

Privacy-preserving over encrypted data for web page phishing detection

Ali Kadhim Bermani *

Department of Information Networks

College of Information Technology

University of Babylon

Babylon

Iraq

and

Department of Computer Techniques Engineering

College of Engineering and Technologies

Al-Mustaqbal University

Babylon

Iraq

Ahmed M. Al-Salih [†]

Department of Information Networks

College of Information Technology

University of Babylon

Babylon

Iraq

Hiba Ameer Jabir [§]

Department of Information Networks

College of Information Technology

University of Babylon

Babylon

Iraq

* E-mail: ali.k.berman@uobabylon.edu.iq (Corresponding Author)

[†] E-mail: ahmed.alsalih@uobabylon.edu.iq

[§] E-mail: hiba.alkhafaji@uobabylon.edu.iq

Mehdi Ebady Manaa[†]
*Department of Information Networks
College of Information Technology
University of Babylon
Babylon
Iraq*

and

*Department of Intelligent Medical System
College of Sciences
Al-Mustaqbal University
Babylon
Iraq*

Abstract

It tricked users into disclosing sensitive information through web page phishing. A traditional phishing detection method involves analyzing web pages and user behaviour using machine learning models, which require access to raw data, raising privacy concerns. Using homomorphic encryption and multiparty computation, this paper protects privacy in phishing detection. Security and privacy are enhanced by encrypting sensitive user data during the detection process. With this approach, Extreme Learning Machine (ELM) classification is combined with TF-IDF feature selection to detect phishing while maintaining optimal computational efficiency. According to our experimental results, this approach provides an accuracy of 93.5% and has low computational and communication costs, making it a viable solution for secure real-time phishing detection.

Subject Classification: 93E10, 94A13.

Keywords: Phishing detection, Privacy-preserving, Homomorphic encryption, Secure multiparty computation, Machine learning, Extreme Learning Machine (ELM), TF-IDF.

1. Introduction

A significant threat to cybersecurity is web page phishing, which is becoming more common as more and more people depend on the internet for communication, financial transactions, and sharing information[1]. Phishing websites trick users into disclosing sensitive information such as login credentials, credit card details, and personal details. The detection of phishing attacks traditionally relies on machine learning models that analyze the web page's features and the user's behaviour. Nevertheless, these approaches often require raw data access, raising questions of data security and privacy. We explore a privacy-preserving approach to detecting web page phishing over

[†] E-mail: mahdi.ebadi@uomus.edu.iq

encrypted data in order to address this challenge. Because homomorphic encryption and secure multiparty computation are used in this method, sensitive user data remains encrypted during the detection process. It allows for the detection of phishing attacks more accurately and the protection of user data from unauthorized access. There is a haemorrhage of personal information on the Internet. In the vast majority of cases, transactions conducted on the internet are scanned by a party outside of the parties directly involved. In order to track user movements across websites, advertisers use tracking cookies, web bugs, and other tools [2].

In some cases, less scrupulous organizations use spyware to obtain personal information illegally. It is possible to warehouse, collate, and store indefinitely such information. In this fraudulent act, online users are deceived in order to gain access to their personal information through the use of the Internet [3], [4]. It is also referred to as phishing when an attacker plans a phishing attack. Phishers have been stealing the usernames and passwords of AOL users since 1996 [5]. According to Figure 1, most successful phishing attacks use email spoofing [6] and website spoofing [7].

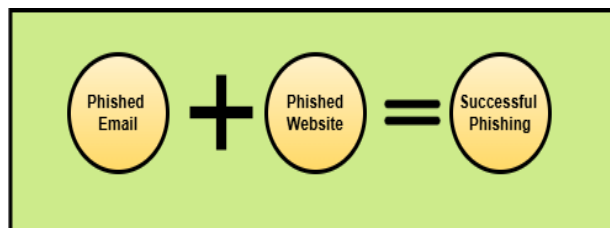


Figure 1
Combining phishing emails and websites

Spoofing involves making a legitimate email address appear to be the source address of the email. Bank managers may use the email address `bankmanager@xyz.co.in` so that phishers can pose as them and fool users into thinking the mail is legitimate. In most cases, the email request calls for recipients to click links, provide their personal information, or reply with their bank account information. Email spoofing is carried out by attackers using the SMTP (Simple Mail Transfer Protocol) server protocol. Since most users do not reply to emails with their personal information, phishing websites have been designed to resemble authentic websites in terms of appearance and functionality.

1.1. Threat Analysis

Privacy and security are threatened by many risks associated with web-based email. We will examine possible adversaries and their motivations and attacks in the following sections. Application providers that provide web-based email services may face threats from corporate adversaries who are interested in profiling their users for revenue-generating purposes. A marketing agency or

another company interested in advertising products to a specific demographic can acquire information about users.

1.2. Email Protection

Aquinas's email message security is described in Figure 2. After email messages are composed, steganography, an encryption method, is used to disguise their content. A trusted third party is necessary to guarantee user identities, select key encryption parameters, and verify credentials in public-key systems[8]. This problem is not trivial, and it has not been fully resolved [9]. Furthermore, this architecture was not compatible with our system objectives in terms of infrastructure and connectivity.

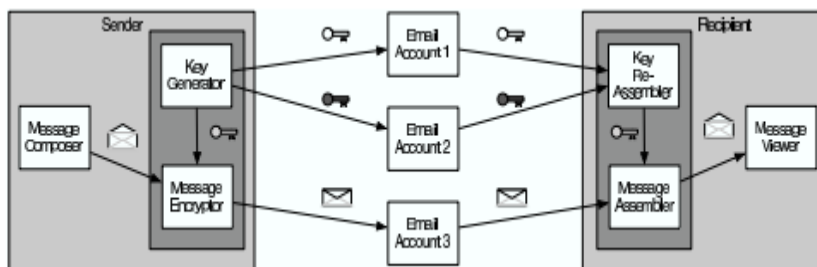


Figure 2

This is a sample flowchart for delivering messages and keys.

These new services, which include free email, allow you to create profiles that go beyond simply measuring an individual's online habits based on their communications. These companies allow them to scan messages passing through their systems, allowing commercial interests unprecedented access to people's private messages. In addition, such communications can be intercepted and examined by repressive regimes. Web-based applications are increasingly using these practices.

In the real world, most people use their browsers to get interesting information, and most users are using their personal banking and work accounts, which can lead to third-party access to our information. We aimed to protect valuable data from phishing by proposing a SAFE BROWSING model [10]. Our user's local history is encrypted as cypher text, and its true meaning is hidden using encryption. In real-time, safe browsing protects users from malicious websites while allowing them to use them freely. It would be impossible to read your browsing history, other personal details, and other details that are stored as encrypted (cypher text) using our model. When requesting their browsing history, users should request an OTP or keyword from the server. If they maintain a registered email address or phone number, they may receive the OTP or keyword via email or phone [11], [12]. Our detection of unsafe URLs uses a third-party blocklist, which provides safe URLs with a guarantee that they are not malicious. We use this key as a valuable key so that data encrypted in the

user's browser can be decrypted by the server (plain text). By enabling this feature, you will be protected from unsafe URLs and will not be able to get through to further content. It is possible to tell the server that some of the keywords or URLs are unsafe by submitting feedback [13].

Cybercriminals impersonate legitimate web pages to obtain sensitive data from businesses and individuals through phishing, a form of social engineering. In recent years, phishing has become more common due to its high reward (e.g., access to bank accounts and credit card numbers), the ease of sending forged email messages impersonating legitimate authorities, and the difficulty of pursuing criminals. In 2004, 1.2 million Americans were affected by phishing attacks, and businesses lost billions of dollars. In addition to business losses caused by phishing, consumer fear also contributes to business losses. As a result of identity fraud, people have become increasingly averse to Internet commerce, even though US companies assume that risk. Users are also increasingly distrusting emails received from financial institutions [14].

2. Related Work

A user can access encrypted data through the CP and generate encrypted queries using her keys for secure NDD, as described in this paper [15]. Stage two involves the user generating an encrypted query to the nearest IS using their encrypted key to find near duplicate data items stored on the encrypted in-network storage. A collaboration between phone vendors and SPs would help detect and remove malware from app stores, as noted in [16]. In spite of this, SPs and vendors are reluctant to share malware signatures, posing a dilemma. Here, we present a method for preserving privacy by combining two commonly used static detection techniques. The Droid Ranger detects malware by analyzing permission requests and extracting semantic footprints, which are details about each app that are based on semantics. The second step involves comparing the apps being scanned with malware samples available on the Internet. A user's smartphone and app store are protected against malicious apps without sharing app information, app runtime behaviour, or malware signatures with third parties. In [17], a system was designed in which crypto service providers (CSPs) issued encryption and decryption keys to all other parties as part of the system initialization process. There are instances in which remote servers perform secure interactive protocol sessions during training. By preserving privacy, we should be able to obtain word vectors that are as accurate as those obtained in plaintext. There should be no substantial communication or computation costs associated with the remote server as well as the crypto service provider.

In this query, True or False is returned depending on whether an item belongs to a particular set. Querying these values involves both the server and the client. The Private Membership Test (PMT) protocol allows users to perform membership tests without revealing their search results to database administrators. Computers store databases that measure PMT values. It is possible that preventing the distribution of query set X has a problem. As long

as the distribution of query set X is disclosed, the second scheme will be able to cache intermediate results locally, therefore enhancing security. In [18], The majority of anti-phishing technologies and methods are passive, relying on input from users to identify phishing URLs. Normally, they are incapable of quickly and efficiently detecting and eliminating phishing attacks. The study [19] proposes WC-PAD (Web Crawler-based Phishing Attack Detector), which recognizes attacks using web crawlers. The researchers in this study agreed on the crucial features that should be used to detect phishing, but they ignored other potential aspects that might be useful in predicting attacks.

Using Fuzzy Rough Set (FRS) theory, we determine which attribute is most appropriate among the three datasets. A phishing detection classifier is trained using the features selected. In [20], part of the phishing website classification process, the author uses a polynomial neural network, followed by optimization techniques such as genetic algorithms, gradient descent, and particle swarm optimization; however, genetic algorithm outperforms gradient descent as well as being computationally efficient. Data structures and simple chromosomal recombination operators are incorporated into a genetic algorithm to encode solutions.

3. Proposed Methodology

3.1 Homomorphic Encryption

The HE encryption method can be used on ciphertexts that require data privacy so that third parties can perform arithmetic operations on them directly. To begin with, partial homomorphic encryption (PHE) was developed, which can only perform additions or multiplications. Taking the AHE scheme as an example, it can only perform addition operations:

$$D_{sk_i}(E_{pk_i}(m_1).E_{pk_i}(m_2)) = m_1 + m_2 \quad (1)$$

Decryption using $D_{sk_i}(\cdot)$ is performed by a private key, encryption using $sk_i, E_{pk_i}(\cdot)$ is performed by a public key pk_i , and plaintext is handled by m_i . With (1), the cloud server is capable of performing homomorphic addition operations without decryption. Full homomorphic encryption (FHE) allows both additions and multiplications, which is not possible with polymorphic encryption (PHE) [21]. Adding and multiplying are used in FHE to implement a wide range of operations. Through these HE technologies, PPML algorithms have been developed for machine learning and cloud computing.

3.2. Distributed Homomorphic Cryptosystem

In this type of cryptosystem, homomorphic operations are performed using distributed homomorphic operations supported by secure multiparty computation (SMC). In public-key cryptography, public and private keys are used to encrypt and decrypt communications. DHC distributes a private key across multiple distributed servers by dividing it into several partial private keys. Distributed servers encrypt values using a public key so that partial decryption can be performed. Another distributed server can obtain the plaintext of the

ciphertext, as the ciphertexts have been partially decrypted. An array of homomorphic operations can be enabled through multilateral cooperation. We created a PPFL system based on the DHC that enables both parties to update a global model using homomorphic operations and protects participants' privacy.

3.3. EL for classification

For phishing Web page detection, neural networks have been used as classifiers and have demonstrated their effectiveness [22]. A neural network that uses single hidden layers feed-forward is trained using ELM, which uses least squares as a learning algorithm. A neural network that uses single hidden layers feed-forward is trained using ELM, which uses least squares as a learning algorithm. The ELM assigns a random input weight and bias, and the output weight of the SLFNs is calculated analytically [23].

We refer to the training data set as $\{(x_i, t_i) | x_i \in R^n, t_i \in R^m, i = 1, 2, \dots, N\}$. Using random input biases and weights, the hidden neuron matrix can be calculated by calculating the hidden neuron output matrix [24],

$$H = \begin{bmatrix} g(w_1 \cdot x_1 + b_1) & \cdots & g(w_K \cdot x_1 + b_K) \\ \vdots & \ddots & \vdots \\ g(w_1 \cdot x_n + b_1) & \cdots & g(w_K \cdot x_n + b_K) \end{bmatrix} \quad (2)$$

In this example, $w_j = [w_{j1}, w_{j2}, \dots, w_{jn}]$ ($j = 1, 2, \dots, K$) stands for input weight, b_j stands for hidden neuron bias, N stands for input sample number, and G stands for activation function. A Neural Network's output in relation to x_i is called o_i [24],

$$o_i = \sum_{j=1}^K \beta_j g(w_j \cdot x_i + b_j) \quad i = 1, 2, \dots, N \quad (3)$$

In this case, $\beta_j = [\beta_{j1}, \beta_{j2}, \dots, \beta_{jm}]$ represents the weight used to connect the j th hidden neuron to its corresponding output neuron. It is possible to make the SLFNs suitable for N samples using β with zero error, which means that $\sum_{i=1}^N ||o_i - t_i|| = 0$ is valid [24].

$$\sum_{j=1}^K \beta_j g(w_j \cdot x_i + b_j) \quad i = 1, 2, \dots, N \quad (4)$$

The linear system (4) can be rewritten as follows: [24],

$$H\beta = T \quad (5)$$

It is equivalent to finding a least-squares $B\hat{\beta}$ for the linear system when training ELMs using minimum norm least-squares.

$$||H\hat{\beta} - T|| = \min_{\beta} ||H\beta - T|| \quad (6)$$

Due to the large number of hidden nodes compared to samples, the smallest norm least-squares solution of (5) is produced by,

$$\hat{\beta} = H^\dagger T \quad (7)$$

The Moore-Penrose generalized inverse operator on $H^\dagger$ is denoted by $H^\dagger$. $\hat{\beta}$ is the least-squares solution to (5) with the lowest minimum norm. The decision function then consists of,

$$f(x) = \text{sign}(H(x)\beta) \quad (8)$$

Due to the random assignment of weights and biases, there is a possibility that a single ELM might misclassify some samples. OPELM (Optimal Pruning Extreme Learning Machine) is a robust method for classifying data that shows superior performance on big data. A high-accuracy ELM is proposed based on ensembles. In most experiments, the ELM classifier results are fused with a majority voting strategy, which doesn't take the performance of each classifier into account. Researchers, therefore, study linear combination models to combine ensembles of ELMs.

3.4. Feature Extraction

Cleansed text can be converted into quantifiable features that can be used for spam detection by feature extraction [25]. In statistical data analysis, the TF-IDF metric measures how significant words are in a corpus. In the dataset, the frequency of words and their meaning are considered. With TF-IDF, less significant terms are removed, simplifying the process of building models and reducing input dimensionality. As shown in the following formula, TF-IDF :

$$W_{i'j} = tf_{i'j} * \frac{\log N}{df_i} \quad (9)$$

The $tf_{i'j}$ component counts the variety of occurrences of a particular term in a specific document, and the j, df_i component counts the number of documents that contain that term. The word "part" is valued at 1 in two instances of "Text1". The term "DF" refers to the frequency with which it appears in documents in the corpus. It assesses the document frequency of "part" in a corpus containing two documents, "Text1" and "Text2".

4. Result and Discussion

Using Microsoft's in-house URL detonation tool, we collected screenshots of websites and manually created labels that identified the brands that were being phished. The dataset contains 18,854 malicious screenshots and 1000 legitimate screenshots. In our dataset, we represent 20 companies targeting social media,

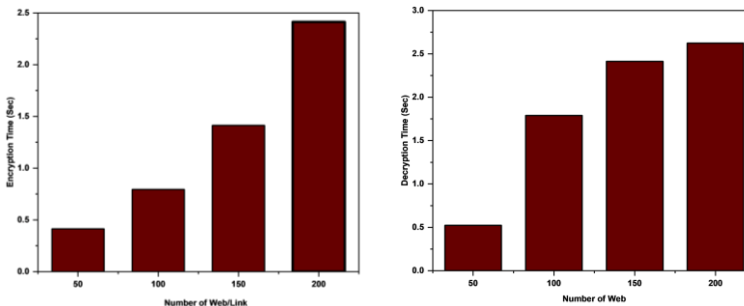


Figure 3
Encryption and decryption time (sec) versus number of web/link.

banking, and enterprise industries, including Facebook, LinkedIn, and American Express, among others. In Figure 3, the encryption and decryption times (sec) for homomorphic encryption are summarized. To run in a normal setting, the proposed model requires 0.0114 seconds and operates on 40,960 operations multiplicatively and 4,106 additively.

By combining encrypting and computing in a single step, the proposed method significantly reduces computation time. The ELM classifier is rescaled with TF-IDF so that encryption time is minimized and computation time is reduced to around 2.2 seconds. Data theft risks can be effectively mitigated by completing the process within the typical time required to enter credentials, enabling a seamless user experience as shown in Figure 4.

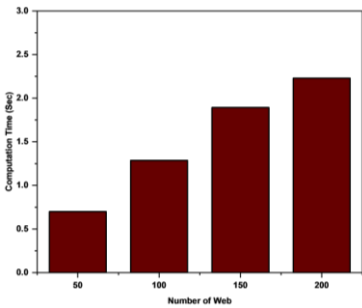


Figure 4
Computation time (sec) versus
number of web/links.

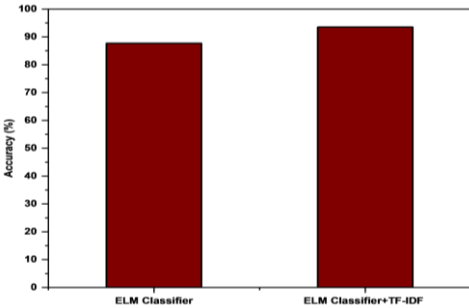


Figure 5
Comparative analysis of the proposed model
in terms of accuracy (%).

In our experiments, homomorphic encryption optimizations, along with feature selection for TF-IDF, have been shown in Figure 5, to improve the effectiveness of the ELM classifier. As a result of open-source models and software, the pipeline achieved over 93.5% accuracy while running in under a second, with communication costs under a megabyte. A TF-IDF classifier with ELM alone produces suboptimal accuracy, while a combination of both significantly improves it. Based on these improvements, it appears that TF-IDF captures textual features while ELM extracts visual characteristics, resulting in a more robust, accurate model.

5. Conclusion

The prevalence of phishing in cyberspace necessitates robust detection mechanisms that balance accuracy and user privacy. As part of the analysis presented in this paper, homomorphic encryption was used to ensure that sensitive data remained encrypted throughout the process of phishing detection. As a result, the proposed method enhances the accuracy of the ELM classifier while minimizing the computations and encryptions required. The model's accuracy in detecting phishing exceeds 93.5% with a minimal computational footprint while maintaining a high level of accuracy. Combining privacy-

preserving cryptography with machine learning offers a promising avenue for phishing detection that is both secure and efficient. Efforts will be made to optimize encryption overhead further, as well as extend the model to detect evolving phishing techniques in real-life scenarios.

References

- [1] P. Rani, S. Verma, S. P. Yadav, B. K. Rai, M. S. Naruka, and D. Kumar, "Simulation of the lightweight blockchain technique based on privacy and security for healthcare data for the cloud system," *International Journal of E-Health and Medical Communications (IJEHMC)*, vol. 13, no. 4, pp. 1–15 (2022).
- [2] K. Butler, W. Enck, J. Plasterr, P. Traynor, and P. McDaniel, "Privacy Preserving Web-Based Email," in *Information Systems Security*, vol. 4332, A. Bagchi and V. Atluri, Eds., in *Lecture Notes in Computer Science*, vol. 4332., Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 116–131 (2006). doi: 10.1007/11961635_8.
- [3] E. E. Lastdrager, "Achieving a consensual definition of phishing based on a systematic review of the literature," *Crime Sci*, vol. 3, no. 1, p. 9 (Dec. 2014), doi: 10.1186/s40163-014-0009-y.
- [4] B. Bhola, S. Misra, M. S. Obaidat, and A. Mukherjee, "Quality-enabled decentralized dynamic IoT platform with scalable resources integration," *IET Communications* (2022).
- [5] M. Khonji, Y. Iraqi, and A. Jones, "Phishing Detection: A Literature Survey," *IEEE Commun. Surv. Tutorials*, vol. 15, no. 4, pp. 2091–2121 (2013), doi: 10.1109/SURV.2013.032213.00009.
- [6] K. Pandove, A. Jindal, and R. Kumar, "Email spoofing," *International Journal of Computer Applications*, vol. 5, no. 1, pp. 27–30 (2010).
- [7] J. Hong, "The state of phishing attacks," *Commun. ACM*, vol. 55, no. 1, pp. 74–81 (Jan. 2012), doi: 10.1145/2063176.2063197.
- [8] P. Rani, C. Sharma, J. V. N. Ramesh, S. Verma, R. Sharma, A. Alkhayyat, and S. Kumar, "Federated Learning-Based Misbehaviour Detection for the 5G-Enabled Internet of Vehicles," *IEEE Transactions on Consumer Electronics* (2023).
- [9] C. Ellison and B. Schneier, "Ten risks of PKI: What you're not being told about public key infrastructure," *Comput Secur J*, vol. 16, no. 1, pp. 1–7 (2000).

- [10] L. Demir, A. Kumar, M. Cuncu, and C. Lauradoux, "The pitfalls of hashing for privacy," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 1, pp. 551–565 (2017).
- [11] T. Gerbet, A. Kumar, and C. Lauradoux, "A privacy analysis of Google and Yandex safe browsing," in 2016 46th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN), IEEE, 2016, pp. 347–358. Accessed: Feb. 22 (2025). [Online]. Available: <https://ieeexplore.ieee.org/abstract/document/7579754/>
- [12] P. Rani, P. N. Singh, S. Verma, N. Ali, P. K. Shukla, and M. Alhassan, "An implementation of modified blowfish technique with honey bee behaviour optimization for load balancing in the cloud system environment," *Wireless Communications and Mobile Computing*, vol. 2022, pp. 1–14 (2022).
- [13] A. Singh, S. Putluri, P. Rani, N. K. Agrawal, R. Sharma, E. Kariri, and D. G. Aray, "Smart Traffic Monitoring Through Real-Time Moving Vehicle Detection Using Deep Learning via Aerial Images for Consumer Application," *IEEE Trans. Consumer Electron.*, pp. 1–1 (2024), doi: 10.1109/TCE.2024.3445728.
- [14] G. Goth, "Phishing attacks rising, but dollar losses down," *IEEE Secur. Privacy Mag.*, vol. 3, no. 1, p. 8 (Jan. 2005), doi: 10.1109/MSP.2005.21.
- [15] S. Keelveedhi, M. Bellare, and T. Ristenpart, "{DupLESS}:{Server-Aided} encryption for deduplicated storage," in 22nd USENIX security symposium (USENIX security 13), pp. 179–194 (2013). Accessed: Feb. 22, 2025. [Online]. Available: <https://www.usenix.org/conference/usenixsecurity13/technical-sessions/presentation/bellare>
- [16] X. Yuan, X. Wang, C. Wang, C. Yu, and S. Nutanong, "Privacy-Preserving Similarity Joins Over Encrypted Data," *IEEE Trans. Inform. Forensic Secur.*, vol. 12, no. 11, pp. 2763–2775 (Nov. 2017), doi: 10.1109/TIFS.2017.2721221.
- [17] Q. Wang, W. Guo, X. Du, and M. Guizani, "Privacy-Preserving Collaborative Model Learning: The Case of Word Vector Training," *IEEE Trans. Knowl. Data Eng.*, vol. 30, no. 12, pp. 2381–2393 (Dec. 2018), doi: 10.1109/TKDE.2018.2819673.
- [18] MQ. Wang, W. Guo, X. Du, and M. Guizani, "An efficient algorithm to detect DDoS amplification attacks," *IFS*, vol. 39, no. 6, pp. 8565–8572 (Dec. 2020), doi: 10.3233/JIFS-189173.

- [19] T. Floyd, M. Grieco, and E. F. Reid, "Mining hospital data breach records: Cyber threats to U.S. hospitals," in 2016 IEEE Conference on Intelligence and Security Informatics (ISI), Tucson, AZ, USA: IEEE, Sep. 2016), pp. 43–48. doi: 10.1109/ISI.2016.7745441.
- [20] D. Kothandaraman, A. Kumar Sivaraman, S. Sundar, and K. Dhanalakshmi, "Energy and bandwidth based link stability routing algorithm for IoT," *Computers, Materials & Continua*, vol. 70, no. 2, pp. 2817–2833 (2022). [Online]. Available: https://www.researchgate.net/profile/Arun-Kumar-Sivaraman/publication/354859362_Energy_and_Bandwidth_Based_Link_Stability_Routing_Algorithm_for_IoT/links/61518d7d154b3227a8b02408/Energy-and-Bandwidth-Based-Link-Stability-Routing-Algorithm-for-IoT.pdf. Accessed: Feb. 22, 2025.
- [21] X. Liu, R. H. Deng, K.-K. R. Choo, and J. Weng, "An Efficient Privacy-Preserving Outsourced Calculation Toolkit With Multiple Keys," *IEEE Trans. Inform. Forensic Secur.*, vol. 11, no. 11, pp. 2401–2414 (Nov. 2016), doi: 10.1109/TIFS.2016.2573770.
- [22] R. M. Mohammad, F. Thabtah, and L. McCluskey, "Predicting phishing websites based on self-structuring neural network," *Neural Comput & Applic*, vol. 25, no. 2, pp. 443–458 (Aug. 2014), doi: 10.1007/s00521-013-1490-z.
- [23] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: theory and applications," *Neurocomputing*, vol. 70, no. 1–3, pp. 489–501 (2006).
- [24] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: a new learning scheme of feed-forward neural networks," in 2004 IEEE International Joint Conference on neural networks (IEEE Cat. No. 04CH37541), *Ieee*, pp. 985–990 (2004). Accessed: Feb. 26, 2025. [Online]. Available: <https://ieeexplore.ieee.org/abstract/document/1380068/>
- [25] I. Guyon and A. Elisseeff, "An Introduction to Feature Extraction," in Feature Extraction, vol. 207, I. Guyon, M. Nikravesh, S. Gunn, and L. A. Zadeh, Eds., in Studies in Fuzziness and Soft Computing, vol. 207., Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 1–25 (2006). doi: 10.1007/978-3-540-35488-8_1.

Received February, 2025