Graph Unlearning: Mechanism and Future Direction for Machine Unlearning with Complex Relationships

Chang Qi *

Posted Date: 23 September 2024

doi: 10.20944/preprints202409.1733.v1

Keywords: Graph Unlearning; Knowledge Graph Embeddings; Machine Learning; Privacy Preserving; Digital Asset

Article

Graph Unlearning: Mechanism and Future Direction for Machine Unlearning with Complex Relationships

Chang Qi

University of Washington, Seattle, WA; changchannieqi1@gmail.com

Abstract: Graph unlearning has emerged as a crucial technique in privacy-preserving applications, particularly in scenarios where sensitive data must be erased from graph-based systems. This paper explores the fundamental concepts, challenges, and methodologies associated with graph unlearning, including its application to domains such as social networks, financial transaction systems, and healthcare networks. By delving into various algorithms, including local and global unlearning methods, we analyze the trade-offs between accuracy, privacy, and computational efficiency. The paper further investigates the security risks associated with graph unlearning, including adversarial and inference attacks, and proposes mitigation strategies to safeguard unlearning processes. We also examine the unique challenges posed by federated learning systems, where unlearning requires coordination across decentralized clients. Finally, the paper evaluates graph unlearning techniques using performance metrics such as time complexity, accuracy, and privacy guarantees, supported by real-world examples from diverse applications. The findings emphasize the importance of developing robust, scalable, and secure unlearning mechanisms to ensure data privacy and compliance with regulations. Evaluation and metrics for Graph Unlearning will also be discussed in this paper.

Keywords: Graph Unlearning; Knowledge Graph Embeddings; Machine Learning; Privacy Preserving; Digital Asset

1. Introduction

In the era of data-driven technologies, vast amounts of information are stored and processed in the form of graphs, ranging from social networks and knowledge graphs to financial transaction systems. As graph data grows in importance, so too does the need for effective mechanisms to manage and modify this information. One critical challenge is the ability to "unlearn" specific data from a graph — a process that involves selectively removing nodes, edges, or subgraphs to comply with privacy regulations, user requests, or corporate policies. This concept, known as graph unlearning, is gaining traction in both academic and industrial research due to increasing concerns about data privacy and the right to be forgotten.

Graph unlearning focuses on the structured and often interdependent nature of graph data, where the removal of one element can impact the entire system. It goes beyond simple data deletion, requiring sophisticated algorithms to ensure that the process does not compromise the integrity of the remaining data or the model's performance. Furthermore, graph unlearning must be efficient, scalable, and secure, given the vast and complex nature of real-world graphs.

The motivation for graph unlearning stems from various sources, including legal frameworks such as the General Data Protection Regulation (GDPR), which grants individuals the right to request the removal of personal information. In networked data, such as social media platforms or blockchain transactions, ensuring compliance with such regulations poses unique challenges. This paper explores the emerging field of graph unlearning, delving into the technical, legal, and ethical considerations that drive its development. The primary focus will be on understanding the core algorithms and methodologies that enable graph unlearning, as well as the security and privacy implications of these processes in modern applications.

2. Literature Review

Graphs, as a data structure, represent relationships between entities in numerous domains, including social networks, knowledge bases, biological systems, and financial networks. Each graph

 doi:10.20944/preprints202409.1733.v1

comprises nodes (or vertices) and edges (links between nodes), capturing intricate relationships between entities. Over the years, graphs have played a pivotal role in machine learning and artificial intelligence (AI), especially in applications like recommender systems, fraud detection, and network analysis.

The concept of unlearning in machine learning refers to the selective removal of specific data points, ensuring that they no longer influence model predictions or analysis outcomes. While unlearning is well understood in the context of traditional machine learning models like decision trees or neural networks, applying this to graph structures is significantly more complex due to the interdependent nature of graph nodes and edges. Unlearning a single node may necessitate revising adjacent nodes and relationships, leading to complex cascading effects.

Research in data privacy, especially with the advent of the General Data Protection Regulation (GDPR), has driven the need for efficient unlearning mechanisms. This need is exacerbated in graph data where the erasure of one node or edge could impact the global structure. Related work in data deletion, differential privacy, and adversarial robustness provides a foundation, but applying these techniques to graph data remains a challenge. Emerging work in graph learning, including Graph Neural Networks (GNNs), has only recently begun addressing privacy-preserving and unlearning mechanisms. [1] has provided a summary of the most relevant research on federated unlearning. [2] has proved that GNN is very successful in representing complex relationships in machine learning. When GNN framework is combined with treasury [3] and crypto trading [4], it becomes very powerful.

[5] [6] [7] use unique methods called PROJECTOR and GRAPHEDITOR. In PROJECTOR [5], it uses projection techniques to remove specific nodes, ensuring no trace in the model parameters. In GRAPHEDITOR [6], it manages dynamic graphs and enables node/edge deletion and feature updates. The next major categories is the guaranteeing certified unlearning. The most famous is the CEU framework [8] [9], which introduces a single-step update methodology for the removal of specific edges [1].

The diverse in graph unlearning is further exemplified by GUIDE [10], FedLU [11] and GNNDELETE [12]. [10] uses inductive graph unlearning in dynamic graphs. [11] features federated learning for knowledge-graph embedding for data heterogeneity. [12] introduces optimization strategies for node/edge deletions without loss of knowledge. The optimization strategies introduced in [12] is like the one used in self-supervised learning in [13].

[14] introduces Graph Scattering Transform which focus on mathematical robustness in unlearning, while [15] proposes Graph Influence Function that emphasizes influence function in unlearning. [16] proposes GraphEraser which stresses efficient partitioning mechanisms for unlearning in graph data.

3. Graph Unlearning: Fundamentals and Concepts

Graph unlearning refers to the process of removing specific data points, such as nodes, edges, or subgraphs, from a graph and ensuring that any machine learning model trained on the graph no longer retains knowledge of the removed elements. This process is crucial in applications where compliance with privacy regulations is essential, such as removing sensitive user data from social networks or confidential transactions from financial networks. One key challenge in graph unlearning is dealing with the interdependencies inherent in graph structures. Deleting a node may impact connected nodes and edges, requiring careful management of cascading effects. Another challenge is preserving the graph's integrity; for instance, removing a single node in a social network could disrupt clusters of users or communities, creating fragmentation in the network. Unlearning must be efficient, particularly in large-scale graphs, where millions of nodes and edges are interconnected. Real-world scenarios for graph unlearning often arise from user requests to delete personal data in compliance with legal frameworks like the General Data Protection Regulation (GDPR). In financial networks, unlearning is necessary when removing records of sensitive or erroneous transactions without disrupting the overall system’s functionality.

 doi:10.20944/preprints202409.1733.v1

4. Algorithms and Techniques for Graph Unlearning

Various algorithms have been proposed to facilitate efficient and accurate graph unlearning. Traditional approaches rely on simple node or edge deletion, which, while conceptually straightforward, often leaves traces of the removed elements in machine learning models that were previously trained on the graph. These residual traces could still infer the deleted information, rendering the deletion incomplete from a privacy perspective. To address this, more sophisticated algorithms have emerged. These include techniques like node splitting, where a node's influence is redistributed to neighboring nodes before deletion, and subgraph partitioning, which isolates the deleted subgraph from the main graph. Additionally, random edge rewiring ensures that graph properties like clustering coefficients and degree distribution remain unchanged post-unlearning.

Unlearning approaches can be categorized into two broad types: local and global methods. Local methods target specific nodes or edges for removal while ensuring that the graph's surrounding structure remains intact. These methods are particularly useful for applications focused on individual privacy, such as when a user requests deletion from a social network. Global methods, on the other hand, involve structural modifications to the entire graph, often through reconfiguring clusters or recalculating metrics like betweenness centrality. These methods are more suited to large-scale unlearning tasks where significant parts of the graph need to be altered. Balancing accuracy and privacy is crucial, as unlearning must completely erase traces of deleted data while preserving the overall graph's functionality.

5. Graph Unlearning in Privacy-Preserving Applications

Graph unlearning has significant implications in privacy-preserving applications, where sensitive information needs to be erased to comply with privacy laws or ensure data security. One of the most prominent use cases is social networks, where user data is represented as nodes connected through edges representing friendships, interactions, or content sharing. When a user requests their data to be deleted, all relevant nodes and edges must be removed from the graph without disrupting the structure for other users. Graph unlearning algorithms enable platforms like Facebook, Twitter, or LinkedIn to erase user data efficiently while maintaining the functionality and integrity of the larger social graph.

Financial transaction networks also benefit from graph unlearning techniques. These graphs, which capture relationships between different financial entities based on transactions, often need to remove sensitive data while preserving the network's analytical capabilities. For example, when regulatory policies mandate the deletion of certain transactional records, the system must unlearn those records without affecting the broader analysis or reporting functions. Blockchain networks present a unique challenge here, as their decentralized and immutable nature makes unlearning more complex.

In healthcare networks, graph unlearning is crucial when dealing with patient data. A patient's medical history may need to be removed, especially in the context of data privacy regulations like HIPAA. In such cases, the graph unlearning algorithm needs to eliminate all traces of the patient from the network, which often involves removing nodes representing treatments, diagnoses, or medical professionals they interacted with. At the same time, the remaining healthcare network must continue to function for other patients.

6. Security and Threats in Graph Unlearning

Graph unlearning, while essential for privacy, introduces several security risks that need to be addressed. One significant threat comes from adversarial attacks on unlearning algorithms. In these attacks, adversaries attempt to exploit vulnerabilities in the unlearning process to either access sensitive information or disrupt the integrity of the graph. For instance, attackers could inject noise or corrupt adjacent nodes to interfere with the deletion process, leaving behind residual traces of supposedly unlearned data. These attacks can lead to severe privacy violations, especially in applications like social networks or financial networks.

 doi:10.20944/preprints202409.1733.v1

Inference attacks are another significant threat to graph unlearning. In such attacks, an adversary might analyze the structure of the graph surrounding the deleted node or edge to infer the missing information. This is particularly concerning in large, interconnected networks, where the removal of one element could still leave enough clues for an attacker to reconstruct sensitive data. For example, even if a node representing a financial transaction is removed, the remaining transactional relationships might still provide enough context for an attacker to infer the deleted information.

To mitigate these threats, secure unlearning protocols must be developed. Techniques such as differential privacy, which introduces noise to the data to obscure sensitive information, and cryptographic methods that ensure data integrity can help protect against inference and adversarial attacks. Additionally, using randomized edge removal or partitioning strategies can further strengthen the security of graph unlearning algorithms.

7. Federated Learning and Graph Unlearning

Federated learning, which enables multiple clients to collaboratively train a machine learning model without sharing raw data, introduces additional challenges for graph unlearning. In federated graph learning, each client typically operates on a local graph, sharing only model updates with a central server. When it comes to unlearning, the challenge lies in ensuring that any data removed locally from one client's graph is also reflected in the global model shared across all clients. This becomes particularly complex when unlearning affects nodes or edges that are integral to the structure of multiple clients' graphs.

Federated unlearning in graph-based models requires careful coordination between local and global updates. For instance, if a client unlearns specific nodes or edges from its graph, the central server must ensure that the global model is updated accordingly without compromising the accuracy or integrity of the remaining data. Efficient graph partitioning and communication protocols are essential to minimize the resource overhead of federated unlearning. Moreover, the process must ensure that unlearning on one client does not negatively impact the performance or security of the global model for other clients. This is especially important in privacy-sensitive applications like healthcare or finance, where federated learning is often used to comply with data privacy regulations.

[17] has a Federated Clusters method that accelerates the unlearning process, providing a significant speed-up compared to retraining. [18] uses SFU to perform gradient ascent in orthogonal space of input gradient spaces, negates a target client's contribution without additional storage.

8. Evaluation and Metrics for Graph Unlearning

Evaluating the effectiveness of graph unlearning algorithms requires a comprehensive set of metrics that measure time efficiency, privacy guarantees, and the impact on the graph's structural integrity. One of the primary performance metrics is time complexity, which evaluates how efficiently an unlearning algorithm can delete nodes and edges from a graph. For large-scale graphs, this metric is particularly important, as unlearning must be performed quickly and without consuming excessive computational resources [1].

Another critical metric is accuracy, which measures whether the unlearning process leaves behind any residual information that could be used to infer the deleted data. An unlearning algorithm must ensure that the removed nodes or edges are not just superficially deleted but completely erased from any machine learning models that were trained on the graph.

Privacy guarantees are also crucial when evaluating graph unlearning algorithms. These guarantees, often supported by differential privacy techniques, ensure that attackers cannot recover any unlearned information, even though sophisticated inference attacks. In applications like social networks or financial systems, strong privacy guarantees are essential for maintaining user trust and complying with regulatory requirements.

Evaluating graph unlearning algorithms typically involves experimental setups that use both synthetic and real-world datasets. Graphs from domains like social networks, financial systems, and healthcare are commonly used to test the scalability and robustness of unlearning methods. Performance metrics such as node degree distribution, clustering coefficients, and graph diameter

 doi:10.20944/preprints202409.1733.v1

are analyzed before and after unlearning to ensure that the remaining graph's structure is preserved. For example, in a social network graph, the unlearning of a user should remove all associated friendships without significantly disrupting community structures or user interactions. Similarly, in a financial transaction network, unlearning a transaction should eliminate all traces of that interaction while preserving the overall flow of transactions.

9. Conclusion

Graph unlearning represents a critical advancement in ensuring privacy and compliance with regulatory frameworks in modern graph-based systems. As more applications rely on graphs to represent complex relationships, the need to securely and efficiently erase sensitive data becomes increasingly urgent. This paper has presented a comprehensive analysis of graph unlearning, covering its fundamental concepts, the challenges posed by node and edge interdependencies, and the importance of maintaining graph integrity post-unlearning. We have also explored various algorithms that balance the trade-offs between efficiency, accuracy, and privacy, as well as the security risks, such as adversarial and inference attacks, that must be addressed to ensure robust unlearning mechanisms.

The application of graph unlearning spans across several critical domains, from social networks, where user data privacy is paramount, to financial systems and healthcare networks, where sensitive transactions and medical information must be erased securely. Additionally, the integration of graph unlearning into federated learning systems adds complexity, requiring coordinated efforts across decentralized clients to ensure global models are updated without retaining traces of unlearned data. As federated learning grows in popularity, developing effective federated graph unlearning methods will be essential for widespread adoption.

Ultimately, the success of graph unlearning lies in the development of advanced algorithms that can handle large-scale graphs while ensuring no residual information is left behind. The evaluation of these techniques, through metrics like time complexity, accuracy, and privacy guarantees, will be critical in determining their effectiveness in real-world applications. Future research must continue to focus on optimizing these algorithms for scalability and security, enabling more applications to benefit from privacy-preserving graph unlearning.

References

[1] N. Li, C. Zhou, Y. Gao, H. Chen, A. Fu, Z. Zhang and Y. Shui, "Machine Unlearning: Taxonomy, Metrics, Applications, Challenges, and Prospects," *arXiv preprint arXiv:2403.08254,* 2024.

[2] Z. Wang, Y. Zhu, Z. Li, Z. Wang, H. Qin and X. Liu, "Graph neural network recommendation system for football formation," *Applied Science and Biotechnology Journal for Advanced Research,* vol. 3, p. 33–39, 2024.

[3] Z. Li, B. Wang and Y. Chen, "Incorporating economic indicators and market sentiment effect into US Treasury bond yield prediction with machine learning," *Journal of Infrastructure, Policy and Development,* vol. 8, p. 7671, 2024.

[4] Z. Li, B. Wang and Y. Chen, "A Contrastive Deep Learning Approach to Cryptocurrency Portfolio with US Treasuries," *Journal of Computer Technology and Applied Mathematics,* vol. 1, pp. 1-10, 2024.

[5] W. Cong and M. Mahdavi, "Efficiently forgetting what you have learned in graph representation learning via projection," in *International Conference on Artificial Intelligence and Statistics*, 2023.

[6] W. Cong and M. Mahdavi, "Grapheditor: An efficient graph representation learning and unlearning approach".

[7] Y. Wei, X. Gu, Z. Feng, Z. Li and M. Sun, "Feature Extraction and Model Optimization of Deep Learning in Stock Market Prediction," *Journal of Computer Technology and Software,* vol. 3, 2024.

[8] E. Chien, C. Pan and O. Milenkovic, "Certified graph unlearning," *arXiv preprint arXiv:2206.09140,* 2022.

[9] K. Wu, J. Shen, Y. Ning, T. Wang and W. H. Wang, "Certified edge unlearning for graph neural networks," in *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2023.

[10] X. Zhu, G. Li and W. Hu, "Heterogeneous federated knowledge graph embedding learning and unlearning," in *Proceedings of the ACM web conference 2023*, 2023.

[11] C.-L. Wang, M. Huai and D. Wang, "Inductive graph unlearning," in *32nd USENIX Security Symposium (USENIX Security 23)*, 2023.

[12] J. Cheng, G. Dasoulas, H. He, C. Agarwal and M. Zitnik, "Gnndelete: A general strategy for unlearning in graph neural networks," *arXiv preprint arXiv:2302.13406*, 2023.

[13] H. Zhao, Y. Lou, Q. Xu, Z. Feng, Y. Wu, T. Huang, L. Tan and Z. Li, "Optimization Strategies for Self-Supervised Learning in the Use of Unlabeled Data," *Journal of Theory and Practice of Engineering Science*, vol. 4, p. 30–39, 2024.

[14] C. Pan, E. Chien and O. Milenkovic, "Unlearning graph classifiers with limited data resources," in *Proceedings of the ACM Web Conference 2023*, 2023.

[15] J. Wu, Y. Yang, Y. Qian, Y. Sui, X. Wang and X. He, "Gif: A general graph unlearning strategy via influence function," in *Proceedings of the ACM Web Conference 2023*, 2023.

[16] M. Chen, Z. Zhang, T. Wang, M. Backes, M. Humbert and Y. Zhang, "Graph unlearning," in *Proceedings of the 2022 ACM SIGSAC conference on computer and communications security*, 2022.

[17] C. Pan, J. Sima, S. Prakash, V. Rana and O. Milenkovic, "Machine unlearning of federated clusters," *arXiv preprint arXiv:2210.16424*, 2022.

[18] G. Li, L. Shen, Y. Sun, Y. Hu, H. Hu and D. Tao, "Subspace based federated unlearning," *arXiv preprint arXiv:2302.12448*, 2023.