Indiana Journal of Multidisciplinary Research
Abbreviated Tittle- Ind J Multi Res
ISSN *(Online)*- 2583-3820
Journal Homepage Link- https://indianapublications.com/Journals/IJMR
DOI: https://doi.org/10.5281/zenodo.21603812

INDIANA PUBLICATIONS
PRODUCTIVE AND QUALITY RESEARCH

Review Article

Volume-06|Issue-04|2026

A Review on Privacy Preserving Analysis

Akanksha Mehta[1], Devika Biradar[2], Gnanavi P[3], Apeksha Bharti[4], Dr. Shashidhar Virupaksha[5]

[1,2,3,4]UG Student, Department of Computer Science and Engineering, RV Institute of Technology and Management (RVITM), Bengaluru, Karnataka, India
[5]Assistant Professor, Department of Computer Science and Engineering, RV Institute of Technology and Management (RVITM), Bengaluru, Karnataka, India

Article History
Received: 05.05.2026
Accepted: 22.06.2026
Published: 25.07.2026

Citation
Mehta, A., Biradar, D., Gnanavi, P., Virupaksha, S. (2026). A Review on Privacy Preserving Analysis. *Indiana Journal of Multidisciplinary Research, 6*(4), 235-238.

Abstract: Maintaining privacy in data governance has become a necessary condition for securely sharing and using sensitive information. The increasing use of data-sharing practices, along with risks relating to data breaches, misuse, and other factors, has posed serious challenges to maintaining privacy while at the same time retaining the benefits of data. This review provides a structured, systematic investigation of privacy-preserving methods, highlighting their capabilities, limitations, and possible future directions. A thorough assessment of privacy-preserving techniques is undertaken with a particular focus on data aggregation and data generalization. The study is conducted using publicly available datasets and incorporates classification algorithms such as the Decision Tree algorithm to maximize accuracy while still protecting sensitive data. The study also identifies several critical gaps, such as scalability issues and computational shortcomings, which limit the effectiveness of current methods. Furthermore, this paper compares the proposed privacy-preserving approach with existing privacy-preserving solutions and evaluates performance metrics such as accuracy, privacy levels, and computational efficiency.

Keywords: Privacy preserving, data privacy, sensitive data protection.

INTRODUCTION

Nowadays, data has become a valuable asset in many fields such as healthcare, finance, transportation, and smart infrastructure. As organizations become increasingly dependent on data, the collection, processing, and sharing of personal and sensitive information have become more susceptible to abuse. However, this enormous demand for and accumulation of data has also raised concerns regarding data privacy, inappropriate use, unauthorized access, surveillance, and misuse. The consequences of data breaches and privacy violations not only reduce users' trust but also cause substantial economic and reputational damage. As a result, privacy preservation has gained renewed attention. Numerous data privacy and protection techniques, including anonymization, differential privacy, homomorphic encryption, federated learning, and secure multiparty computation, have been introduced.

These approaches help improve individual privacy while preserving data utility for analysis and decision-making. Despite considerable progress, challenges remain in balancing privacy with utility, ensuring the scalability of solutions for large datasets, and applying these methods effectively in real-world scenarios. This review aims to provide an overview of contemporary privacy-preserving mechanisms, examining their applicability, effectiveness, and suitability. By reviewing existing literature, this paper summarizes recent research, identifies current trends, discusses the advantages and disadvantages of various approaches, and highlights open research problems and future research directions.

PRIVACY PRESERVING ANALYSIS METHODS

Privacy Preserving Smart Grids

This paper proposes the use of homomorphic encryption to enable the training of deep learning and machine learning models for electrical applications while ensuring data security. The method is evaluated for fault detection, fault localization, and load forecasting in smart grids. Performance results show that encryption-based fault localization achieves an accuracy of 97%–98%, which is nearly identical to the 98%–99% accuracy obtained using unencrypted data. For load forecasting, the root mean square error (RMSE) with encryption is 0.0352 MWh, compared with 0.0248 MWh without encryption. The study demonstrates that homomorphic encryption enables secure model training without significantly compromising accuracy. This allows smart grid applications to leverage advanced AI models while maintaining confidentiality and complying with privacy requirements. [1]

This paper introduces a Secure and Privacy-Preserving Data Aggregation and Classification (SP-DAC) scheme over fog and cloud infrastructure. Aggregation is performed at the fog-node level, while classification is carried out in the cloud using three machine learning

***Corresponding Author:** Akanksha Mehta

classifiers. The research compares cryptographic overhead and classification accuracy through simulation studies. Experiments conducted using the real-world UMass Smart dataset achieved classification accuracy, precision, recall, and F1-score values of 88%, 87%, 90%, and 88%, respectively. Comparisons with existing models demonstrate that the SP-DAC framework provides superior security, privacy, and overall performance. [2]

Privacy Preserving in Federated Learning

This paper proposes FedGRU, a privacy-centered framework that combines Federated Learning with a Gated Recurrent Unit (GRU) neural network. Unlike traditional centralized systems that exchange raw data, FedGRU updates the global model through the secure aggregation of model parameters. The framework also reduces communication overhead and incorporates a joint aggregation protocol to improve scalability. Experimental results on real-world datasets show that FedGRU achieves prediction errors of only 0.76 km/h compared with existing models, demonstrating high predictive performance while preserving user privacy. [3]

Privacy-Preserving Federated Learning (PPFL) is a secure and decentralized learning approach. However, it remains vulnerable to model poisoning attacks, in which malicious participants, known as Byzantine attackers, intentionally degrade model performance. This paper introduces ShieldFL, a defense mechanism based on two-trapdoor homomorphic encryption that protects against encrypted model poisoning attacks while preserving privacy. ShieldFL employs a cosine similarity measure to calculate distances between encrypted gradients, followed by an aggregation technique that performs effectively across various data distributions. Experimental results demonstrate that ShieldFL significantly outperforms existing methods, improving model accuracy by 30%–80% against advanced poisoning attacks. [4]

Privacy Preservation in IoT

This paper presents a privacy-preserving machine learning framework called Heda, which enables secure model training without compromising user privacy. The framework utilizes partial homomorphic encryption, allowing machine learning models to be trained securely without relying on untrusted servers. Heda guarantees data privacy under the honest-but-curious threat model. Security analysis confirms that Heda provides strong privacy protection while maintaining model accuracy. Large-scale experiments demonstrate that the framework successfully preserves model performance while ensuring data confidentiality, making it an efficient and secure solution for privacy-preserving machine learning in smart IoT-driven cities. [5]

The Industrial Internet of Things (IIoT) plays a significant role in modern industries. However, machine learning models trained on sensitive industrial data remain vulnerable to privacy attacks. To address this challenge, the PriModChain framework combines differential privacy, federated learning, blockchain (Ethereum), and smart contracts to preserve IIoT data privacy and integrity. Python-based simulations using socket programming were conducted to evaluate security, reliability, and resilience. Ganache v2.0.1 was used for local blockchain testing, while the Kovan test network was used for public blockchain experiments. The proposed security protocol was also verified using Scyther v1.1.3. Overall, PriModChain enhances IIoT security while maintaining scalability and computational efficiency. [6]

Privacy Preserving in Logistic Regression

This paper investigates a parallel approach for computing logistic regression using a distributed asynchronous task framework. Unlike existing methods, the proposed approach does not depend on a third-party coordinator, thereby improving security and enabling multiple training scenarios. Logistic regression is implemented using homomorphic encryption in Python to ensure both accuracy and security in vertical federated learning. The proposed model is evaluated using the **MNIST** dataset, where experimental results demonstrate that it learns effectively and delivers robust, efficient, and reliable performance. [7]

This paper introduces a Privacy-Preserving Machine Learning (PPML) model for vertically partitioned data, where different devices provide different features for the same user. Such settings are more challenging than horizontally partitioned data because they require secure feature aggregation and sample synchronization. Existing methods generally require multiple rounds of communication between users and servers, leading to high communication and computational costs. To address these limitations, the paper proposes VPPLR, a privacy-preserving logistic regression framework. VPPLR securely performs gradient updates, allowing clients to leave the training process after submitting local data without compromising security. It also incorporates a vectorization technique that improves the efficiency of plaintext global parameter updates compared with encryption-intensive approaches. Experimental results demonstrate that VPPLR successfully builds accurate machine learning models using distributed IoT data. [8]

Privacy Preservation in Natural Language Processing

This paper presents Privacy-Preserving System Secure NLP, which focuses on a Recurrent Neural Network (RNN)-based sequence-to-sequence model with an attention mechanism for neural machine translation. The paper also presents two efficient Secure Multi-Party Computation (MPC) protocols for non-linear functions (sigmoid, tanh). It further proves the security of the two protocols—namely, the privacy-preserving Long Short-Term Memory (PrivLSTM) network and the privacy-

***Corresponding Author:** Akanksha Mehta

preserving sequence transformation (PrivSEQ2SEQ) model—under the semi-honest adversary model. The system is implemented in C++ and Python. [9]

Privacy Preservation in Graph Mining

Mining data from the Social Internet of Things (SIoT) is difficult because it suffers from large-scale noise, and there are possibilities of privacy leakage. To secure user data, the authors make use of Brakerski–Gentry–Vaikuntanathan (BGV) homomorphic encryption and store private information in encrypted form throughout the process. The encrypted graph structure is then divided using entropy-based partitioning algorithms to cluster nodes with similarities without compromising privacy. Normalized structural information and node partition similarity measures are also suggested by the research for assessing the accuracy of the resulting clustering. Experimental evaluation confirms that the proposed solution is highly effective in clustering while providing robust privacy assurance and reliable results, making it a suitable solution for secure SIoT data mining. [11]

Graph Neural Networks (GNNs) are effective in representing complex relationships, particularly in personalized domains such as recommendation systems. The paper proposes FedPerGNN, a federated learning-based GNN model that supports decentralized training while preserving privacy. The model incorporates a privacy-preserving model update approach, enabling collaboration without revealing raw data. It also includes a privacy-protected graph expansion protocol that allows models to leverage high-order interactions beyond local data while maintaining privacy. Experiments across six datasets show that FedPerGNN reduces errors by 4.0% to 9.6% compared with existing federated methods, demonstrating its effectiveness in privacy-sensitive, decentralized graph learning for personalized tasks. [12]

Privacy Preservation in Social Networks

It is difficult for privacy protection schemes to achieve an effective allocation of noise while maintaining data utility and execution efficiency due to the large volume of data in social network graph structures. Based on the differential privacy model, a privacy protection approach called PBCN is proposed. This approach comprises five algorithms, including random disturbance based on clustering and other techniques. Simulation experiments are conducted to compare the performance of PBCN with similar existing approaches. The experimental results demonstrate improved data utility and execution efficiency. [13]

Online Social Networks (OSNs) are widely used, but privacy issues arise due to the unauthorized disclosure of user data. This paper introduces BPP, a blockchain-based privacy-preserving framework that provides secure data transactions using public-key cryptography. It supports secure data sharing, retrieval, and access while maintaining user privacy. It employs a secure keyword search algorithm to ensure fairness and accurate results without requiring additional verification. A prototype of BPP was implemented on a test network, and the experimental results demonstrate its security, efficiency, and effectiveness. [14]

Privacy Preservation in Trajectory Mining

This paper introduces LSTMTrajGAN, a deep learning model for generating synthetic trajectory data while preserving user privacy. The model incorporates a dedicated loss function, TrajLoss, to measure trajectory similarity and optimize the training process. It is evaluated on a real-world trajectory dataset and outperforms state-of-the-art geomasking techniques in resisting re-identification attacks. Unlike conventional methods, it preserves essential spatial, temporal, and thematic characteristics of real trajectories. This approach effectively balances data utility with privacy preservation, providing valuable insights into AI-based privacy mechanisms for geolocation data. [15]

The increasing use of location-aware devices such as mobile phones, RFID tags, and GPS has led to the rapid growth of trajectory data across various applications. This work proposes WINR2D, a clustering-based solution that provides personalized privacy protection for individual trajectories. The algorithm assigns a privacy level to each trajectory, clusters them accordingly, and then anonymizes the data. The resulting generalized and distorted trajectory data achieve a balance between privacy preservation and data utility. Experimental results demonstrate that WINR2D effectively prevents successful user re-identification while maintaining the usefulness of the published data. [16]

CONCLUSION

This survey comprehensively reviews the evolving landscape of privacy-preserving technologies across several data-intensive domains, including smart grids, the Internet of Things (IoT), federated learning, natural language processing (NLP), logistic regression, graph mining, social networks, and trajectory data analysis. Despite the diversity of these application domains, three common challenges consistently emerge: balancing data utility with privacy protection, addressing scalability issues in resource-constrained and distributed environments, and defending against increasingly sophisticated adversarial attacks.

Advanced techniques such as homomorphic encryption, differential privacy, adversarial obfuscation, and federated learning have demonstrated significant potential for protecting data confidentiality while maintaining computational performance. However, existing implementations often face challenges related to computational latency, limited personalization, restricted cross-domain applicability, and evolving regulatory requirements. Hybrid approaches, such as blockchain-integrated federated learning, show considerable promise but also introduce additional concerns regarding

***Corresponding Author:** Akanksha Mehta

interoperability, computational overhead, and architectural complexity.

To establish a sustainable path forward, future research should focus on developing adaptive, lightweight, and context-aware privacy-preserving frameworks capable of dynamically adjusting protection levels based on data sensitivity, user requirements, and application-specific constraints. There is also a strong need for standardized evaluation protocols that quantitatively assess the trade-offs among privacy, utility, and scalability, enabling fair benchmarking across different approaches.

Furthermore, combating emerging adversarial techniques requires proactive defense mechanisms such as structure-preserving anonymization, secure synthetic data generation, and robust privacy budget management suitable for real-world deployments. Future systems should also emphasize transparency, interpretability, and regulatory compliance, particularly in high-stakes domains such as healthcare, transportation, and critical infrastructure.

By integrating domain-specific innovations with reproducible cross-domain solutions, supported through interdisciplinary collaboration among privacy researchers, system architects, and policymakers, the field can advance toward secure, intelligent, and ethically grounded data-driven technologies.

REFERENCES

1. Syed, D., Refaat, S. S., & Bouhali, O. (2020). Privacy preservation of data-driven models in smart grids using homomorphic encryption. *Information, 11*(7), 357. https://doi.org/10.3390/info11070357
2. Singh, K., & Kumar, J. (2023). A secure and privacy-preserving data aggregation and classification model for smart grid. *Multimedia Tools and Applications, 82*, 22997–23015. https://doi.org/10.1007/s11042-023-14590-3
3. Liu, Y., Yu, J. J. Q., Kang, J., Niyato, D., & Zhang, S. (2020). Privacy-preserving traffic flow prediction: A federated learning approach. *IEEE Internet of Things Journal, 7*(8), 7751–7763. https://doi.org/10.1109/JIOT.2020.2991401
4. Ma, Z., Ma, J., Miao, Y., Li, Y., & Deng, R. H. (2022). ShieldFL: Mitigating model poisoning attacks in privacy-preserving federated learning. *IEEE Transactions on Information Forensics and Security, 17*, 1639–1654. https://doi.org/10.1109/TIFS.2022.3167699
5. Zhu, L., Tang, X., Shen, M., Gao, F., Zhang, J., & Du, X. (2021). Privacy-preserving machine learning training in IoT aggregation scenarios. *IEEE Internet of Things Journal, 8*(15), 12106–12118. https://doi.org/10.1109/JIOT.2021.3065916
6. Arachchige, P. C. M., Bertok, P., Khalil, I., Liu, D., Camtepe, S., & Atiquzzaman, M. (2020). A trustworthy privacy-preserving framework for machine learning in industrial IoT systems. *IEEE Transactions on Industrial Informatics, 16*(9), 6092–6102. https://doi.org/10.1109/TII.2020.2974763
7. He, D., Du, R., Zhu, S., Zhang, M., Liang, K., & Chan, S. (2022). Secure logistic regression for vertical federated learning. *IEEE Internet Computing, 26*(2), 61–68. https://doi.org/10.1109/MIC.2021.3120884
8. Zhang, Y., & Tang, M. (2024). VPPLR: Privacy-preserving logistic regression on vertically partitioned data using vectorization sharing. *Journal of Information Security and Applications, 82*, Article 103725. https://doi.org/10.1016/j.jisa.2024.103725
9. Feng, Q., He, D., Liu, Z., Wang, H., & Choo, K.-K. R. (2020). SecureNLP: A system for multi-party privacy-preserving natural language processing. *IEEE Transactions on Information Forensics and Security, 15*, 3709–3721. https://doi.org/10.1109/TIFS.2020.3003237
10. Li, J., *et al.* (2022). A federated learning-based privacy-preserving smart healthcare system. *IEEE Transactions on Industrial Informatics, 18*(3), 2021–2031. https://doi.org/10.1109/TII.2021.3108353
11. Tian, Y., Zhang, Z., Xiong, J., Chen, L., Ma, J., & Peng, C. (2022). Achieving graph clustering privacy preservation based on structure entropy in social IoT. *IEEE Internet of Things Journal, 9*(4), 2761–2777. https://doi.org/10.1109/JIOT.2021.3105903
12. Wu, C., Wu, F., Lyu, L., *et al.* (2022). A federated graph neural network framework for privacy-preserving personalization. *Nature Communications, 13*, Article 3091. https://doi.org/10.1038/s41467-022-30714-9
13. Huang, H., Zhang, D., Xiao, F., Wang, K., Gu, J., & Wang, R. (2020). Privacy-preserving approach PBCN in social networks with differential privacy. *IEEE Transactions on Network and Service Management, 17*(2), 931–945. https://doi.org/10.1109/TNSM.2020.2979718
14. Zhang, S., Yao, T., Arthur Sandor, V. K., Weng, T. H., Liang, W., & Su, J. (2021). A novel blockchain-based privacy-preserving framework for online social networks. *Connection Science, 33*(3), 555–575. https://doi.org/10.1080/09540091.2021.1888990
15. Rao, J., Gao, S., Kang, Y., & Huang, Q. (2020). *LSTMTrajGAN: A deep learning approach to trajectory privacy protection* (arXiv Preprint No. arXiv:2006.10521). https://arxiv.org/abs/2006.10521
16. Mahdavifar, S., Deldar, F., & Mahdikhani, H. (2022). Personalized privacy-preserving publication of trajectory data by generalization and distortion of moving points. *Journal of Network and Systems Management, 30*, Article 10. https://doi.org/10.1007/s10922-021-09647-8

***Corresponding Author:** Akanksha Mehta