel OO, Samuel KO. AI- Powered Phishing Detection and Prevention. *Path of Science*. 2024 Dec 31;10(12):4001-10. Available from: http://dx.doi.org/10.22178/pos.112-7

[25] Yang J, Fang B, Lu H, Tian Z. Context-Aware Phishing-Resistant Authentication for Federated Identity in Internet of Things Platforms. *IEEE Internet of Things Journal*. 2024 Dec 11;12(8):11121-34. Available from: https://doi.org/10.1109/JIOT.2024.3515079

[26] Langford T, Payne B. Phishing faster: Implementing chatgpt into phishing campaigns. In *Proceedings of the Future Technologies Conference*. 2023 Oct 19 (pp. 174-187). Cham: Springer Nature Switzerland. Available from: https://doi.org/10.1007/978-3-031-47454-5_13

[27] Wyss E, Davidson D, De Carli L. What's in a URL? An Analysis of Hardcoded URLs in npm Packages. In *Proceedings of the 2024 Workshop on Software Supply Chain Offensive Research and Ecosystem Defenses*. 2023 Nov 19 (pp. 26-32). Available from: https://doi.org/10.1145/3689944.3696168

[28] Sahay R, Meng W, Li W. A Comparative Analysis of Phishing Tools: Features and Countermeasures. In *International Conference on Information Security Practice and Experience*. 2024 Oct 25 (pp. 365-382). Singapore: Springer Nature Singapore. Available from: https://doi.org/10.1007/978-981-97-9053-1_21

[29] Bell S, Komisarczuk P. An analysis of phishing blacklists: Google safe browsing, openphish, and phishtank. In *Proceedings of the Australasian Computer Science Week Multiconference*. 2020 Feb 4 (pp. 1-11). Available from: https://dl.acm.org/doi/10.1145/3373017.3373020

[30] Vidyakeerthi S, Nabeel M, Elvitigala C, Keppitiyagama C. Phishchain: a decentralized and transparent system to blacklist phishing urls. In *Companion proceedings of the web conference 2022*. 2022 Apr 25 (pp. 286-289). Available from: https://doi.org/10.1145/3487553.3524235

[31] Ashraf QM, Habaebi MH. Autonomic schemes for threat mitigation in Internet of Things. *Journal of Network and Computer Applications*. 2015 Mar 1;49:112-27. Available from: https://doi.org/10.1016/j.jnca.2014.11.011

[32] I. H. Sarker *et al.*, "Sarker IH, Kayes AS, Badsha S, Alqahtani H, Watters P, Ng A. Cybersecurity data science: an overview from machine learning perspective. *Journal of Big data*. 2020 Jul 1;7(1):41. Available from: https://doi.org/10.1186/s40537-020-00318-5

[33] Zwilling M, Klien G, Lesjak D, Wiechetek Ł, Cetin F, Basim HN. Cyber security awareness, knowledge and behavior: A comparative study. *Journal of Computer Information Systems*. 2022 Jan 2;62(1):82-97. Available from: https://doi.org/10.1080/08874417.2020.1712269

[34] Quayyum F, Cruzes DS, Jaccheri L. Cybersecurity awareness for children: A systematic literature review. International Journal of Child-Computer Interaction. 2021 Dec 1;30:100343. Available from: https://doi.org/10.1016/j.ijcci.2021.100343

[35] Alharbi T, Tassaddiq A. Assessment of cybersecurity awareness among students of Majmaah University. *Big Data and Cognitive Computing*. 2021 May 10;5(2):23. Available from: https://doi.org/10.3390/bdcc5020023

[36] Bongu SR. Real-Time Behavioral Biometrics and Continuous Authentication Framework for Secure Financial Transaction Ecosystems. *Journal of Applied Sciences and Modelling*. 2025 Dec. 31;1(1):40-5. Available from: https://doi.org/10.71426/jasm.v1.i1.pp40-50