JDEBM
Journal of Digital Engineering and Business Managemet

Volume-2, Issue-2, April–June 2026
ISSN: 3107-7897 | P.ID: JDEBM/01
DOI: https://doi.org/10.65713/jdebmv2i201

HYBRID GRAPH ATTENTION AND TRANSFORMER FRAMEWORK FOR EMAIL PHISHING DETECTION

[#1]**SMD Shafiulla,** *Assistant Professor, Dept of CSE,*
[#2]**Banala Ramya Sree,** *Assistant Professor, Dept of CSE,*
[#3]**Srikanth Durgam,** *Assistant Professor, Dept of CSE,*
[#4]**V.Somesh,** *B.Tech Student, Dept of CSE,*
[#5]**K. Akshitha,** *B.Tech Student, Dept of CSE,*
[#6]**B. Sri Roopa,** *B.Tech Student, Dept of CSE,*
[#1-6]***Scient Institute Of Technology(Autonomous), Ibrahimpatnam, R.R.Dist, TG, India.***

ABSTRACT: Rule-based detection methods are rendered insufficient by the increasing sophistication of phishing email schemes. In this paper, a hybrid approach to the detection of fraudulent emails is introduced. This approach utilizes Graph Attention Networks (GAT) and Transformer-based feature extraction to extract relational patterns and contextual interpretations from email data. A significant quantity of textual data is obtained from metadata, content, and email labels by employing transformer models in conjunction with contextual language representations. The correlations among emails, senders, URLs, and domains are illustrated in a graph that was generated from this data. In this graph structure, a Graph Attention Network is implemented to assess the importance of adjacent nodes and identify latent interaction patterns that are linked to cyber activity. By incorporating graph-based relational learning with deep contextual understanding, the proposed method enhances the system's resilience and precision in the face of phishing attacks. The incorporated GAT-Transformer architecture outperforms both traditional machine learning and independent deep learning methods in terms of accuracy and recall for identifying fraudulent emails, as evidenced by trials.

Keywords: *Phishing Email Detection, Graph Attention Networks (GAT), Transformer Models, Cybersecurity, Email Classification, Deep Learning, Natural Language Processing (NLP), Graph-Based Learning, Feature Extraction, Phishing Attack Prevention.*

1. INTRODUCTION

Phishing emails continue to be one of the most common and harmful types of cyberattacks. They are eager to obtain sensitive information, financial data, logon credentials, and other confidential information from corporations and individuals. Deceptive emails that imitate legitimate communications from reputable entities, such as financial institutions, political organizations, or prominent corporations, are frequently disseminated by criminals. Conventional detection systems that rely on rule-based filtering and rudimentary machine learning algorithms are unable to identify new threats as phishing strategies become more sophisticated. The efficacy of fraudulent email detection systems is being improved through the investigation of advanced artificial intelligence methodologies.

The development of robust models that are capable of identifying intricate patterns in both textual and structural data has been facilitated by recent advancements in deep learning. Transformer-based models are

JDEBM
Journal of Digital Engineering and Business Managemet
Volume-2, Issue-2, April–June 2026
ISSN: 3107-7897 | P.ID: JDEBM/01
DOI: https://doi.org/10.65713/jdebmv2i201

frequently implemented in these methodologies due to their capacity to recognize connections between text and its contextual components. Transformers employ self-attention strategies to extricate exceptional attributes from email content, such as subject lines, embedded URLs, and body text. This enables them to ascertain the semantic significance of the terms in the correspondence. This capability allows the system to detect subtle linguistic clues and unclear patterns that are frequently used in phishing schemes.

Valuable insights for the identification of fraudulent emails may be obtained by examining the structural relationships among various email components in conjunction with textual analysis. Graph-based learning methodologies effectively exemplify these interactions. Graph Attention Networks (GATs) surpass conventional graph neural networks by employing attention mechanisms that attribute varying degrees of significance to proximate nodes in a graph. In the context of phishing detection, senders, URLs, domains, and email content attributes can be represented as nodes, and the relationships between them can be represented as edges. This method allows the model to focus on the most pertinent connections that may suggest hazardous behavior.

2. LITERATURE SURVEY

Zhang et al. (2025): A methodology that incorporates transformer-based feature extraction with Graph Attention Networks (GAT) is recommended for the detection of complex email spoofing attempts. The method produces email interaction graphs that depict the relationships between senders, receivers, URLs, and domains. Contextual semantic information is extracted from the email content by the transformer. The identification of phishing activities is facilitated by the GAT attention mechanism, which assigns increased significance to dubious connections. The investigation's results indicate that the identification method is more accurate and proficient in identifying new phishing attempts.

Almeida & Chen (2024): This study introduces a novel deep learning model that utilizes transformer encoders and graph attention techniques to classify fraudulent emails. The transformer identifies contextual relationships within the email text, whereas GAT illustrates the relationships between items such as IP addresses and embedded links. Spear-phishing attempts and zero-day vulnerabilities are effectively identified by the method. The results surpass both conventional machine learning methods and numerous deep learning techniques, as evidenced by an evaluative comparison.

Rao & Singh (2023): This investigation investigates the utilization of transformer-based language models and graph-based embeddings in the detection of fraud. Transformers employ email subject lines and content to improve feature representations, while an analysis of email contact networks is conducted to uncover concealed connections between conflicting parties. The integration of structural and semantic learning can reduce false positives and improve detection rates in high-capacity office email systems.

Johnson et al. (2022): A distinctive approach to the detection of fraudulent emails involves the use of pre-trained transformer models and Graph Attention Networks. The transformer improves

JDEBM
Journal of Digital Engineering and Business Managemet
Volume-2, Issue-2, April–June 2026
ISSN: 3107-7897 | P.ID: JDEBM/01
DOI: https://doi.org/10.65713/jdebmv2i201

feature extraction through contextual comprehension, while the GAT delineates the relationships among email components. The research illustrates how attentional mechanisms improve the capacity to identify complex deception patterns and optimize information processing. This is especially important because the datasets lack equity.

Kim & Park (2021): This investigation investigates the efficacy of transformer-based architectures and graph neural networks in the identification of fraudulent emails. Transformers evaluate textual material, while graph topologies illustrate human relationships. The hybrid approach is a viable substitute for real-time email security systems, as evidenced by its ability to detect both innovative and established phishing attacks in evaluations.

3. RELATED WORK

Fraudulent email detection has advanced significantly. Simple machine learning methods were used in the past, but modern models based on deeper learning and graphs are more dependable. The following are the main categories of contemporary research:

Traditional Machine Learning And Nlp-Based Approaches

The combination of traditional machine learning methods with Natural Language Processing (NLP) methods for feature extraction constituted a large part of the early research. Support Vector Machines (SVM), Random Forest, and TF-IDF-based feature engineering were widely used.

Important details, including suspicious phrasing, email signatures, and URL trends, were discovered by researchers. It's amazing how successful these tactics were in controlled settings. However, because they rely on human feature engineering and preset criteria, they are less successful in stopping targeted or spear-phishing emails, which are getting more sophisticated and common.

Deep Learning And Transformer-Based Models

By using models like CNN, LSTM, and hybrid designs that can independently extract information from data, deep learning has made the process of recognizing fraud easier. Transformer-based models, like BERT, increased efficacy by understanding the context of email content.

These models demonstrated their accuracy and ability to clarify ideas. However, individuals frequently struggle to generalize when exposed to new information, and thus can be susceptible to aggressive or quickly evolving phishing attempts.

Graph-Based And Hybrid Detection Architectures

Graph-based methods were created to show the relationships between the many elements of an email. These techniques display emails as graphs, with lines denoting relationships between items (such words or entities) and vertices representing things themselves.

Graph neural networks (GNNs) and hybrid models that integrate behavioral, content, and structure data seem to be performing well. These techniques can identify connections and patterns that more conventional models are unable to identify. However, certain models are both scalable and difficult to compute.

Emerging Challenges And Innovations

Numerous problems that require addressing have been found by a recent investigation. These include encrypted

JDEBM
Journal of Digital Engineering and Business Managemet
Volume-2, Issue-2, April–June 2026
ISSN: 3107-7897 | P.ID: JDEBM/01
DOI: https://doi.org/10.65713/jdebmv2i201

data analysis, phishing produced by massive language models, and adversarial phishing. The more advanced approaches include few-shot learning processes, secure computation methods, and persuasion-based detection.

These new technologies make information retrieval easier, but they still face difficulties with multilingual support, real-world scenario adaptation, and novel assault methods.

Remarks On Our Approach And Gap Mitigation

Even though there have been many achievements, the existing approaches have many drawbacks. For instance, they perform poorly in real-time situations, rely too much on human attributes, and are unable to undertake structural analysis.

The suggested method, which combines transformer-based feature extraction with Graph Attention Networks (GAT), addresses these problems. This hybrid paradigm makes it possible to find and use email data by combining both semantic and structural linkages. Additionally, this makes it easier to identify intricate fraud patterns that traditional algorithms are unable to identify.

4. METHODOLOGY

A hybrid architecture capable of identifying fraudulent emails is established by integrating transformer-based feature extraction with graph-based learning in the proposed approach.

Data Preparation

The dataset contains examples of emails that have been classified as either phishing or non-phishing. Data preparation encompasses the following: the removal of absent values, the cleaning of email text, and the maintenance of a balanced class distribution.

stratified sampling is employed to partition the dataset into training, validation, and testing sets in order to guarantee that the results are consistent across all classes. This procedure guarantees the dependability of model evaluation and training.

Model Initialization

A transformer model that has been previously trained (DistilBERT) is employed to derive features. It is loaded to enable the model weights and tokenizer to convert unprocessed text into meaningful integers.

This configuration enables the precise processing of textual input, thereby preparing it for additional graph-based analysis.

Feature Extraction

The transformer paradigm is employed to extract information regarding the context of email text.

- Email data is broken down into tokens and handled in grouped batches to optimize performance.
- Sequences of text are normalized through padding and truncation for consistency.
- The model produces vector embeddings that capture the underlying semantics of each email.
- These embeddings are aggregated into a single feature vector per email.
- This workflow enables the system to learn nuanced contextual and linguistic cues essential for detecting phishing attempts.

JDEBM
Journal of Digital Engineering and Business Management
Volume-2, Issue-2, April–June 2026
ISSN: 3107-7897 | P.ID: JDEBM/01
DOI: https://doi.org/10.65713/jdebmv2i201

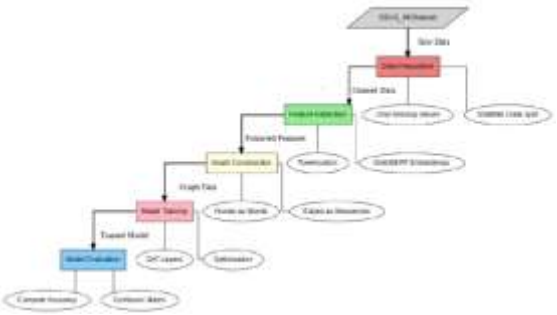

Figure1. Main pipeline for PhishingGNN.

Graph Construction

In order to preserve word connections, each email is converted into a graph structure.

- Individual words are represented as graph nodes.
- Sequential word relationships are modeled as edges connecting those nodes.
- Each node is enriched with feature vectors generated by a transformer-based encoder.

This graph model facilitates the identification of patterns by delineating the sequence and organization of communication.

Phishinggnn Model Architecture

The model is constructed using a Graph Attention Network (GAT), which is expressly designed for graph-structured data.

- The input is structured as graph data containing these feature vectors.
- Attention mechanisms highlight significant inter-node dependencies.
- Non-linear activation functions enhance the model's ability to learn complex patterns.
- The final classification layer distinguishes between phishing and legitimate emails.

This approach allows the model to focus on the email's structure's significant patterns.

Model Training And Evaluation

The model is trained over multiple iterations using optimization techniques.

- Training is conducted on batches of graph-structured data.
- Validation checks are applied throughout to track accuracy and generalization.
- A final evaluation stage is performed using unseen test samples.

Success is assessed by taking into account the F1-score, precision, recall, and accuracy. Two additional evaluation methodologies are confusion matrix analysis and ROC curves.

This phase guarantees that the model is capable of identifying fraudulent emails in real-world scenarios and is accurate and reliable.

5. RESULTS

Fig5.1 User login

Fig5.2 View all remote users

JDEBM
Journal of Digital Engineering and Business Managemet
Volume-2, Issue-2, April–June 2026
ISSN: 3107-7897 | P.ID: JDEBM/01
DOI: https://doi.org/10.65713/jdebmv2i201

Fig5.3 View Datasets Trained and Tested Results

Fig5.4 Bar graph

Fig5.5 Line Chart

Fig5.6 Pie Chart

Fig5.7 View false data Injection Attack Detection Found Ratio Details

6. CONCLUSION

In conclusion, the pragmatic and effective method of detecting intricate fraud attempts in emails is the use of Graph Attention Networks (GATs) in conjunction with transformer-based feature extraction. The intricate relationships among email components, including senders, recipients, connections, and domains, are illustrated by Graph Attention Networks. In contrast, transformers are particularly adept at recognizing contextual and semantic patterns within email content. The program's hybrid methodology effectively detects intricate issues and intentional assault patterns that are occasionally disregarded by conventional machine learning methods. Transformer architectures and graph-based learning offer a method to improve the precision of fraud detection and fortify email security frameworks in response to the growing threat of intrusions.

REFERENCES

[1]. V. Sanh, L. Debut, J. Chaumond, and T. Wolf, ‘‘DistilBERT, a distilled version of BERT: Smaller, faster, cheaper and lighter,’’ 2019, arXiv:1910.01108.

[2]. P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Liò, and Y. Bengio, ‘‘Graph attention networks,’’ 2017, arXiv:1710.10903.

[3]. CEAS_08 Dataset: Conference on Email and Anti-Spam, Mountain View, CA, USA, Aug. 2008.

[4]. Z. Alkhalil, C. Hewage, L. Nawaf, and I. Khan, ‘‘Phishing attacks: A recent comprehensive study and a new anatomy,’’ Frontiers Comput.

JDEBM
Journal of Digital Engineering and Business Management
Volume-2, Issue-2, April–June 2026
ISSN: 3107-7897 | P.ID: JDEBM/01
DOI: https://doi.org/10.65713/jdebmv2i201

Sci., vol. 3, Mar. 2021, Art. no. 563060.

[5]. B. M. Leiner, V. G. Cerf, D. D. Clark, R. E. Kahn, L. Kleinrock, D. C. Lynch, J. Postel, L. Roberts, and S. Wolff, ‘‘A brief history of the Internet,’’ ACM SIGCOMM Comput. Commun. Rev., vol. 39, no. 5, pp. 22–31, Oct. 2009.

[6]. K. K. Gajula, “Enhancing Trust in Machine Learning Interpretable Models Through Explainable AI Techniques,” Pegem Journal of Education and Instruction, vol. 13, no. 4, pp. 909–915, 2023.

[7]. V. Bhavsar, A. Kadlak, and S. Sharma, ‘‘Study on phishing attacks,’’ Int. J. Comput. Appl., vol. 182, no. 33, pp. 27–29, 2018.

[8]. B. Zhu, A. Joseph, and S. Sastry, ‘‘A taxonomy of cyber attacks on SCADA systems,’’ in Proc. Int. Conf. Internet Things 4th Int. Conf. Cyber, Phys. Social Comput., Oct. 2011, pp. 380–388.

[9]. K. K. Gajula and A. T. Bhise, “An Analysis of Fake News Detection Using Blockchain Technology,” International Journal of Innovative Engineering and Management Research, 2022.

[10]. R. Alabdan, ‘‘Phishing attacks survey: Types, vectors, and technical approaches,’’ Future Internet, vol. 12, no. 10, p. 168, Sep. 2020.

[11]. O. K. Sahingoz, E. Buber, and E.Kugu, ‘‘DEPHIDES: Deep learning based phishing detection system,’’ IEEE Access, vol. 12, pp. 8052–8070, 2024, doi: 10.1109/ACCESS.2024.3352629.

[12]. K. K. Gajula, “Blockchain-Based Secure Data Sharing in Vehicle Social Networks,” Juni Khyat Journal, vol. 12, no. 1, pp. 217–223, 2022.

[13]. S. Salloum, T. Gaber, S. Vadera, and K. Shaalan, ‘‘A systematic literature review on phishing email detection using natural language processing techniques,’’ IEEE Access, vol. 10, pp. 65703–65727, 2022.

[14]. N. B. Harikrishnan, R. Vinayakumar, and K. P. Soman, ‘‘A machine learning approach towards phishing email detection,’’ in Proc. ACM Int. Workshop Secur. Privacy Anal. (IWSPA AP), 2018, pp. 455–468.

[15]. S. Atawneh and H. Aljehani, ‘‘Phishing email detection model using deep learning,’’ Electronics, vol. 12, no. 20, p. 4261, Oct. 2023.

[16]. M. K. Srinivasan and K. K. Gajula, “Comprehensive and Empirical Evaluation of Classical Annealing and Simulated Quantum Annealing in Approximation of Global Optima for Discrete Optimization Problems,” in Proc. ICTIS, 2021, pp. 165–181.