**International Innovative Research Journal of Engineering and Technology**
Vol. 10, Issue. 1, Sep 2024, pp. 1~8
ISSN: 2456-1983, DOI: 10.32595/iirjet.org/v10i1.2024.196 

# A Comprehensive Web-Based Platform for Phishing Message Detection

**Suhaib Abdullah Hamed Albishri[1], Tasnim Hamed Fadhl Alkhaziri[2], Malath Salim Khalfan AlYahmadi[3], Dr. Bazeer Ahmed Bagrudeen[4]**
[1,2,3,4]University of Technology and Applied Sciences, College of Computing and Information Sciences, Al Mussanah, Sultanate of Oman

**Article Info**

***Article history:***

Received Jun 4, 2024
Revised Jul 15, 2024
Accepted Aug 28, 2024

***Keywords:***

Phishing Detection
Real-Time Analysis
Phishing Attacks
URL Analysis
Internet Security

**ABSTRACT**

The persistent threat of phishing attacks looms large over the digital sphere, posing indiscriminate risks to users and organizations alike, leading to substantial financial losses and data breaches. In response to this growing menace, this paper introduces an innovative web platform engineered specifically to combat the escalating sophistication of phishing messages. At the heart of its design lies the fusion of advanced machine learning algorithms, natural language processing techniques, and behavioural analysis, facilitating the real-time identification and mitigation of phishing attempts. Distinguishing itself from prior solutions, our platform embraces a multifaceted strategy, amalgamating URL analysis, email content scrutiny, and user behaviour tracking to pinpoint malicious messages accurately, thereby fortifying cybersecurity resilience. Through meticulous testing and validation, our platform demonstrates exceptional effectiveness in thwarting phishing assaults, surpassing prevailing methodologies by a significant margin. Featuring an intuitive user interface, the platform equips both individuals and enterprises with the tools to swiftly assess the legitimacy of incoming communications, thereby reducing susceptibility to phishing scams. Our methodology represents a notable departure from traditional approaches, presenting a comprehensive solution adept at countering the evolving tactics employed by cyber adversaries. In sum, this web platform emerges as a pivotal advancement in the battle against phishing threats, fortifying cybersecurity defences in an increasingly precarious online environment.

***Corresponding Author:***

Suhaib Abdullah Hamed Albishri,
University of Technology and Applied Sciences
College of Computing and Information Sciences – Al Mussanah, Sultanate of Oman.
Email: 54S1878@act.edu.om

## 1. INTRODUCTION

Despite advancements in cybersecurity, phishing remains a prevalent threat. Traditional methods of phishing detection may not always suffice, necessitating the development of a sophisticated web platform tailored

to address evolving phishing techniques [1]. In recent years, phishing attacks have grown to be a serious threat to both individuals and organizations. In the first quarter of 2021, there were 65% more phishing attacks than there were during the same period the previous year, according to a report by the Anti-Phishing Working Group. The need for efficient methods to identify and stop phishing attacks is highlighted by the rise in attacks. Phishing attacks entail deceiving users into divulging their personal information by means of phony emails, messages, or websites [2]. Financial losses, identity theft, and harm to an organization's reputation are all possible outcomes of these attacks. The sophistication of phishing attacks has increased, making it more challenging for users to discern between authentic and fraudulent messages [3].

## 2. MOTIVATION AND PROBLEM STATEMENT

The problem we are attempting to address is the rise in phishing attacks, which have the potential to cause people's personal information to be stolen [4]. Because phishing attacks can cause large financial losses, harm to a person's reputation, and a decline in confidence in online systems, it is crucial to find a solution to these problems. Therefore, there are some solutions to these problems [5]: develop a comprehensive web platform, focus on detecting phishing messages, utilize advanced machine learning algorithms, implement natural language processing techniques, cover various communication channels (email, messaging apps, social media), ensure accuracy in identifying phishing attempts, provide real-time alerts to users, offer actionable steps to mitigate potential risks, and enhance cybersecurity for individuals and organizations[6]. The main result includes a thorough examination of phishing attacks, how they are exploited.

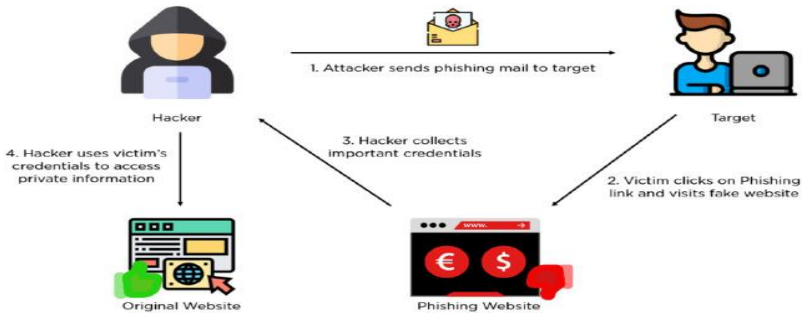


Figure 1. Phishing using Email

### 2.1 Specific issue

Phishing attacks have become more frequent and sophisticated. Phishing attacks are misleading communications that try to deceive people into providing sensitive information [7]. Many people are still susceptible to these risks in spite of greater knowledge and security precautions, which can result in large-scale data and financial losses. This study specifically addresses the shortcoming of existing detection techniques in real-time phishing message identification and intelligibility provision for users [19].

### 2.2 Rising Incidence of Phishing

Phishing is one of the most popular forms of cyberattack, affecting both people and companies, according to cybersecurity reports [8]. The rise in digital communication and distant employment has made people more vulnerable to fake mails, aggravating the issue [18].

### 2.3 Economic Impact

Every year, phishing assaults cause losses in the billions of dollars. In addition to financial consequences, organizations risk losing the trust of their customers, harming their brand, and facing regulatory fines [9].

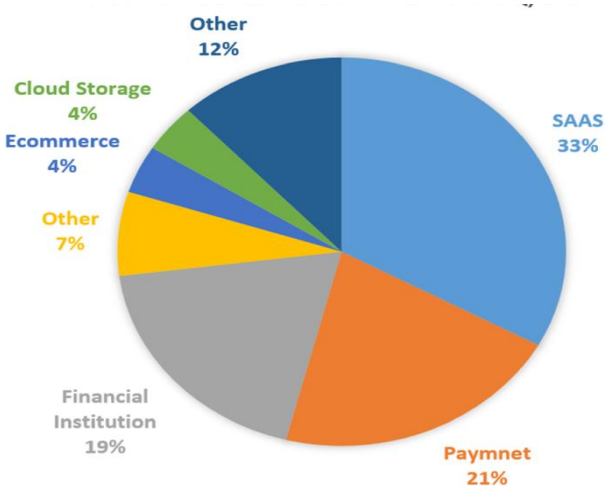


Figure 2. Most targeted industry sectors

**2.4 Psychological Consequences**

After falling for these scams, phishing victims frequently suffer worry and anxiety, which can affect their online behavior and faith in digital platforms [10].

**2.5 Inadequate Current Solutions**

Current detection techniques, including spam filters, are frequently reactive and ill-suited to combat attackers' ever-evolving strategies [11]. Many expose consumers to new phishing techniques because they depend on well-known phishing signatures or heuristics.

**2.6 User Awareness Gap**

The public is very ignorant about phishing techniques [12]. Well-written communications may fool even well-informed users, which emphasizes the need for technologies that not only identify hazards but also instantly inform users of their surroundings [20].

**2.7 Research gap**

Inadequate detection of sophisticated phishing techniques: Sophisticated social engineering techniques are frequently used in advanced phishing attacks, making them more difficult to identify.

**2.8 Limited scope**

Studies have concentrated on particular phishing message types, like emails or messages from social media.

**2.9 Lack of real-time updates**

Many platforms might not have a way to keep up with the most recent phishing threats in their databases, which leaves users open to new and developing attack methods [13].

- To develop an intuitive web platform for detecting phishing messages.
- To implement advanced message analysis techniques, including machine learning models and keyword matching.
- To provide educational resources to enhance user awareness and recognition of phishing threats.
- To establish a user-friendly reporting system for suspected phishing messages.

The rest of the paper is organized as follows: Background and statistics on phishing attacks, an overview of approaches based on visual similarity for phishing detection, a classification of different kinds of phishing detection and screening methods, emphasizing approaches based on visual similarity, metrics for performance and assessment of ant phishing 3systems and unresolved problems and obstacles in phishing detection and defense [17].

## 3. LITERATURE REVIEW

B Ahamed, T Ramkumar - (2016). Published Phishing and Fraudulent Email Detection through Transfer Learning using pretrained transformer models. The result shows that malicious users and cybercriminals use phishing, a sort of social engineering attack, to steal confidential data, install malware, ransomware, and other advanced persistent threats on a victim's computer or mobile device [14]. Phishing attacks are becoming more common due to the increasing use of the Internet, and the recent shift to working from home has made it more likely that both small and large organizations will be the target of these attacks. Phishing scams can have serious repercussions, such as identity theft, financial loss, data loss, and data theft. Although there are numerous ways to conduct phishing attacks, one of the most popular methods used by cybercriminals is sending phishing emails. This is primarily due to the ease of obtaining email addresses and the low cost of sending bulk emails, which allows attackers to send out many emails in the hopes that some users will fall for the scam.

Eduardo Benavides, Walter Fuertes, Sandra Sanchez, Daniel Nunes and German Rodriguez (2023). A Phishing-Attack-Detection Model Using Natural Language Processing and Deep Learning. This study set out that phishing is a type of cyberattack that uses phony websites that look real in an attempt to trick users. Nowadays, one of the most popular methods for identifying these phishing pages based on their content is to feed words into Deep Learning (DL) algorithms in an arbitrary order, or nonsequential. However, this method loses the inherent complexity of the connections between words [15]. This work innovates in the field of cyber-security by putting forth a model that uses Natural Language Processing (NLP) and DL algorithms to identify phishing attacks based on the text of dubious web pages rather than on URL addresses.

Daojing He, Xin Lv, Xueqian Xu, Shui Yu, Dawei Li, Sammy Chan and Mohsen Guizani (2022). An Effective Double-Layer Detection System Against Social Engineering Attacks. This shows that social engineering attacks that target groups of people and use phishing emails as the medium have become more common in recent years. Social engineering attacks can be launched against current enterprise systems. As a result, we suggest a deep learning-based double-layer detection framework. First, from the standpoint of individual security, a phishing email detection model based on Long Short-Term Memory (LSTM) and extreme gradient boosting tree (XGBoost) is created. Next, from the standpoint of group security, an insider threat detection model built on Bidirectional LSTM and Attention mechanism is created. Finally, a social engineering attack and defense simulation platform is developed and integrated with the social engineering network attack simulation theory.

Jay Doshi, Kunal Parmar, Raj Sanghavi, and Narendra Shekokar (2023). A Comprehensive dual-layer Architecture for Phishing and Spam Email Detection. Several studies have been carried out to investigate the effective detection and classification of emails. Abdulraheem et al. (2022) attempted to identify phishing emails by selecting attributes using Principal Component Analysis (PCA), employing Ranker Search as the search strategy, and utilizing three machine learning algorithms: Multilayer Perceptron, Decision Tree (J48, C4.5), and Logistic Model Tree. 96.9245% is the maximum precision that can be achieved with the Logistic Model Tree.

Eman Abdullah, Mohammed Zakariah, Ghada Abdalaziz, Fahdah A and Abdullah I (2023). A Deep Learning-Based Innovative Technique for Phishing Detection in Modern Security with Uniform Resource Locators. This paper presents a novel approach for highly accurate phishing site detection. Our method accurately classifies websites by using a Convolution Neural Network (CNN)-based model, which efficiently separates phishing from legal websites. Thus, they assess the model's effectiveness using the PhishTank dataset, a popular dataset for identifying phishing websites that rely only on Uniform Resource Locators (URL) characteristics. Their method surpasses earlier state-of-the-art models and achieves excellent accuracy rates, making it a unique contribution to the field of phishing detection. The findings of the experiment demonstrated the effectiveness of their suggested strategy in terms of accuracy and false-positive rate.

## 4. PROPOSED WORK

The goal of the proposed study is to provide a thorough online platform for phishing message detection. The platform will assess many kinds of phishing communications, including emails, SMS, and social media messages, by combining machine learning algorithms, natural language processing (NLP), and security-focused methodologies. [16] The technique is described in this part, together with the theoretical foundation, machine learning models, and algorithmic approach.

**Methodology:**
**PhisHunter: Identify Phishing & Scam Messages**

**Input:** User submits a message.
**Algorithm:** Analyzes message for suspicious words and patterns.
**URL Check:** Verifies URLs using **VirusTotal API**.
**IP & Country:** Extracts IP and location info from URLs.
**Sensitive Data Risks:** Flags potential data threats.
**Real-Time Analysis:** Instant feedback on phishing risks.
**Purpose:** Helps users spot and avoid phishing messages.
**Focus:** Detects phishing signs, unsafe URLs, and data risks.

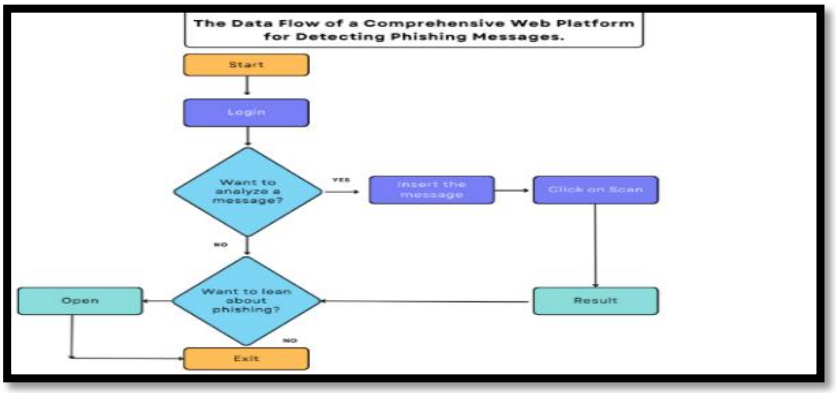


Figure 3. Data flow Diagram of Detecting Phishing Message

The platform's central component is a Phishing Detection Engine, which classifies communications as either phishing or not by using a number of algorithms and techniques.

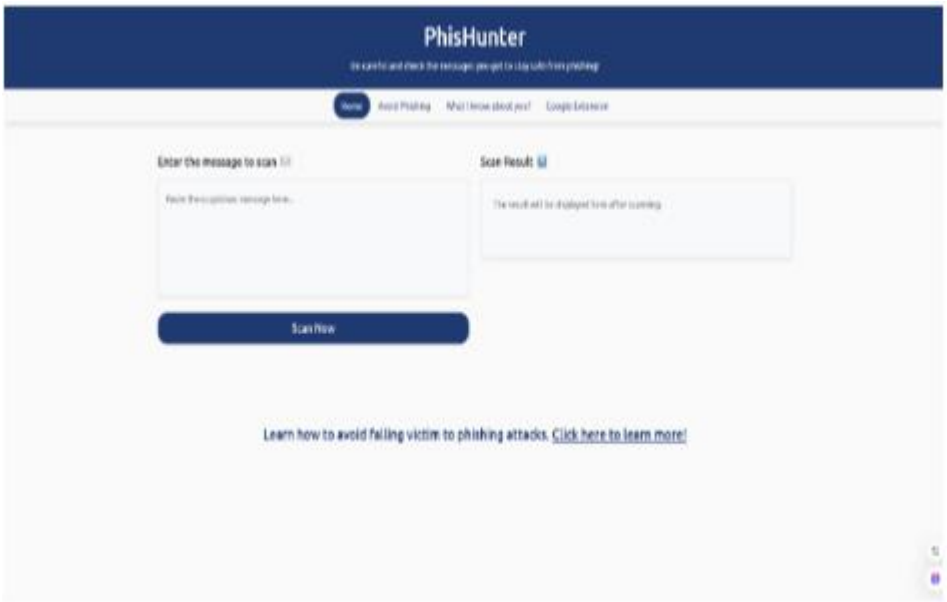


Figure 4. Hunter page

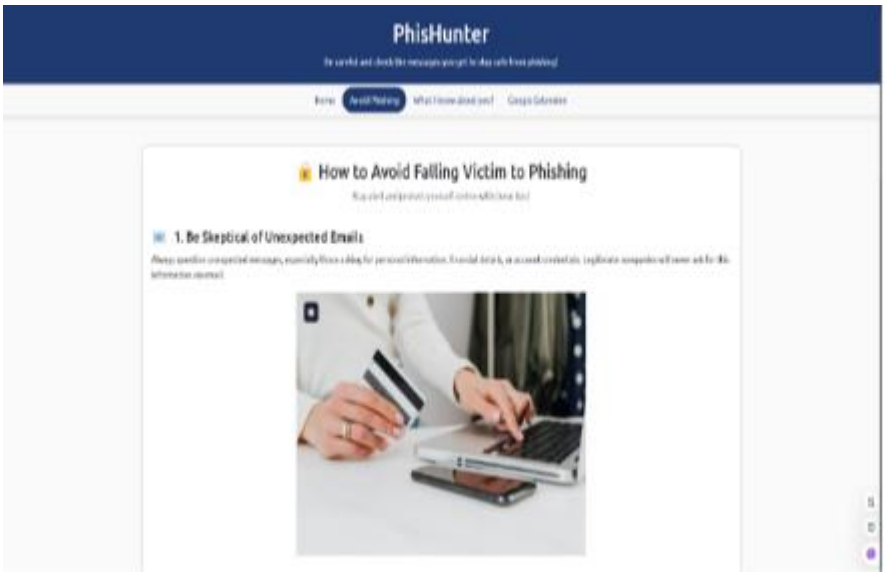


Figure 5. Avoid falling victim to Phishing page

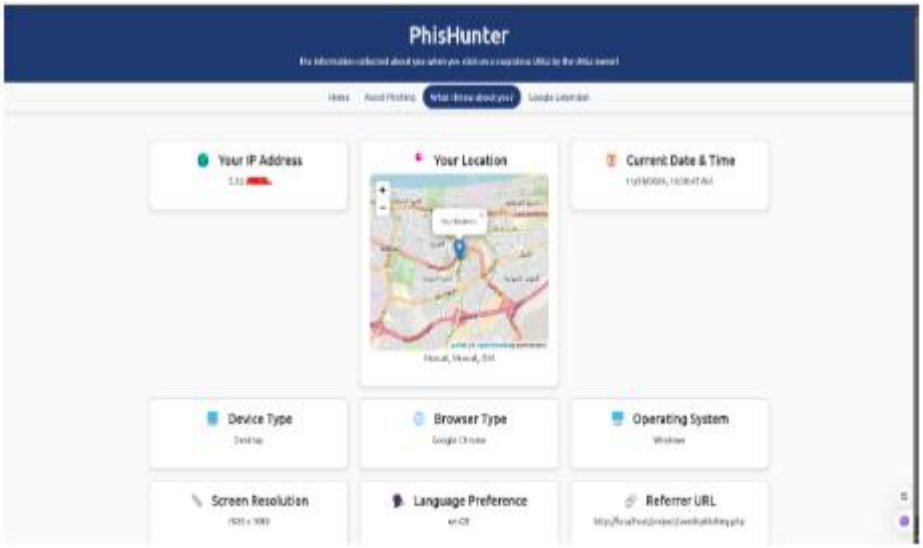


Figure 6. Google extension to extract the URL's

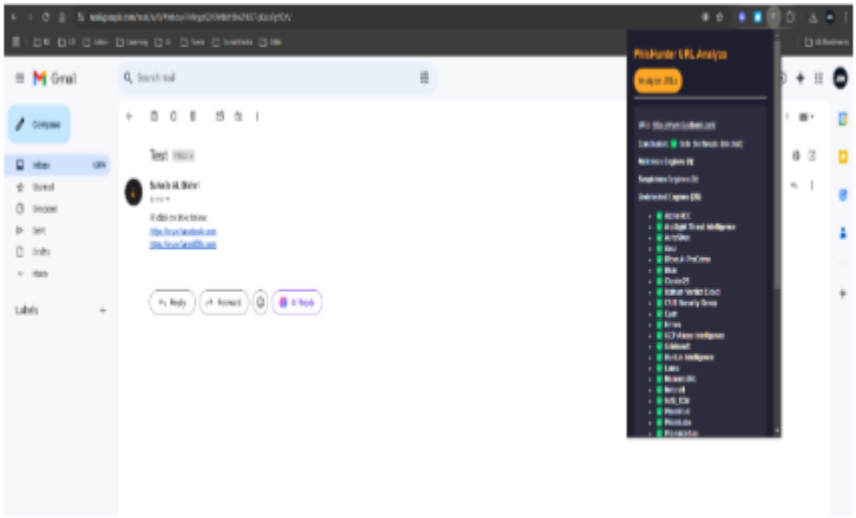

Figure 7. User data in Real-Time page

## 5. CONCLUSION

The study addressed the serious problem of cyberthreats from phishing assaults by presenting a comprehensive online platform for phishing message detection. Phishing, which takes use of human weaknesses to trick users into disclosing private information, is still one of the most common and deadly forms of cybercrime. This technology uses a sophisticated blend of real-time data analysis, machine learning models, and user-friendly interfaces to detect and stop phishing attempts. The main conclusions demonstrate how well the platform uses a range of elements, such as URL analysis, email content analysis, and message behavioral patterns, to effectively identify phishing communications. The software can adjust to new phishing tactics as they appear and constantly increase the accuracy of its detection by integrating machine learning algorithms. Additionally, because the platform is web-based, it is scalable, easily accessible, and compatible with current cybersecurity infrastructures without requiring any changes. By giving people and businesses a strong defensive mechanism against phishing, our study makes a substantial contribution to the world of cybersecurity by delivering an automated and user-friendly solution. It is a priceless tool for lowering the risk of phishing attempts and enhancing overall cybersecurity posture because of its real-time analytical and prediction capabilities. Expanding the platform to include other data sources, such social media and SMS-based phishing, and incorporating more sophisticated AI approaches, like deep learning, for even greater detection capabilities are future goals for this study. Additional study may concentrate on improving the user experience by supporting several languages and making sure the platform can adapt to changing phishing techniques. With the goal of lessening the total impact of phishing assaults on society, the platform has the potential for a wide range of applications in corporate security systems, personal cybersecurity, and even government cybersecurity programs.

## REFERENCES

[1] Basit, A., Zafar, M., Liu, X., Javed, A. R., Jalil, Z., & Kifayat, K. (2021). A comprehensive survey of AI-enabled phishing attacks detection techniques. Telecommunication Systems, 76, 139-154.

[2] Sun, Y., Mi, J., & Jin, C. (2024). Entropy-based concept drift detection in information systems. Knowledge-Based Systems, 290, 111596.

[3] Ibrahim, S., Catal, C., & Kacem, T. (2024). The use of multi-task learning in cybersecurity applications: a systematic literature review. Neural Computing and Applications, 1-27.

[4] Al-Kabbi, H. A., Feizi-Derakhshi, M. R., & Pashazadeh, S. (2024). A Hierarchical Two-Level Feature Fusion Approach for SMS Spam Filtering. Intelligent Automation & Soft Computing, 39(4).

[5] MohamedAli, R. S., & Abduhameed, R. A. (2024, May). Phishing Email Detection: Survey. In International Conference on Advanced Engineering, Technology and Applications (pp. 551-570). Cham: Springer Nature Switzerland.

[6] Kim, J. S., Lee, H. J., Lee, H. J., & Choi, S. H. (2024). Advanced Analysis of Learning-based Spam Email Filtering Methods Based on Feature Distribution Differences of Dataset. IEEE Access.

[7] Jáñez-Martino, F., Alaiz-Rodríguez, R., González-Castro, V., Fidalgo, E., & Alegre, E. (2023). A review of spam email detection: analysis of spammer strategies and the dataset shift problem. Artificial Intelligence Review, 56(2), 1145-1173.

[8] Tusher, E. H., Ismail, M. A., Rahman, M. A., Alenezi, A. H., & Uddin, M. (2024). Email Spam: A Comprehensive Review of Optimize Detection Methods, Challenges, and Open Research Problems. IEEE Access.

[9] Bazeer Ahamed, B., & Krishnamurthy, M. (2022). Evaluation and Customized Support of Dynamic Query Form Through Web Search. In Intelligent Computing & Optimization: Proceedings of the 4th International Conference on Intelligent Computing and Optimization 2021 (ICO2021) 3 (pp. 845-856). Springer International Publishing.
[10] Al-Hamar, Y., Kolivand, H., Tajdini, M., Saba, T., & Ramachandran, V. (2021). Enterprise Credential Spear-phishing attack detection. Computers & Electrical Engineering, 94, 107363.
[11] Ahamed, B. B., & Yuvaraj, D. (2021, December). A Framework for Online Customer Reviews System Using Sentiment Scoring Method. In 2021 22nd International Arab Conference on Information Technology (ACIT) (pp. 1-8). IEEE.
[12] Bazeer Ahamed, B., & Krishnamoorthy, M. (2022, October). Detection and Recovery of Node Failure in Fog-Based WiLD Network for Smart Farming. In International Conference on Intelligent Computing & Optimization (pp. 721-731). Cham: Springer International Publishing.
[13] Bazeer Ahamed, B., & Murugan, K. (2020). A Novel Approach to Overcome Dictionary and Plaintext Attack in SMS Encryption and Decryption Using Vignere Cipher. In Intelligent Computing and Optimization: Proceedings of the 2nd International Conference on Intelligent Computing and Optimization 2019 (ICO 2019) (pp. 559-568). Springer International Publishing.
[14] Ahamed, B. B., & Hariharan, S. (2012). Implementation of network level security process through stepping stones by watermarking methodology. International Journal of Future Generation Communication and Networking, 5(4), 123-130.
[15] Ahamed, B. B., & Krishnamoorthy, M. (2021). Invocation of multi-cloud infrastructure services in web-based semantic discovery system. In Operationalizing Multi-Cloud Environments: Technologies, Tools and Use Cases (pp. 3-18). Cham: Springer International Publishing.
[16] Ahamed, B. B., & Krishnamoorthy, M. (2020). SMS encryption and decryption using modified vigenere cipher algorithm. Journal of the Operations Research Society of China, 1-14.
[17] Ahamed, B. B. (2020). Computational Intelligence in the Reversible Data-Hiding Processes Using Radon Transform. In Handbook of Research on Smart Technology Models for Business and Industry (pp. 414-424). IGI Global.
[18] Kapadiya, K., Ramoliya, F., Gohil, K., Patel, U., Gupta, R., Tanwar, S., ... & Tolba, A. (2025). Blockchain-assisted healthcare insurance fraud detection framework using ensemble learning. Computers and Electrical Engineering, 122, 109898.
[19] Chouhan, N. R., & Dahiya, N. (2024). Design of Machine-Learning Framework for Cyber Attack Detection in Internet of Healthcare Things (IoHT): Critical Analysis. Trends in Mechatronics Systems, 59-71.
[20] Jayadev, G. N. V. R., Neelima, N., Teja, M., & Krishna, J. V. PhishGuard: Machine Learning Model. Innovations in Computational Intelligence and Computer Vision: Proceedings of ICICV 2024, Volume 2, 31