

A Collective Inference Framework for Phishing URL Detection Using Loopy Belief Propagation

Riya Dandawate, Rutuja Hole, Rohit Patil

Department of Information technology and Engineering,
Nutan Maharashtra Institute of Engineering and Technology,
Pune, India

riyadandwate101@gmail.com, rutujahole2412@gmail.com, rohitpatil95702@gmail.com

Abstract : This study presents a collective inference framework for phishing URL detection built upon Loopy Belief Propagation (LBP) to establish a reliable and scalable phishing detection system. Instead of treating URLs as isolated entities, the proposed framework models relational dependencies among URLs using graph-based representations and probabilistic inference. Machine learning techniques are used to estimate initial phishing probabilities, while iterative message passing refines predictions across interconnected nodes. Experimental evaluation on large-scale phishing and legitimate URL datasets demonstrates that the proposed approach improves detection accuracy and supports efficient real-time phishing analysis.

Keywords— Phishing URL Detection, Loopy Belief Propagation, Graph-Based Learning, Collective Inference, Random Forest, Cyber Security.

I. INTRODUCTION

Phishing attacks continue to pose a major cybersecurity threat by exploiting deceptive URLs and fraudulent web pages to steal sensitive user information. As digital services continue to expand, attackers increasingly use sophisticated phishing techniques that bypass traditional security mechanisms. Conventional phishing detection approaches, including blacklist systems and rule-based filtering, often struggle to detect newly generated or dynamically evolving phishing URLs in real time.

Machine learning-based phishing detection methods have improved detection capabilities by learning lexical, structural, and domain-level patterns from URL datasets. However, most existing approaches treat URLs as isolated entities and fail to capture the hidden relationships that frequently exist among phishing campaigns. Malicious URLs often share common hosting infrastructure, domain characteristics, structural similarities, or behavioural patterns that cannot be effectively analysed through independent classification alone.

Graph-based learning provides a more effective approach for modelling such relational dependencies by representing URLs and their associated features as interconnected nodes within a graph structure. Probabilistic graphical models further enhance this representation by enabling collective inference under uncertainty. In particular, Loopy Belief Propagation (LBP) allows iterative probabilistic message passing across cyclic graph structures, making it suitable for large-scale phishing detection environments.

This study presents **PhiURL**, a collective inference framework for phishing URL detection using Loopy Belief Propagation.

The proposed framework integrates machine learning-based feature extraction, probabilistic graph modelling, and real-time

inference to improve phishing detection accuracy. Unary phishing probabilities are estimated using Random Forest, while pairwise dependencies among related URLs are modelled through graph-based similarity relationships. LBP is then applied to iteratively refine phishing predictions across interconnected nodes.

Unlike traditional phishing detection systems that rely solely on isolated feature analysis, the proposed framework combines graph-based collective reasoning with probabilistic inference to support scalable and reliable phishing detection. Experimental evaluation conducted on large-scale phishing and legitimate URL datasets demonstrates the effectiveness of the proposed approach for real-time phishing analysis and probabilistic threat detection.

II. RELATED WORK

Phishing URL detection has been extensively studied using blacklist systems, machine learning algorithms, and deep learning approaches. Traditional blacklist-based methods maintain databases of previously identified phishing websites and block user access to known malicious URLs. Although these techniques are effective against existing threats, they struggle to detect newly generated phishing websites and rapidly evolving attack patterns. Rule-based filtering systems similarly rely on predefined heuristics, making them less adaptable to sophisticated phishing strategies.

Machine learning-based phishing detection approaches have significantly improved automated threat identification by analysing lexical, host-based, and structural URL features. Several studies have applied classification algorithms such as Logistic Regression, Support Vector Machines (SVM), Random Forest, and Naive Bayes for phishing detection. These approaches extract features including URL length, domain age, suspicious keywords, special characters, and HTTPS usage to classify URLs as phishing or legitimate. While machine learning models improve detection accuracy, most existing systems process URLs independently and fail to capture relationships among phishing entities.

Recent research has explored deep learning techniques for phishing detection using neural networks and feature embedding methods. Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and transformer-based architectures have been used to automatically learn phishing patterns from large-scale datasets. Although deep learning approaches achieve strong classification

performance, they often require large computational resources and extensive training data, limiting their scalability and interpretability.

Graph-based learning has emerged as an effective solution for modelling relational dependencies in cybersecurity applications. Several studies have represented URLs, domains, and web entities as interconnected graph structures to analyse hidden phishing relationships. Graph mining and Graph Neural Network (GNN)-based approaches have demonstrated improved phishing detection by

leveraging structural and behavioural similarities among malicious URLs. However, many graph-based systems primarily focus on graph embeddings and classification performance while providing limited probabilistic reasoning capabilities.

Probabilistic graphical models provide an alternative approach for handling uncertainty and collective inference in relational systems. Loopy Belief Propagation (LBP) has been widely used as an approximate inference algorithm for cyclic graph structures in domains such as network analysis, computer vision, and cybersecurity. LBP enables iterative message passing among connected nodes, allowing probabilistic beliefs to be refined based on neighbouring relationships. Despite its effectiveness, limited research has integrated LBP-based collective inference specifically for large-scale phishing URL detection combined with real-time probabilistic analysis.

The limitations of existing phishing detection systems highlight the need for a framework that combines machine learning, graph-based modelling, and probabilistic inference within a unified architecture. To address these challenges, the proposed system, **PhiURL**, integrates feature-based machine learning, relational graph construction, and Loopy Belief Propagation to improve phishing detection accuracy and support scalable real-time phishing analysis.

III. SYSTEM ARCHITECTURE

The proposed **PhiURL** framework is designed as an integrated phishing URL detection system that combines machine learning, graph-based relational modelling, and probabilistic inference to identify malicious URLs with improved accuracy and scalability. The architecture consists of multiple interconnected components that collectively perform dataset preprocessing, feature analysis, probabilistic classification, graph construction, collective inference, and real-time phishing prediction.

The detection process begins with the dataset collection and preprocessing stage, where phishing and legitimate URLs are gathered from trusted repositories and publicly available sources. The collected data undergoes normalization and cleaning procedures, including duplicate removal, invalid URL filtering, and inconsistency handling, to ensure high-quality input for subsequent analysis. This preprocessing stage improves data reliability and minimizes noise during model training and inference.

Following preprocessing, the feature extraction engine analyses each URL by extracting lexical, domain-based, TF-IDF, and statistical features. These features capture critical phishing-related characteristics such as suspicious keywords, abnormal URL structures, domain irregularities, and behavioural patterns commonly observed in phishing campaigns. The extracted feature vectors are then provided to the Random Forest classifier, which estimates the initial phishing probability for each URL. These probabilities serve as unary belief values that represent the preliminary confidence associated with phishing or legitimate behaviour.

To model hidden dependencies among URLs, the graph construction component represents URLs as interconnected graph nodes linked through similarity-based weighted edges. Relationships among URLs are established using structural similarities, lexical correlations, and behavioural dependencies. The pairwise potential modelling stage further computes

relational influence scores between connected nodes, enabling the system to capture collective phishing behaviour rather than relying solely on isolated URL analysis.

The central component of the proposed architecture is the **Loopy Belief Propagation (LBP) Inference Engine**, which performs iterative probabilistic message passing across the constructed graph structure. During each iteration, neighbouring nodes exchange belief information to refine phishing probabilities collectively. This collective inference process enables the framework to propagate phishing evidence across related URLs, thereby improving detection consistency and robustness in complex graph environments. By integrating machine learning-based unary probabilities with graph-driven relational inference, the LBP engine significantly enhances the overall phishing detection capability of the system.

After convergence of the inference process, the prediction and classification layer generates the final phishing verdict and confidence score for each analyzed URL. The effectiveness of the proposed framework is evaluated using standard performance metrics including accuracy, precision, recall, F1-score, confusion matrix analysis, and ROC curve evaluation.

Finally, the real-time prediction interface enables users to perform live phishing analysis by submitting URLs through the detection platform. The interface displays phishing probabilities, classification outcomes, and threat visualization results generated by the PhiURL framework, thereby supporting efficient, scalable, and user-friendly real-time phishing detection.

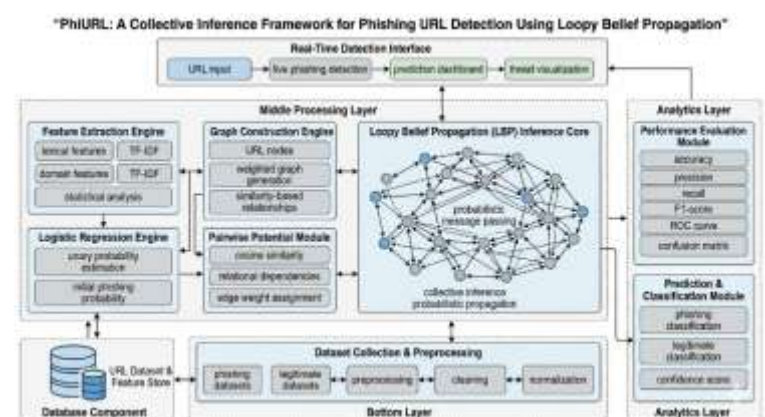


Fig. 1. System architecture of the proposed PhiURL framework for phishing URL detection using Loopy Belief Propagation

IV. PROPOSED METHODOLOGY

The proposed PhiURL framework performs phishing URL detection through a collective inference pipeline that integrates feature-based machine learning, graph construction, and probabilistic message passing. Instead of classifying URLs independently, the framework models relational dependencies among URLs using a probabilistic graph structure and refines phishing predictions through Loopy Belief Propagation (LBP).

A. Feature Extraction

The detection process begins with feature extraction, where multiple lexical, domain-based, TF-IDF, and statistical attributes are extracted from each URL. Lexical features capture characteristics such as URL

length, special symbols, suspicious keywords, and abnormal structural patterns commonly associated with phishing URLs. Domain-based features analyse properties including HTTPS usage, domain information, and URL hierarchy, while TF-IDF and statistical indicators capture textual and behavioural patterns within the dataset.

The extracted feature vectors provide a numerical representation of each URL and serve as the input for the initial phishing probability estimation stage.

B. Unary Potential Estimation Using Random Forest

After feature extraction, a Random Forest classifier is utilized to estimate the initial phishing probability for each URL. These probabilities represent unary potentials within the probabilistic graphical model and define the initial belief associated with each graph node.

The unary potential of a URL node is represented as:

$$\phi_i(x_i) = P(y_i | x_i)$$

where $\phi_i(x_i)$ denotes the local phishing belief generated from the extracted feature vector x_i . These initial probabilities provide the foundational evidence for subsequent graph-based collective inference.

C. Graph Construction and Pairwise Potential Modeling

To capture hidden dependencies among URLs, the framework represents URLs as nodes within an undirected probabilistic graph. Similar URLs are connected through weighted edges based on lexical similarity, structural relationships, and behavioural correlations observed within the dataset. Pairwise potential functions are used to model relational influence between neighbouring nodes. These potentials define how strongly the phishing behaviour of one URL influences connected URLs within the graph structure.

The pairwise potential is represented as:

$$\psi_{ij}(x_i, x_j) = \exp(\lambda \cdot \text{sim}(x_i, x_j))$$

where $\text{sim}(x_i, x_j)$ represents the similarity relationship between two connected URL nodes and λ controls the influence strength of neighbouring beliefs.

D. Loopy Belief Propagation Framework

After graph construction, probabilistic inference is performed using Loopy Belief Propagation (LBP). The LBP framework iteratively propagates belief information across interconnected graph nodes through message passing operations. During each iteration, neighbouring nodes exchange probabilistic evidence, allowing phishing beliefs to be refined collectively based on both local feature evidence and relational dependencies.

The message passed from node i to node j is defined as:

$$m_{i \rightarrow j}(x_j) = \sum_{x_i} \phi_i(x_i) \psi_{ij}(x_i, x_j) \prod_{k \in N(i) \setminus j} m_{k \rightarrow i}(x_i)$$

where $\phi_i(x_i)$ denotes the unary potential, $\psi_{ij}(x_i, x_j)$ represents the pairwise potential, and $N(i)$ indicates the neighbouring nodes connected to node i .

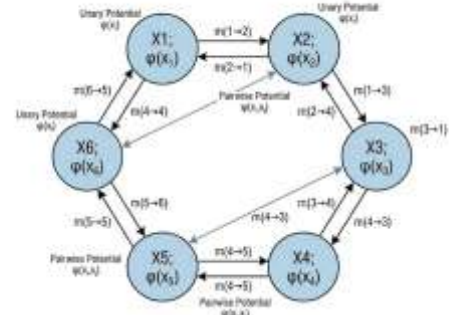


Fig. 2. Loopy Belief Propagation inference

Through iterative message propagation, the framework updates node beliefs until convergence is achieved. This collective inference process enables the system to propagate phishing evidence across related URLs, thereby improving detection robustness and reducing isolated classification errors.

V. PROPOSED METHODOLOGY

A. Dataset Configuration

Dataset Category	Number of URLs
Phishing URLs	1,00,945
Legitimate URLs	1,34,850
Total URLs	2,35,795

Table 1. Dataset distribution used for experimental analysis.

Table 1 presents the dataset distribution used during experimentation.

B. Model Configuration

During experimentation, extracted URL features were provided as input to the Random Forest classifier to generate initial phishing probabilities. These probabilities were then integrated into the graph-based inference process.

The major experimental configurations are summarized below:

- Random Forest used for unary probability estimation.
- Similarity-based weighted graph construction performed using extracted URL features.
- Loopy Belief Propagation (LBP) executed iteratively for collective inference.
- Maximum inference iterations applied until convergence conditions were satisfied.
- Performance evaluated using standard classification metrics.

VI. SIMULATION RESULTS

In this section, we conduct a comprehensive evaluation of the proposed PhiURL framework for phishing URL detection using graph-based collective inference and Loopy Belief Propagation. The implementation of the proposed framework was carried out using Python-based machine learning and graph processing libraries. The experimental setup and evaluation results are presented using multiple performance metrics and visualization techniques.

The proposed framework was implemented using Python within the Jupyter Notebook and Visual Studio Code environment. The experiments were performed using Scikit-learn for Random Forest-based phishing probability estimation and NetworkX for graph construction and relational modelling.

Additional libraries including Pandas, NumPy, and Matplotlib were utilized for data preprocessing, numerical computation, and visualization purposes. The extracted feature vectors were first provided to the Random Forest classifier to generate initial phishing probabilities. These probabilities were then integrated into the graph inference stage, where URLs were represented as interconnected nodes linked through similarity-based weighted edges. The Loopy Belief Propagation (LBP) framework iteratively refined node beliefs through probabilistic message passing until convergence conditions were satisfied.

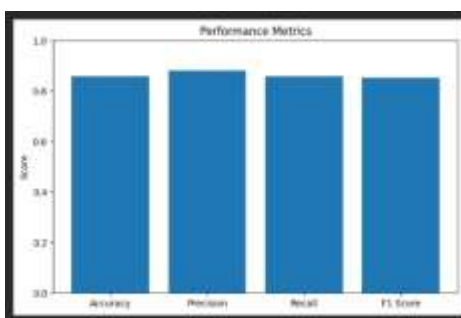


Fig. 3. Performance Metrics

Fig. 3 illustrates the performance metrics obtained using the proposed PhiURL framework. The framework achieved high classification performance across accuracy, precision, recall, and F1-score metrics, demonstrating the effectiveness of integrating graph-based collective inference with machine learning-based phishing detection.

The benchmark accuracy comparison between traditional machine learning models and the proposed framework is shown in Fig. 4. Logistic Regression achieved comparatively lower classification performance due to isolated URL prediction behaviour. Random Forest improved phishing detection

performance through ensemble learning; however, the integration of Loopy Belief Propagation further enhanced classification consistency by propagating phishing evidence

among related URL nodes.

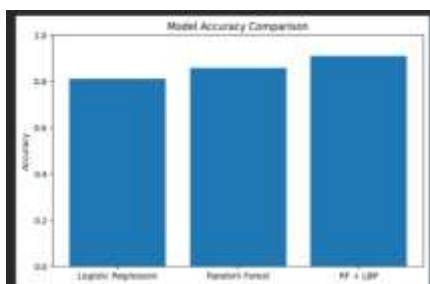


Fig. 4. Model Accuracy

To further evaluate the classification capability of the framework, ROC curve analysis was performed using the True Positive Rate (TPR) and False Positive Rate (FPR). The ROC curve shown in Fig. 5 demonstrates strong classification

performance with an AUC value of approximately 0.91, indicating effective separation between phishing and legitimate URLs across different threshold values.

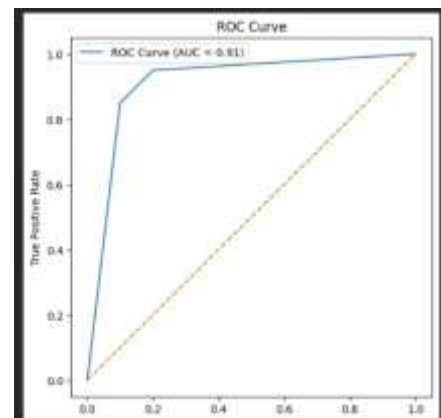


Fig. 5. ROC Curve

The confusion matrix generated for the proposed framework is presented in Fig. 6. The results indicate that the framework correctly classified a large number of phishing URLs while maintaining relatively low false negative rates. The integration of graph-based collective inference improved probabilistic consistency among neighbouring URLs and reduced isolated prediction errors.

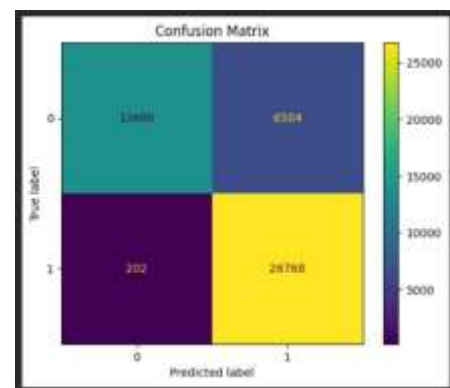


Fig. 6. Confusion Matrix

Table 2 presents the benchmark comparison between traditional machine learning models and the proposed RF + LBP framework.

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Logistic Regression	81.00	79.71	73.57	76.14
Random Forest	86.00	87.76	83.58	85.58
Proposed RF + LBP	91.00	94.12	93.35	94.73

Table 2. Benchmark Comparison

From the obtained results, it can be observed that the proposed PhiURL framework outperformed traditional machine learning approaches across all evaluation metrics. The incorporation of graph-based relational modelling and probabilistic message passing significantly improved phishing detection capability by enabling collective inference across interconnected URL nodes.

The experimental analysis further demonstrates that Loopy Belief Propagation enhances phishing prediction stability by iteratively refining node beliefs using neighbouring URL relationships. This

collective inference mechanism improves classification robustness and enables the framework to effectively detect phishing URLs under large-scale cybersecurity environments.

VII. CONCLUSION

In this paper, we developed PhiURL, a phishing URL detection framework leveraging Loopy Belief Propagation, graph-based relational modelling, and machine learning-based probabilistic inference. The proposed framework integrates lexical, domain-based, TF-IDF, and statistical URL features to improve phishing detection accuracy and classification stability. URLs are represented as interconnected graph nodes, where similarity-based pairwise relationships enable collective inference across the graph structure. Experimental results demonstrate that integrating Random Forest classification with Loopy Belief Propagation significantly improves phishing detection performance compared to traditional machine learning approaches.

The proposed framework achieved high accuracy, precision, recall, and F1-score while maintaining effectiveness on large-scale phishing datasets. Furthermore, the real-time prediction capability and probabilistic graph inference mechanism make the proposed PhiURL framework suitable for scalable cybersecurity and phishing detection applications.

VIII. REFERENCES

1. Efficient Phishing URL Detection Using Graph Based Machine Learning and Loopy Belief Propagation.
2. Wenye Guo, Qun Wang, Hao Yue, Haijian Sun, and Rose Qingyang Hu, "Efficient Phishing URL Detection Using Graph-Based Machine Learning and Loopy Belief Propagation," arXiv preprint, Jan. 2025
3. A systematic literature review on phishing website detection techniques Safi and S. Singh, "A Systematic Literature Review on Phishing Website Detection
4. Techniques," J. King Saud Univ. Comput. Inf. Sci., vol. 35, no. 2, pp. 590–611, Jan. 2023.
5. A Survey of Machine Learning-Based Solutions for Phishing Website Detection
6. L. Tang and Q. H. Mahmoud., "A Survey of Machine Learning-Based Solutions for Phishing Website Detection," Information, vol. 3, no. 3, p. 34, 2023.
<https://www.mdpi.com/2504-4990/3/3/34>
7. Phishing Website Detection Based on Multidimensional Features Driven by Deep Learning
8. P. Yang, G. Zhao and P. Zeng, "Phishing Website Detection Based on Multidimensional Features Driven by Deep Learning," IEEE Access, 2018.
<https://ieeexplore.ieee.org/abstract/document/8610190>
9. A graph-theoretic approach for the detection of phishing webpages
10. C. C. L. Tan, K. L. Chiew, K. S. C. Yong, S. N. Sze, J. Abdullah, and Y. Sebastian, "A Graph-Theoretic Approach for the Detection of Phishing Webpages," Computers & Security, vol. 92, 2020.
11. Web Phishing detection based on graph mining F. Zou, Y. Gang, B. Pei, L. Pan and L. Li, "Web Phishing Detection Based on Graph Mining," 2016 IEEE Conference Convolutional Graph Network-Based Feature Extraction to Detect Phishing Attacks
12. M. U. Younis, M. A. Azam, M. U. Ali and H. Jamjoom, "Convolutional Graph Network- Based Feature Extraction to Detect Phishing Attacks," Future Internet, vol. 17, no. 8, p. 331, 2024, doi: 10.3390/fi17080331.
13. Security Optimization of Dynamic Networks with Probabilistic Graph Modeling and Linear Programming
14. Z. Xiong, J. Wang, and L. Li, "Security Optimization of Dynamic Networks with
15. Probabilistic Graph Modeling and Linear Programming," in Proc. 2015 IEEE International Conference on Communication Software and Networks (ICCSN),
16. Loopy Belief Propagation for Approximate Inference: An Empirical Study
17. K. P. Murphy, Y. Weiss, and M. I. Jordan, "Loopy Belief Propagation for Approximate Inference: An Empirical Study," Proceedings of UAI, 1999.
18. Loopy Belief Propagation: Convergence and Effects of Message Errors
19. D. A. Ihler, J. W. Fisher III, and A. S. Willsky, "Loopy Belief Propagation: Convergence and Effects of Message Errors," J. Mach. Learn. Res., vol. 6, pp. 905–936, 2005.
<https://www.jmlr.org/papers/volume6/ihler05a/ihler05a.pdf>
20. ZooBP: Belief Propagation for Heterogeneous NetworksM. Eswaran, B. Zong, C. C. Aggarwal, H. Park, S. Parthasarathy and Y. Yang,, "ZooBP: Belief Propagation for Heterogeneous Networks," Proc. VLDB Endow., vol. 10, no. 6, pp. 625–636, 2017.
21. Lexical feature based phishing URL detection using online learning.
22. P. Kumar, S. K. Singh, A. K. Tripathi and A. Mehfuz, "Lexical Feature Based Phishing URL Detection Using Online Learning," Proc. ACM Conference, 2011.
<https://dl.acm.org/doi/abs/10.1145/1866423.1866434>