Rutuja Hole, 2026, 14:1
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science, Engineering and Technology
An Open Access Journal

PhiURL – Graph based phishing url detection using Loopy Belief Propogation

Rutuja Hole, Riya Dandawate, Rohit Patil
SPPU - Information Technology

Abstract - Phishing attacks continue to pose a significant threat to online security by exploiting user trust through deceptive URLs and malicious web resources. Traditional phishing detection approaches largely rely on isolated feature-based classification techniques, which fail to capture the relational dependencies that naturally exist among web entities such as domains, URLs, and structural attributes. This paper presents a theory-driven framework that models phishing detection as a probabilistic inference problem over a graph structure. URLs and their associated characteristics are represented as interconnected nodes within a graphical model, enabling relational reasoning across the network. To infer the likelihood of phishing behaviour, Loopy Belief Propagation (LBP) is employed as an approximate probabilistic inference mechanism capable of handling cyclic graph structures. The proposed framework emphasizes formal graph construction, probabilistic modelling, and message-passing inference without reliance on implementation-specific heuristics. By reasoning collectively over relational dependencies, the model provides a robust theoretical foundation for phishing detection under uncertainty. This work contributes a formalized approach that bridges graph- based learning and probabilistic inference for cybersecurity applications.

Keywords - Phishing Detection, Graph-Based Learning, Probabilistic Graphical Models, Loopy Belief Propagation, Cybersecurity, URL Analysis.

I. INTRODUCTION

Phishing remains one of the most prevalent and effective cyberattack vectors, targeting users through deceptive URLs that imitate legitimate online services. As phishing techniques evolve rapidly, static detection mechanisms such as blacklists and rule-based filters struggle to provide timely and accurate protection. Machine learning-based approaches have improved detection capabilities by learning patterns from historical data; however, most existing models treat URLs as independent instances, ignoring the relational structure inherent in the web ecosystem.

In practice, phishing URLs often share common characteristics, hosting infrastructure, or domain-level behaviours that form implicit relationships among them. Modelling such dependencies requires a representation that extends beyond traditional feature vectors. Graph- based learning offers a natural abstraction for capturing these relationships by representing URLs and their attributes as nodes connected through meaningful edges.

To reason over such graph structures under uncertainty, probabilistic graphical models provide a principled framework. In particular, Loopy Belief Propagation (LBP) enables approximate inference in graphs containing cycles, making it suitable for complex relational domains such as web security. This paper proposes a theoretical framework that formulates phishing detection as a probabilistic inference problem over a graph, leveraging LBP to propagate belief information across interconnected entities.

The primary contribution of this work lies in formalizing the graph representation, probabilistic modelling, and inference process, thereby providing

a theoretical foundation for relational phishing detection systems.

Contributions:

This work makes the following contributions.

(1) It formulates phishing URL detection as a collective inference problem over a relational graph.(2) It presents a formal probabilistic graphical model integrating unary and pairwise dependencies. (3) It applies Loopy Belief Propagation as an approximate inference mechanism for cyclic graph structures in phishing detection.

Related Work

Early phishing detection techniques relied heavily on signature-based methods and manually crafted rules, which were effective only against known attack patterns. Machine learning approaches later introduced automated feature learning using lexical, structural, and host-based attributes. While these methods improved detection accuracy, they largely operated under the assumption of independent observations.

Recent research has explored graph-based representations for cybersecurity problems, including malware detection, intrusion analysis, and domain reputation modelling. These approaches demonstrate that relational information can significantly enhance detection performance. However, many graph-based phishing detection systems focus on empirical performance and implementation-specific optimizations, offering limited theoretical insight into the underlying inference mechanisms.

Probabilistic graphical models, such as Markov Random Fields and Bayesian Networks, provide a mathematically grounded framework for modelling uncertainty and dependencies among variables. Belief Propagation has been successfully applied in domains including error correction, computer vision, and network analysis. Despite its approximate nature in loopy graphs, LBP remains a powerful inference tool when exact inference is intractable.

This work differentiates itself by emphasizing a formal probabilistic formulation of phishing detection using graph-based inference, rather than focusing on dataset-specific engineering choices.

Collective inference and graph-based learning have been widely explored in cybersecurity domains where relational dependencies play a critical role. Graph-based malware detection techniques have demonstrated that modelling interactions among files, processes, or network entities can significantly improve robustness compared to independent classification approaches [1], [5], [6]. Similarly, collective classification methods leverage correlations among linked instances to refine predictions in relational settings, enabling label propagation and consistency across connected entities [2]. Probabilistic graphical models provide a principled foundation for such collective reasoning [8], [9]. Early work on belief propagation established message-passing algorithms for inference in graphical models, with subsequent studies extending these techniques to loopy graph structures where exact inference is intractable [9],[11]. These advances motivate the use of Loopy Belief Propagation as an effective approximate inference mechanism for relational phishing detection.

Limitations of Feature-Independent Phishing Detection

Traditional phishing detection models assume conditional independence among URLs, treating each instance as an isolated observation. While this assumption simplifies model design, it neglects the collective behaviour exhibited by phishing campaigns, where multiple URLs often share infrastructure, lexical patterns, or domain-level characteristics.

Such independence assumptions limit robustness, particularly in adversarial settings where attackers intentionally manipulate individual features to evade detection.

Graph-Based Learning in Cybersecurity

Graph-based learning has emerged as a powerful paradigm for modelling relational data in cybersecurity. By representing entities and their interactions explicitly, graph models enable

reasoning over structural dependencies that are otherwise inaccessible to flat feature representations. Prior research has demonstrated the effectiveness of graph modelling in malware detection, botnet identification, and network intrusion analysis.

However, many existing approaches emphasize empirical performance while offering limited formal analysis of inference mechanisms.

Probabilistic Inference on Graphs

Probabilistic graphical models provide a mathematically grounded approach for handling uncertainty in relational systems. Inference algorithms such as Belief Propagation enable the computation of marginal distributions over interconnected variables. While exact inference is tractable only for acyclic graphs, approximate methods such as Loopy Belief Propagation have proven effective in practice for cyclic structures.

This work builds upon these foundations by formalizing phishing detection as a probabilistic inference task over a graph.

Problem Formulation

Let $U = \{u1, u2, \ldots, un\}$ denote a set of URLs under consideration. Each URL is associated with a latent label Xi, where $Xi = 1$denotes a phishing URL and $Xi = 0$denotes a legitimate URL. Equivalently, $Xi \in$ {Phishing,Legitimate}. Each random variable Xirepresents the phishing state of the i-th URL. Unlike independent classification, the phishing likelihood of a URL may depend on its relationships with other URLs or shared characteristics. Therefore, the problem is formulated as collective inference over a graph structure rather than individual prediction.

Graph Construction Model

The detection framework is modelled as an undirected graph $G = (V, E)$, where each node $vi \in V$corresponds to a URL or an associated feature entity. Edges represent relational dependencies such as shared structural patterns, domain-level associations, or similarity relationships. While the graph may include feature and domain entities for construction, probabilistic inference is performed over URL label variables, with auxiliary nodes influencing dependencies.

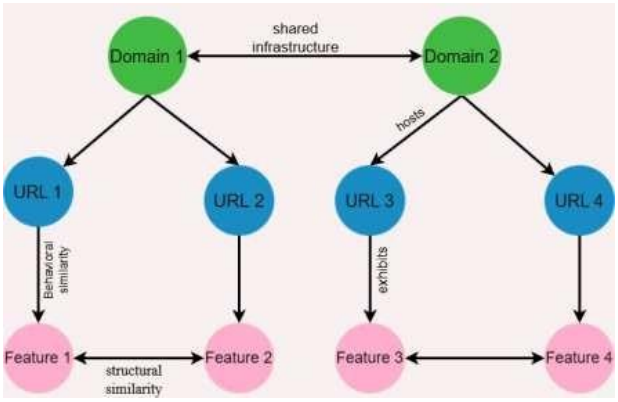

Figure 1 illustrates the heterogeneous graph structure capturing relational dependencies among domains, URLs, and structural features.

The graph may be heterogeneous, consisting of multiple node types, or homogeneous, depending on the abstraction level. Edge weights encode the strength of dependency between connected nodes, reflecting how strongly the state of one node influences another. This relational structure enables the model to propagate information beyond local observations.

Probabilistic Graphical Modelling

Each node in the graph is treated as a random variable with an associated probability distribution. The joint probability distribution over all node labels is defined using a pairwise Markov Random Field formulation:

$$P(X) = \frac{1}{Z} \prod_{i} \phi_i(X_i) \prod_{(i,j)\in E} \phi_{ij}(X_i, X_j)$$

where:

$\phi i(Xi)$represents the unary potential capturing local evidence,

$\phi ij(Xi, Xj)$represents the pairwise potential encoding relational influence,

$Z$is the partition function ensuring normalization.

This formulation allows the integration of both individual URL characteristics and relational dependencies within a unified probabilistic framework.

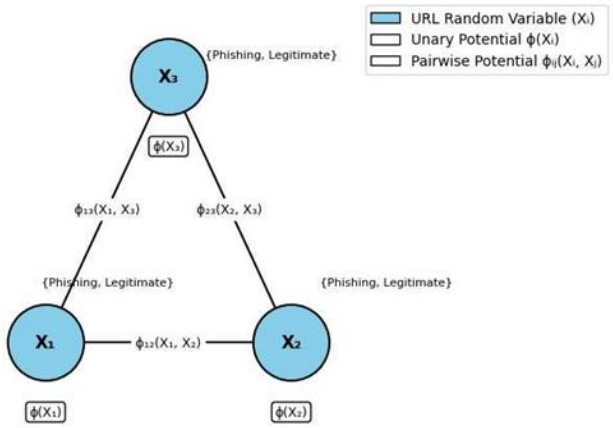

Figure 2 : Probabilistic Graphical Model

As shown in Figure 2, the joint distribution is modelled using unary and pairwise potential functions within a Markov Random Field formulation.

Loopy Belief Propagation for Inference

Exact inference in loopy graphs is computationally intractable; therefore, approximate inference is performed using Loopy Belief Propagation. LBP operates by iteratively exchanging messages between neighbouring nodes to update belief estimates.

The message from node *i* to node *j* at iteration *t* is defined as:

$$m_{i\to j}^{(t)}(x_j) = \sum_{x_i} \phi_i(x_i)\,\phi_{ij}(x_i, x_j) \prod_{k\in N(i)\setminus j} m_{k\to i}^{(t-1)}(x_i)$$

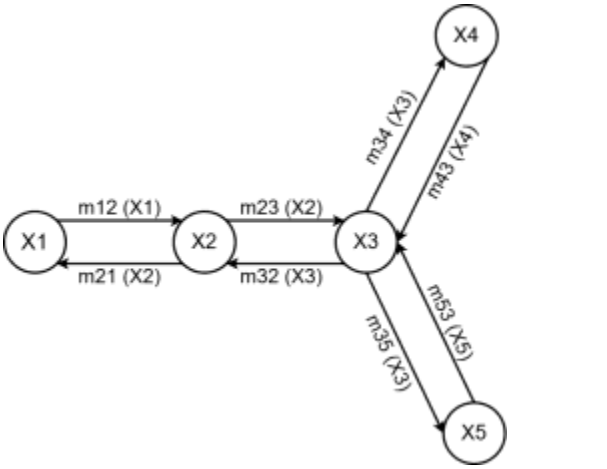

Figure 3 : Loopy Belief Propogation

Figure 3 depicts the iterative message-passing process employed by Loopy Belief Propagation across cyclic graph structures.

Belief estimates are updated until convergence or a predefined stopping criterion is met. Although convergence is not guaranteed in loopy graphs, empirical and theoretical studies suggest that LBP often yields accurate marginal approximations in practice.

Approximate Inference in Loopy Graphs

Loopy Belief Propagation operates without guarantees of convergence in general graphs. Despite this limitation, extensive empirical and theoretical studies suggest that LBP often converges to useful fixed points in practice. In phishing detection, the redundancy of relational evidence contributes to inference stability.

Role of Collective Reinforcement

LBP enables collective reinforcement, where consistent evidence across related URLs amplifies belief confidence, while conflicting evidence is smoothed through probabilistic averaging. This property is particularly valuable when individual URLs provide weak or ambiguous signals.

Illustrative Message-Passing Example

Consider a small graph of three URLs connected by shared structural features. Initial unary potentials encode local phishing evidence, while pairwise potentials represent similarity relationships. During LBP, messages propagate iteratively between URLs, reinforcing consistent phishing signals and smoothing conflicting evidence. This process converges to stable marginal probabilities reflecting collective inference.

Convergence and Complexity Analysis

The computational complexity of LBP depends on the number of nodes, edges, and possible label states. Each iteration involves message updates proportional to the number of edges in the graph. While convergence behaviour varies based on graph topology and potential functions, damping strategies can improve stability.

From a theoretical standpoint, the framework scales linearly with graph size per iteration, making it suitable for large relational datasets when approximate inference is acceptable.

Discussion

The proposed framework highlights the importance of relational reasoning in phishing detection. By modelling URLs within a probabilistic graph, the system can infer phishing likelihoods even when individual evidence is weak or incomplete. Unlike traditional classifiers, the approach captures collective behaviour, which is particularly valuable in adversarial environments.

While the framework is theoretical, it provides a foundation upon which practical systems can be built. Limitations include approximate inference accuracy and sensitivity to graph construction choices, which present opportunities for further research.

II. CONCLUSION AND FUTURE WORK

This paper presented a theory-driven framework for phishing URL detection using graph-based probabilistic modelling and Loopy Belief Propagation. By formalizing phishing detection as a collective inference problem, the approach moves beyond isolated feature analysis and leverages relational dependencies among web entities.

Future work may extend this framework to dynamic graphs, incorporate temporal dependencies, or explore hybrid inference techniques that combine probabilistic reasoning with deep learning. The theoretical foundation established here supports the development of more robust and adaptive cybersecurity systems.

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