Generative adversarial network-based phishing URL detection with variational autoencoder and transformer

Jishnu Kaitholikkal Sasi, Arthi Balakrishnan
Department of Computing Technologies, College of Engineering and Technology, SRM Institute of Science and Technology,
Kattankulathur Campus, Chengalpattu, TN, India

Article Info

Article history:

Received Jun 9, 2023
Revised Oct 31, 2023
Accepted Dec 2, 2023

Keywords:

Cyber crimes
Generative adversarial networks
Phishing URLs
Transformers
Variational autoencoders

ABSTRACT

Phishing attacks pose a constant threat to online security, necessitating the development of efficient tools for identifying malicious URLs. In this article, we propose a novel approach to detect phishing URLs employing a generative adversarial network (GAN) with a variational autoencoder (VAE) as the generator and a transformer model with self-attention as the discriminator. The VAE generator is trained to produce synthetic URLs. In contrast, the transformer discriminator uses its self-attention mechanism to focus on the different parts of the input URLs to extract crucial features. Our model uses adversarial training to distinguish between legitimate and phishing URLs. We evaluate the effectiveness of the proposed method using a large set of one million URLs that incorporate both authentic and phishing URLs. Experimental results show that our model is effective, with an impressive accuracy of 97.75%, outperforming the baseline models. This study significantly improves online security by offering a novel and highly accurate phishing URL detection method.

Corresponding Author:

Arthi Balakrishnan
Department of Computing Technologies, College of Engineering and Technology
SRM Institute of Science and Technology
Kattankulathur Campus, SRM Nagar, Chengalpattu, Chennai 603203, Tamilnadu, India
Email: arthib@srmist.edu.in

1. INTRODUCTION

In this modern digital environment, cyber crimes are more widespread than ever. Among this, phishing attacks are scandalous because of their wide range of attacking levels and the anonymity of attackers [1]. This poses a serious threat to individuals, organisations, and their sensitive data. In order to trick users into disclosing private information, such as passwords, credit card numbers, or personal information, phishing attempts frequently use fraudulent emails, messages, or websites that act as reliable sources [2]. A record-breaking 1,270,883 phishing attacks were reported during the third quarter, according to the anti-phishing working group (APWG) report from 2022. This is the highest quarterly figure ever recorded, highlighting the need for quick and efficient solutions to stop this growing threat. With 23.2% of all reported incidents, the financial sector in particular has been a top target for phishing attacks. Business email compromise (BEC) attacks, a sophisticated form of phishing, have continued, and during the third quarter, there was a 59% increase in wire transfer BEC attacks. The report also reveals a startling 1,000% increase in advance fee fraud scams that are sent via email, highlighting the constantly evolving methods used by cybercriminals. It is essential to recognise and block these phishing URLs in order to ensure online security and prevent users from becoming victims of cybercriminals [3].

Modern methods for identifying phishing URLs frequently rely on manual analysis, heuristics, or rule-based algorithms, which find it difficult to keep up with the attackers' constantly changing strategies [4]. Therefore, it is imperative to create reliable and automated techniques for efficiently detecting phishing URLs. Machine learning (ML) and deep learning techniques have recently demonstrated excellent results in a number of fields, and their use in phishing URL identification has enormous promise [5]. Since ML can learn automatically from the training dataset, numerous researchers have been looking at it for phishing detection. Features are taken from the URL in this. To achieve the highest level of accuracy, many of them created various feature extraction strategies for ML algorithms. With the help of ML algorithms like decision tree, support vector machine, random forest, and K-nearest neighbor, some researchers are able to detect phishing URLs with an accuracy of more than 90% [6]-[11]. The manual feature engineering of these models is their primary flaw. It implies that because the data were extracted based on manual interpretations, it's possible that crucial features that machines might have picked up on were missed. After that point, more researchers proposed deep learning-based techniques due to the automatic feature extraction characteristic of deep learning. Phishing detection systems based on convolutional neural network (CNN) were given in methods [12]-[16], while some researchers [17]-[21] utilized recurrent neural network (RNN), multilayer perceptron (MLP), long short-term memory (LSTM), and its hybrid models. The majority of the approaches were more than 95% accurate. However, the system's biggest flaw was its inability to recognize dynamic phishing URLs and use imbalanced data sets.

In this work, we propose a novel method for phishing URL detection using a generative adversarial network (GAN) that comprises a variational autoencoder (VAE) as the generator and a transformer model with self-attention as the discriminator. While VAEs can capture latent representations and generate realistic URL samples, generator parts have demonstrated remarkable success in producing synthetic URLs. The transformer model uses self-attention mechanisms to distinguish between legitimate and phishing URLs. The primary goal of this research is to implement a reliable and accurate phishing URL detection system that can adapt to changing phishing tactics and offer a strong defence against them. We aim to increase the accuracy and effectiveness of phishing URL detection, enhancing online security for people and organisations, by using the capabilities of GAN with VAE and transformer.

2. METHOD

The proposed methodology uses a GAN architecture with VAE as the generator and a transformer model with self-attention as the discriminator. The system aims to accurately detect phishing URLs and enhance online security against attacks [18]. Figure 1 shows the workflow of the proposed model and the following paragraphs outline the key components and steps of the proposed method.

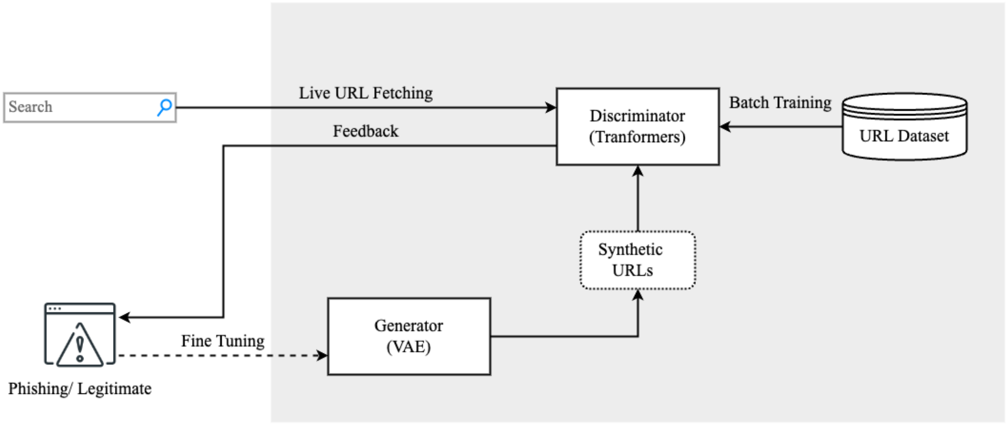

Figure 1. Workflow of the proposed model

2.1. Dataset preparation

Dataset preparation focuses on collecting and preparing the dataset. It comprises one million URLs, including both phishing URLs from different sources as shown in Table 1 [22]. The CSV file consists of a URL column and a label column with corresponding labels (0 for legitimate URLs and 1 for phishing URLs). The tokenizer from Keras is used to tokenize URLs at the character level. Sequences of tokens are generated using texts_to_sequences. pad_sequences are used to ensure that all urls are of the same length. In order to clean, normalise, remove duplicates, and ensure a balanced distribution between phishing and legitimate URLs, various data preprocessing methods are used. The data set is split into 80% training and 20% testing sets using train_test_split [23].

Table 1. Dataset source details

URL data set	Source
Legitimate URLs	Majestic Million, Common Crawl, Kaggle, GitHub, and Alexa
Phishing URLs	PhishTank, Kaggle, Common Crawl, and OpenPhish

2.2. Variational autoencoder as generator

With the aim of developing a latent representation of URLs and producing synthetic URLs that closely resemble genuine ones, the VAE is trained on the prepared dataset. By limiting the reconstruction loss, the generator's performance is optimised, resulting in the generation of realistic URLs. In order to learn a low-dimensional representation (latent space) of the input URL sequences, the VAE model is built [24]. The maximum sequence length of the URLs in the dataset determines the input shape of the VAE. This guarantees that each URL sequence will be processed with the same length. The encoder portion of the VAE is added as a thick layer. The encoder layer reduces the input URL sequences' dimensionality to the designated latent dimension [25]. This is accomplished by projecting the input data into a lower-dimensional space via a non-linear transformation (activation function). The decoder portion of the VAE is added as a further dense layer. The latent representation created by the encoder is used by the decoder layer to reassemble the URL sequences. It returns the data's original dimensionality [26].

The Adam optimizer, a well-liked option for training neural networks, is used to create the VAE model. The VAE's aim is the binary cross-entropy loss function. It calculates the discrepancy between actual input URL sequences and predicted URL sequences. The VAE model is trained on the training URLs using a predetermined batch size and number of epochs. The model gains the ability to encode URLs into a lower-dimensional representation and decode them again to recreate the original sequences during training. The model is prompted to produce precise reconstructions of the input URLs by minimising the binary cross-entropy loss during training. In this method, the VAE model is trained to recognise the key traits and patterns in the input URL sequences. The remaining steps of the phishing URL detection system can make use of the efficient encoding and reconstruction of the URLs made possible by the reduced-dimensional latent space [27].

2.3. Transformer model with self-attention as discriminator

As the discriminator element of the GAN, the transformer model with self-attention is introduced. The transformer discriminator is pre-trained on a sizable corpus of text data, enabling it to extract URL context. To enable the discriminator to accurately identify between legitimate and phishing URLs, fine-tuning is carried out on the training set of URLs using a binary classification goal [28]. The discriminator model is constructed to classify URLs. The discriminator model consists of several layers. Figure 2 represents the structure of the discriminator.

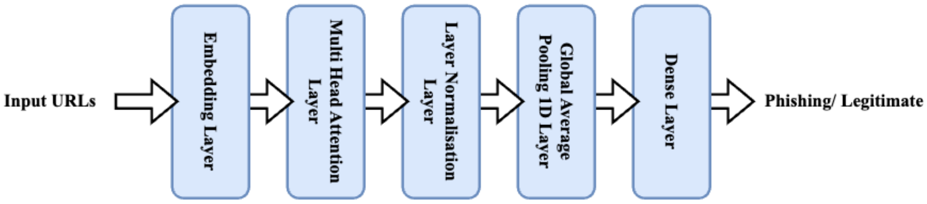
Input URLs
Embedding Layer
Multi Head Attention Layer
Layer Normalisation Layer
Global Average Pooling 1D Layer
Dense Layer
Phishing/ Legitimate

Figure 2. Structure of discriminator

- Embedding layer: the input URLs are converted to detailed vector representations by this layer. It helps in capturing the semantic intent behind the URLs.
- Multi head attention layer: the self-attention technique is used by this layer to extract key details from the embedded URL sequences. It enables the model to concentrate on different aspects of the input during processing. Figure 3 shows the multi head attention mechanism [29]. In actuality, we simultaneously compute the attention function on a collection of queries that are gathered into a matrix *Q*. In matrices *K* and *V*, the keys and values are also condensed together. Multiple simultaneous attention layers combine to create multi-head attention. The output matrix is calculated as follows:

$$MultiHead(Q,K,V) = Concat(h_1, h_2, ..., h_8)W^O \tag{1}$$

$$h_i = Attention(QW_i^Q, KW_i^K, VW_i^V) \tag{2}$$

where the projections are parameter matrices:

$$W_i^Q \in \mathbb{R}^{d_{model} \times d_k} \tag{3}$$

$$W_i^K \in \mathbb{R}^{d_{model} \times d_k} \tag{4}$$

$$W_i^V \in \mathbb{R}^{d_{model} \times d_v} \tag{5}$$

$$W^O \in \mathbb{R}^{hd_v \times d_{model}} \tag{6}$$

In our system, h=8 attention layers are there and they are arranged parallel. The value of $d_k = d_v = d_{model}/h = 64$.

- Normalisation layer: the outputs from the attention layer are normalised in this layer, guaranteeing stable training and enhanced performance.
- Global average pooling 1D layer: by calculating the average value over the time dimension, this layer can produce output with smaller spatial dimensions and URLs with defined lengths.
- Dense layer: for binary classification, a final dense layer with sigmoid activation is added. It generates a probability score that indicates whether a URL is likely to be real or phishing.

The Adam optimizer, an effective method for neural network training, is used to create the discriminator model. As the loss function, binary cross-entropy loss is employed. It calculates the difference between the actual labels and the expected probabilities. The model's training performance in classification is also measured using the accuracy metric. On the training URLs and their respective labels for the number of epochs, with a 256 batch size, the discriminator is trained. Based on the retrieved attributes, the model develops the ability to distinguish between legal and phishing URLs during training. The Adam optimizer adjusts the weights of the model to enhance classification performance while minimising the binary cross-entropy loss. The accuracy statistic aids in tracking the model's development throughout training. This method of training the discriminator teaches it to recognise patterns and characteristics that distinguish between legitimate and phishing URLs. In order to focus on crucial information in the URL sequences, the model uses the self-attention mechanism. The model then generates predictions based on the learned representations.

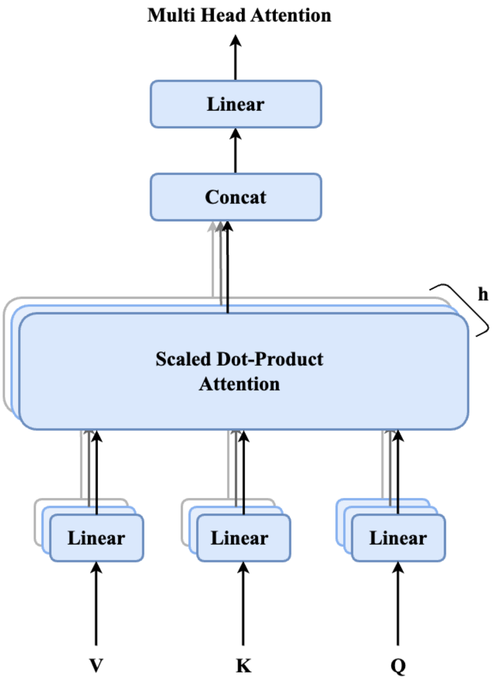

Figure 3. Multiple simultaneous attention layers combine to create multi-head attention

2.4. Adversarial training

The generator and discriminator are used to train the GAN using adversarial training procedures. To enhance their individual skills, the generator and discriminator are iteratively trained in a competitive manner. While the discriminator learns to distinguish between actual and fake URLs, the generator creates fake phishing URLs to trick it [30]. In this method, an adversarial training strategy is used to train a combined model that combines the VAE and discriminator models.

- Freezing the discriminator weights: the discriminator's trainable parameter is set to false prior to training the combined model. By doing this, the discriminator weights are guaranteed to remain constant during the adversarial training procedure. The VAE may concentrate on enhancing its reconstruction capability without being impacted by the discriminator's feedback because the discriminator's weights are frozen and are not modified.
- Model stacking: the combined model is built by combining the VAE and discriminator models. The discriminator for categorization uses the output of the VAE as its input. This stacking enables the combined model to carry out both the discriminator's classification task and the VAE's reconstruction mission.
- Compilation: the Adam optimizer, which is a well-liked optimizer for training neural networks, is used to create the merged model. To measure the difference between the genuine labels and the expected output of the discriminator, binary cross-entropy loss is utilised as the loss function.
- Training: the training URLs and labels for the training URLs are used to train the combined model over a predetermined number of epochs and with a predetermined batch size. The integrated model is simultaneously tuned during training to correctly classify the URLs and rebuild them accurately (VAE objective). The system's overall performance is enhanced by the model by taking use of the antagonistic interactions between the discriminator and VAE.

The combined model is trained in such a way that the VAE learns to produce convincing URL reconstructions that can fool the discriminator. At the same time, the discriminator gains proficiency in accurately distinguishing between trustworthy and phishing URLs. This adversarial training procedure enhances the system's capacity to identify and distinguish between malicious and trustworthy URLs.

3. RESULTS AND DISCUSSION

A dataset of one million URLs served as the basis for testing our suggested technique for detecting phishing URLs. The major objective was to improve online safety against phishing attacks by accurately differentiating between authentic and fraudulent URLs. The trial results show the efficiency of the suggested strategy, detecting phishing URLs with an astonishing accuracy of 97.75%. To assess the superiority of our model, we compared it with two other crucial deep learning-powered phishing detection models, as demonstrated in Table 2. A graphical representation of this is provided in Figure 4, with our model exhibiting clear dominance in every performance metric when compared to the other two.

Table 2. Comparison with baseline models

Models	Used algorithm	Dataset size	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Aljofey *et al.* [31]	Character level Convolution neural network	5,58,962	95.02	92.35	97.09	95.13
Ariyadasa *et al.* [32]	Long-term recurrent Convolutional network & Graph convolutional network	1,36,096	96.42	96.40	96.44	96.42
Proposed system	GAN using VAE and transformers	1 Million	97.75	97.91	97.62	97.77

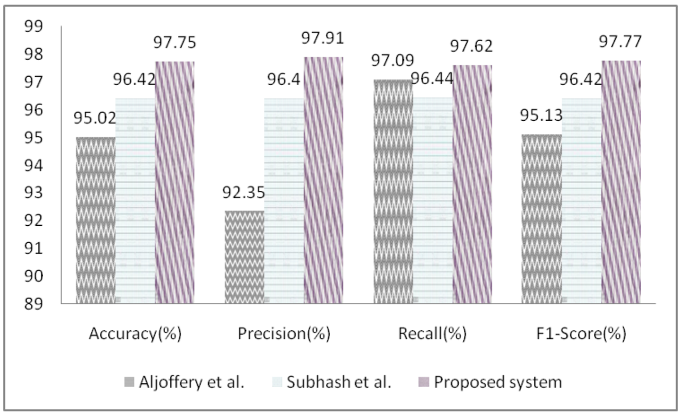

Figure 4. Comparison with baseline models

4. CONCLUSION

Using a transformer model with self-attention as the discriminator and a GAN with a VAE as the generator, we developed a unique method for phishing URL detection in this paper. Through intensive testing on a sizable sample of one million URLs, we attained a phenomenal accuracy of 97.75%. This illustrates how our suggested methodology might enhance internet security by accurately identifying phishing URLs. Our research demonstrates the potential of GANs for phishing URL detection. By exploiting the VAE's generation capabilities, we were able to create believable synthetic URLs that evaluated the discriminator's capacity to generate precise classifications. By using the transformer model with self-attention, we were able to extract important elements from the text and recognise the characteristic features of phishing URLs. The adversarial training increased the model's ability to discriminate, leading to better results than baseline models. Considering the great outcomes of our suggested approach, there are still many of areas that can be researched and improved. First, testing out different generator and discriminator designs might help the model perform better. We also want to develop a deep learning-based browser plugin for real-time phishing URL identification.

REFERENCES

[1] H. Abroshan, J. Devos, G. Poels, and E. Laermans, "COVID-19 and Phishing: Effects of Human Emotions, Behavior, and Demographics on the Success of Phishing Attempts During the Pandemic," *IEEE Access*, vol. 9, pp. 121916–121929, 2021, doi: 10.1109/ACCESS.2021.3109091.

[2] Z. Zhang, H. Al Hamadi, E. Damiani, C. Y. Yeun, and F. Taher, "Explainable Artificial Intelligence Applications in Cyber Security: State-of-the-Art in Research," *IEEE Access*, vol. 10, pp. 93104–93139, 2022, doi: 10.1109/ACCESS.2022.3204051.

[3] P. López-Aguilar, C. Patsakis, and A. Solanas, "The Role of Extraversion in Phishing Victimisation: A Systematic Literature Review," in *2022 APWG Symposium on Electronic Crime Research (eCrime)*, 2022, pp. 1–10, doi: 10.1109/eCrime57793.2022.10142078.

[4] N. Q. Do, A. Selamat, O. Krejcar, E. Herrera-Viedma, and H. Fujita, "Deep Learning for Phishing Detection: Taxonomy, Current Challenges and Future Directions," *IEEE Access*, vol. 10, pp. 36429–36463, 2022, doi: 10.1109/ACCESS.2022.3151903.

[5] R. Zieni, L. Massari, and M. C. Calzarossa, "Phishing or Not Phishing? A Survey on the Detection of Phishing Websites," *IEEE Access*, vol. 11, pp. 18499–18519, 2023, doi: 10.1109/ACCESS.2023.3247135.

[6] A. Basit, M. Zafar, X. Liu, A. R. Javed, Z. Jalil, and K. Kifayat, "A comprehensive survey of AI-enabled phishing attacks detection techniques," *Telecommunication Systems*, vol. 76, no. 1, pp. 139–154, 2021, doi: 10.1007/s11235-020-00733-2.

[7] S. Anupam and A. K. Kar, "Phishing website detection using support vector machines and nature-inspired optimization algorithms," *Telecommunication Systems*, vol. 76, no. 1, pp. 17–32, 2021, doi: 10.1007/s11235-020-00739-w.

[8] H. Shirazi, S. R. Muramudalige, I. Ray, and A. P. Jayasumana, "Improved Phishing Detection Algorithms using Adversarial Autoencoder Synthesized Data," in *2020 IEEE 45th Conference on Local Computer Networks (LCN)*, 2020, pp. 24–32, doi: 10.1109/LCN48667.2020.9314775.

[9] M. S. M. Prince, A. Hasan, and F. M. Shah, "A New Ensemble Model for Phishing Detection Based on Hybrid Cumulative Feature Selection," in *2021 IEEE 11th IEEE Symposium on Computer Applications & Industrial Electronics (ISCAIE)*, 2021, pp. 7–12, doi: 10.1109/ISCAIE51753.2021.9431782.

[10] A. Butnaru, A. Mylonas, and N. Pitropakis, "Towards lightweight url-based phishing detection," *Future Internet*, vol. 13, no. 6, pp. 1–15, 2021, doi: 10.3390/fi13060154.

[11] Ö. Kasim, "Automatic detection of phishing pages with event-based request processing, deep-hybrid feature extraction and light gradient boosted machine model," *Telecommunication Systems*, vol. 78, no. 1, pp. 103–115, 2021, doi: 10.1007/s11235-021-00799-6.

[12] M. Korkmaz, E. Kocyigit, O. K. Sahingoz, and B. Diri, "Phishing Web Page Detection Using N-gram Features Extracted From URLs," in *2021 3rd International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*, 2021, pp. 1–6, doi: 10.1109/HORA52670.2021.9461378.

[13] C. Opara, B. Wei, and Y. Chen, "HTMLPhish: Enabling Phishing Web Page Detection by Applying Deep Learning Techniques on HTML Analysis," in *2020 International Joint Conference on Neural Networks (IJCNN)*, 2020, pp. 1–8, doi: 10.1109/IJCNN48605.2020.9207707.

[14] W. Wei, Q. Ke, J. Nowak, M. Korytkowski, R. Scherer, and M. Woźniak, "Accurate and fast URL phishing detector: A convolutional neural network approach," *Computer Networks*, vol. 178, pp. 1–13, 2020, doi: 10.1016/j.comnet.2020.107275.

[15] S. Abdelnabi, K. Krombholz, and M. Fritz, "VisualPhishNet: Zero-Day Phishing Website Detection by Visual Similarity," in *Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security*, 2020, pp. 1681–1698, doi: 10.1145/3372297.3417233.

[16] S. Singh, M. P. Singh, and R. Pandey, "Phishing Detection from URLs Using Deep Learning Approach," in *2020 5th International Conference on Computing, Communication and Security (ICCCS)*, 2020, pp. 1–4, doi: 10.1109/ICCCS49678.2020.9277459.

[17] P. Yang, G. Zhao, and P. Zeng, "Phishing website detection based on multidimensional features driven by deep learning," *IEEE Access*, vol. 7, pp. 15196–15209, 2019, doi: 10.1109/ACCESS.2019.2892066.

[18] A. AlEroud and G. Karabatis, "Bypassing Detection of URL-based Phishing Attacks Using Generative Adversarial Deep Neural Networks," in *Proceedings of the Sixth International Workshop on Security and Privacy Analytics*, 2020, pp. 53–60, doi: 10.1145/3375708.3380315.

[19] T. Feng and C. Yue, "Visualizing and Interpreting RNN Models in URL-based Phishing Detection," in *Proceedings of the 25th ACM Symposium on Access Control Models and Technologies*, 2020, pp. 13–24, doi: 10.1145/3381991.3395602.

[20] I. Saha, D. Sarma, R. J. Chakma, M. N. Alam, A. Sultana, and S. Hossain, "Phishing Attacks Detection using Deep Learning Approach," in *2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT)*, 2020, pp. 1180–1185, doi: 10.1109/ICSSIT48917.2020.9214132.

[21] S. J. Bu and S. B. Cho, "Deep character-level anomaly detection based on a convolutional autoencoder for zero-day phishing url detection," *Electronics*, vol. 10, no. 12, pp. 1–16, 2021, doi: 10.3390/electronics10121492.

[22] A. Safi and S. Singh, "A systematic literature review on phishing website detection techniques," *Journal of King Saud University - Computer and Information Sciences*, vol. 35, no. 2, pp. 590–611, 2023, doi: 10.1016/j.jksuci.2023.01.004.

[23] B. B. Gupta, K. Yadav, I. Razzak, K. Psannis, A. Castiglione, and X. Chang, "A novel approach for phishing URLs detection using lexical based machine learning in a real-time environment," *Computer Communications*, vol. 175, pp. 47–57, 2021, doi: 10.1016/j.comcom.2021.04.023.

[24] M. K. Prabakaran, P. M. Sundaram, and A. D. Chandrasekar, "An enhanced deep learning-based phishing detection mechanism to effectively identify malicious URLs using variational autoencoders," *IET Information Security*, vol. 17, no. 3, pp. 423–440, 2023, doi: 10.1049/ise2.12106.

[25] L. Bergamin, T. Carraro, M. Polato, and F. Aiolli, "Novel Applications for VAE-based Anomaly Detection Systems," in *2022 International Joint Conference on Neural Networks (IJCNN)*, 2022, vol. 2022-July, pp. 1–8, doi: 10.1109/IJCNN55064.2022.9892879.

[26] G. Qi and H. Yu, "CMVAE: Causal Meta VAE for Unsupervised Meta-Learning," in *The Thirty-Seventh AAAI Conference on Artificial Intelligenc*, pp. 9480–9488, 2023, doi: 10.1609/aaai.v37i8.26135.

[27] L. Zhang, P. Zhang, L. Liu, and J. Tan, "Multiphish: Multi-Modal Features Fusion Networks for Phishing Detection," in *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2021, pp. 3520–3524, doi: 10.1109/ICASSP39728.2021.9415016.

[28] P. Maneriker, J. W. Stokes, E. G. Lazo, D. Carutasu, F. Tajaddodianfar, and A. Gururajan, "URLTran: Improving Phishing URL Detection Using Transformers," in *MILCOM 2021 - 2021 IEEE Military Communications Conference (MILCOM)*, 2021, pp. 197–204, doi: 10.1109/MILCOM52596.2021.9653028.

[29] A. Vaswani et al., "Attention is all you need," in *Advances in Neural Information Processing Systems*, 2017, pp. 5999–6009.

[30] S. Al-Ahmadi, A. Alotaibi, and O. Alsaleh, "PDGAN: Phishing Detection With Generative Adversarial Networks," *IEEE Access*, vol. 10, pp. 42459–42468, 2022, doi: 10.1109/ACCESS.2022.3168235.

[31] A. Aljofey, Q. Jiang, Q. Qu, M. Huang, and J. P. Niyigena, "An effective phishing detection model based on character level convolutional neural network from URL," *Electronics*, vol. 9, no. 9, pp. 1–24, 2020, doi: 10.3390/electronics9091514.

[32] S. Ariyadasa, S. Fernando, and S. Fernando, "Combining Long-Term Recurrent Convolutional and Graph Convolutional Networks to Detect Phishing Sites Using URL and HTML," *IEEE Access*, vol. 10, pp. 82355–82375, 2022, doi: 10.1109/ACCESS.2022.3196018.

BIOGRAPHIES OF AUTHORS

Jishnu Kaitholikkal Sasi is pursuing a Ph.D. in Computer Science from the SRM Institute of Science and Technology. He completed his B.Tech. in Computer Science from KMCT College of Engineering in 2017 and received an M.Tech. degree in Computer Science from Government Engineering College, Wayanad, in 2021. He has good programming skills and is proficient in software engineering, deep learning, and blockchain. His research mainly focuses on detecting phishing URLs using deep learning. He can be contacted at email: js2963@srmist.edu.in.

Arthi Balakrishnan holds a Ph.D. degree in the field of Computer Science and Engineering from Anna University. She has 15+ years of experience in teaching. Her areas of interest include artificial intelligence, machine learning, software engineering, and Internet of Things. She has published several articles in various reputed national and international, such as Scopus and SCI journals. She has also authored book chapters in CRC and Springer Publications. She has presented papers in various national and international conferences and attended many workshops, seminars, and faculty development programs to be in track with the changing technology and teaching methodology. She is a member of various scientific and professional bodies. She has been awarded the IET Inspiring Young Teacher Award for the year 2016-2017 for the IET Chennai. She can be contacted at email: arthib@srmist.edu.in.