Chapter 8

Book Chapter

# Privacy Enhancing Cross-Silo Federated Learning For FDIA Using ML

Bhoomika C J *[1] and S. PandiKumar S †[2]

[1]Master of Computer Applications, Acharya Institute of Technology, Bangalore
[2]Assistant Professor, Master of Computer Applications, Acharya Institute of Technology, Bangalore

## Abstract

Combined Learning (CL) tackles data privacy by allowing users to store data locally and share only model parameters with a central server to train a global model. However, CL is vulnerable to inference attacks from untrusted aggregators. Existing solutions often require a trusted third party or inefficient protocols. Our work proposes an efficient privacy-preserving federated learning scheme with strong security. By designing a dual-layer encryption scheme without the need for discrete logarithm calculations, using secret sharing only initially and when groups rejoin, and enhancing computational efficiency through parallel processing, we ensure secure federated learning. This method is applied to detect false data injection attacks (FDIA) in smart systems, offering improved security and resistance to private data inference attacks compared to previous methods.

Keywords: False data injection attack. Cross-Silo. Shamir's Secrete Sharing. Privacy Preserve. Federated Learning.

*Email: bhoomikac.22.mcav@acharya.ac.in Corresponding Author
†Email: pandikumar2906@acharya.ac.in

## 1 Introduction

In the era of data-driven decision-making, there are significant privacy concerns.(Gautam & Mittal, 2022). Collaborative machine learning enables the use of distributed datasets while maintaining data privacy. Traditional methods centralize sensitive data, risking privacy breaches and regulatory issues. Federated learning addresses this by allowing multiple parties to train models collaboratively without sharing their data. However, existing federated learning frameworks struggle to balance privacy and efficiency, especially in cross-silo settings where data spans multiple organizations. This project proposes an innovative cross-silo federated learning scheme using double-layer encryption, secret sharing, and parallel computing to enhance privacy and computational performance. Our approach allows participants to train models jointly without compromising data confidentiality and supports dropout and rejoining, improving scalability. We apply our scheme to false data injection attack (FDIA) detection in smart grids, demonstrating its real-world applicability. Through theoretical analysis and empirical evaluation, we prove its privacy guarantees against adversarial threats while achieving effective model performance.

Overall, this project advances privacy-preserving federated learning, providing a robust, efficient framework for secure data collaboration across organizational boundaries, benefiting industries reliant on collaborative data analysis and model training.

## 2 Literature Survey

Valkenburg, Meier, and Beyens's (2022) paper introduces communication-efficient learning of deep networks from decentralized data. It addresses the challenge of training deep learning models on decentralized data sources. This provides insights into communication-efficient techniques, which are crucial in federated learning scenarios. Further the study by Arias-De la Torre et al.'s (2020) presents model inversion attacks that exploit confidence information and basic counter measures. It is important as it sheds light on security vulnerabilities in machine learning models, which is highly relevant in the context of federated learning where privacy is a primary concern. Ivie et al.'s (2020) paper discusses stealing machine learning models via prediction APIs which highlights and explores the vulnerabilities of machine learning models, crucial in the context of federated learning where models are trained across distributed data sources. Cataldo et al.'s (2021) paper investigated information leakage from collaborative deep learning. This is relevant as it examines the risks associated with collaborative learning, an area directly relevant to federated learning where models are trained across distributed data sources.

Researchers have been discussing models on inversion attacks against collaborative inference. Understanding these attacks is crucial in federated learning scenarios, where models are trained across distributed data sources.This may also help to protect the se-


DOI:10.48001/978-81-966500-7-0-8 | ISBN: 978-81-966500-7-0 

curity and integrity of these teaching tools in federated learning scenarios.(Mittal, Kaur, & Jain, 2022). Researchers have examined unintended feature leakage in collaborative learning. This is pertinent to our research as it addresses privacy concerns in collaborative learning scenarios, which are analogous to federated learning setups. The unintended memorization in neural networks addresses the privacy and security aspects of machine learning models, which are essential considerations in federated learning.(Karim et al., 2020). The vulnerability of AC state estimation to false data injection of cyber-attacks examines the security of vulnerabilities in smart grid systems, a domain where federated learning is applied.

Keles, McCrae, and Grealish's (2020) paper provides a review of false data injection attacks against modern power systems discussing the security threats ,federated learning is applied for privacy-preserving analytics. Faelens et al.'s (2021) study discusses transmission management in a deregulated environment is significant as it provides insights into the challenges and requirements of deregulated energy markets, where federated learning can be applied for efficient data analysis. This improved data analysis will further improve public administration. (Mittal, 2020)

## 3 Proposed System

In this research, we propose an efficient cross-silo federated learning system with robust privacy preservation, suitable for the smart grid domain. Our method employs a double-layer encryption scheme and Shamir secret sharing to ensure strong privacy while enabling clients to dynamically drop out and rejoin during training. Key contributions include a general privacy-enhancing federated learning framework with secure weighted aggregation, eliminating the need for discrete logarithms and non-colluding servers. The scheme uses decentralized key generation to enhance privacy. Additionally, our system incorporates a novel double-masking technique to protect against information leakage during transmission delays. Theoretically and empirically, our method demonstrates provable privacy against honest-but-curious servers and maintains model utility. It is efficient in communication and computation, utilizing a logarithmic communication graph to reduce overhead. Our approach is resilient to local training data inference attacks and is particularly useful for False Data Injection Attack (FDIA) detection in smart grids. Extensive experiments validate the efficiency and privacy guarantees of our framework, making it a practical solution for secure federated learning in critical infrastructure.

## 4 Existing System

Existing secure aggregation systems for federated learning face several limitations. Schemes based on ElGamal homomorphic encryption and the Decisional Composite Residuosity


DOI:10.48001/978-81-966500-7-0-8 | ISBN: 978-81-966500-7-0 

Assumption require trusted dealers and struggle with discrete logarithm computation or handle only scalar aggregation, lacking dropout resilience. A vector-based approach using pairwise additive stream ciphers and Shamir secret sharing addressed dropouts but required multiple communication rounds per iteration, with inefficiencies and weak security models. Some systems necessitate multiple non-colluding servers, resulting in significant overheads and vulnerability to collusion. Other methods, such as modified ElGamal and combinations of homomorphic encryption with differential privacy, faced issues with discrete logarithms and could not handle dropouts effectively, while approaches using functional encryption, k-regular graphs, and verifiable computation had drawbacks related to trusted parties, dropout resilience, and collusion assumptions. Additionally, protocols leveraging threshold secret sharing incurred high computation and communication costs. Hybrid schemes using functional encryption and differential privacy were dependent on trusted third parties, posing significant security risks. Finally, additive noise methods compromised differential privacy, failing to provide strong protection against inference attacks. Overall, these limitations highlight the need for a more efficient and resilient secure aggregation framework in federated learning.

## 5 Methodology Used

- White Box Testing: White Box Testing involves an in-depth understanding of the software's inner workings, structure, and language. This method allows testers to explore areas that are not accessible from a black box level, focusing on the internal logic and code of the software.
- Black Box Testing: Black Box Testing is conducted without any knowledge of the internal structure or language of the module being tested. Test cases are derived from definitive source documents such as specifications or requirements, and the software is treated as a black box. Inputs are provided, and outputs are observed without considering the internal workings of the software.
- Unit Testing: Unit testing is carried out as part of the combined code and unit test phase of the software lifecycle. It ensures that individual units of the software function correctly. While coding and unit testing may be conducted as separate phases, they are often performed together.
- Test Strategy and Approach: Field testing will be executed manually, while functional tests will be meticulously scripted. The primary objectives of the testing phase are to ensure that all field entries function properly, pages are activated from the identified links, and the entry screen, messages, and responses are not delayed.
- Integration Testing: Integration testing involves the incremental integration of two or more integrated software components on a single platform to identify failures caused by interface defects.


DOI:10.48001/978-81-966500-7-0-8 | ISBN: 978-81-966500-7-0 

- Acceptance Testing: User Acceptance Testing, a crucial phase of the project, requires significant participation from end-users to ensure that the system meets functional requirements.

Test Results: All the test cases mentioned above passed successfully, and no defects were encountered during testing.

## 6 Architecture

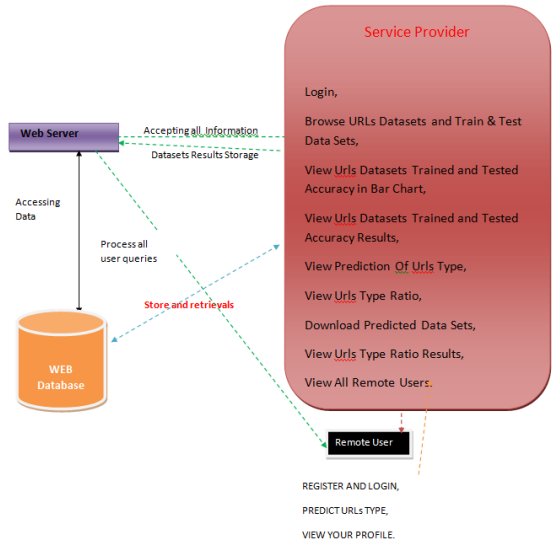


Figure 1. Architecture

This system architecture ensures a secure, efficient, and scalable federated learning environment, making it highly suitable for applications requiring robust privacy measures, such as smart system attack detection.(see figure 1). The architecture employs advanced encryption techniques and decentralized key management to safeguard data privacy and integrity. It supports dynamic client participation, allowing devices to securely join and leave the network without compromising the overall system security. The use of secure multiparty computation ensures that no single entity can access the entire dataset, mitigating risks of data breaches. Additionally, the system's design optimizes communication overhead and computational efficiency, enabling real-time detection and response to potential threats. By leveraging these features, the architecture not only enhances security


DOI:10.48001/978-81-966500-7-0-8 | ISBN: 978-81-966500-7-0 

but also ensures high availability and resilience, crucial for maintaining trust in critical infrastructures.

## 7 Module Description

- Service Provider In this module, the Service Provider has to login by using valid user name and password. After login successful he can do some operations such as Login, Browse URLs Datasets and Train & Test Data Sets, View URL Datasets Trained and Tested Accuracy in Bar Chart, View URL Datasets Trained and Tested Accuracy Results, View Prediction Of URL Type, View URL Type Ratio, Download Predicted Data Sets, View URL Type Ratio Results, View All Remote Users.
- View and Authorize Users In this module, the admin can view the list of users who all registered. In this, the admin can view the user's details such as, user name, email, address and admin authorizes the users.
- Remote User In this module, there are numbers of users are present. User should register before doing any operations. Once user registers, their details will be stored to the database. After registration successful, he has to login by using authorized user name and password. Once Login is successful user will do some operations like REGISTER AND LOGIN, PREDICT URLs TYPE, VIEW YOUR PROFILE.

## 8 Flowchart

The process has been explained in the following flowchart:(see figure 2)

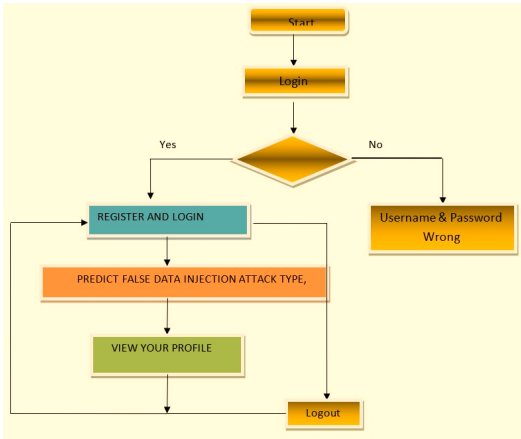


Figure 2. Flowchart


DOI:10.48001/978-81-966500-7-0-8 | ISBN: 978-81-966500-7-0 

- Start: The process begins at the start node.
- Login: The user is prompted to log in with their username and password.(see figure 3)


Figure 3. Login Page

- Login Decision: The system checks if the entered username and password are correct. If the credentials are incorrect, the user is shown a "Username & Password Wrong" message and prompted to log in again. If the credentials are correct, the user proceeds to the next step.
- Register and Login: Upon successful login, the user is either directed to register (if they haven't already) or logged into their account. This step ensures that the user is properly authenticated and registered in the system.
- Predict False Data Injection Attack Type: Once logged in, the user can utilize the system's functionality to predict the type of false data injection attacks. This involves analyzing incoming data to detect and classify any potential false data injections.
- View Your Profile: The user has the option to view their profile, which may contain personal information, system usage statistics, or other relevant data. This step allows users to manage their account settings and monitor their activity within the system.
- Logout: The user can choose to log out of the system. Logging out ensures that the user's session is terminated, and they need to log in again to access the system.

## 9 Results

Functional tests provide systematic demonstrations that functions tested are available as specified by the business and technical requirements, system documentation, and user manuals. Functional testing is centered on the following items:

- Valid Input : identified classes of valid input must be accepted.
- Invalid Input : identified classes of invalid input must be rejected.
- User Authentication: Users are correctly authenticated based on valid credentials, employing multi-factor authentication to enhance security.
- Access Control: Unauthorized access is prevented, ensuring security through role-based


DOI:10.48001/978-81-966500-7-0-8 | ISBN: 978-81-966500-7-0 

access control (RBAC) and least privilege principles.

- Session Management: User sessions are managed correctly, with proper handling of login, logout, and session timeouts, alongside automatic session expiration after a period of inactivity.
- Error Handling: Appropriate error messages are displayed for invalid login attempts, and the system prevents brute force attacks through account lockout mechanisms and CAPTCHA integration.
- Data Encryption: Sensitive user data, both in transit and at rest, is encrypted using strong encryption protocols.
- Audit Logging: All access and authentication attempts are logged for auditing and monitoring purposes, aiding in the detection of unauthorized access attempts.
- Regular Security Updates: The system is regularly updated with security patches encryption. This approach eliminates the need for computing discrete logarithms or relying on multiple non-colluding server settings, addressing the limitations of several related works. Additionally, the secret keys for the two encryption layers are generated by each participant in a decentralized manner, enhancing privacy.

We design and empirically evaluate a practical and efficient privacy-enhancing cross-silo federated learning system that is robust against local data inference attacks, specifically for FDIA detection in smart grids. The proposed scheme provides a framework adaptable to various domains.In future work, we will explore more dynamic adversarial models in mitigate vulnerabilities and protect against emerging threats.


Figure 4. Prediction of URL


DOI:10.48001/978-81-966500-7-0-8 | ISBN: 978-81-966500-7-0 

The output page for the "Privacy Enhancing Cross-Silo Federated Learning for FDIA Using ML" project provides the results of the false data injection (FDIA) detection. After logging into the system, users enter the URL, Headline of the URL, and Body of the URL on the input page. Upon clicking the "Predict" button, the system processes the input data and displays the prediction results below.(see figure 4).The output will clearly indicate whether false data injection has been detected, showing a message like "No false data injection detected" for clean data or appropriate alerts if false data injection is present. Additionally, the page will provide a confidence score for the prediction, indicating the system's certainty level. Users will also see detailed information on the detected anomalies, including the type and location of the false data within the input. The system offers recommendations for mitigating detected FDIA, ensuring users can take immediate and informed action.

## 10 Conclusions

In this paper, we propose a cross-silo privacy-enhancing federated learning system that is secure in the honest-but-curious adversarial model. Our scheme is resilient to client dropouts and rejoining, and it is efficient in terms of communication and computation overhead, leveraging secure multiparty computation techniques like secret sharing and double-layer different federated learning settings, relevant for security in cyber-physical systems.

## References


Arias-De la Torre, J., Puigdomenech, E., García, X., Valderas, J. M., Eiroa-Orosa, F. J., Fernández-Villa, T., Molina, A. J., Martín, V., Serrano-Blanco, A., Alonso, J., & Espallargues, M. (2020). Relationship between depression and the use of mobile technologies and social media among adolescents: Umbrella review. Journal of Medical Internet Research, 22(8). https://doi.org/10.2196/16388

Cataldo, I., Lepri, B., Neoh, M. J. Y., & Esposito, G. (2021). Social Media Usage and Development of Psychiatric Disorders in Childhood and Adolescence: A Review. Frontiers in Psychiatry, 11. https://doi.org/10.3389/fpsyt.2020.508595

Faelens, L., Hoorelbeke, K., Cambier, R., van Put, J., Van de Putte, E., De Raedt, R., & Koster, E. H. (2021). The relationship between Instagram use and indicators of mental health: A systematic review. Computers in Human Behavior Reports, 4. https://doi.org/10.1016/j.chbr.2021.100121

Gautam, S., & Mittal, P. (2022). Comprehensive Analysis of Privacy Preserving Data Mining Algorithms for Future Develop Trends. International Research Journal of Computer Science, 9(10), 367–374. https://doi.org/10.26562/irjcs.2022.v0910.01




DOI:10.48001/978-81-966500-7-0-8 | ISBN: 978-81-966500-7-0

Ivie, E. J., Pettitt, A., Moses, L. J., & Allen, N. B. (2020). A meta-analysis of the association between adolescent social media use and depressive symptoms. Journal of Affective Disorders, 275, 165–174. https://doi.org/10.1016/j.jad.2020.06.014

Karim, F., Oyewande, A., Abdalla, L. F., Chaudhry Ehsanullah, R., & Khan, S. (2020). Social Media Use and Its Connection to Mental Health: A Systematic Review. Cureus. https://doi.org/10.7759/cureus.8627

Keles, B., McCrae, N., & Grealish, A. (2020). A systematic review: the influence of social media on depression, anxiety and psychological distress in adolescents. International Journal of Adolescence and Youth, 25(1), 79–93. https://doi.org/10.1080/02673843.2019.1590851

Mittal, P. (2020). Big data and analytics: a data management perspective in public administration. International Journal of Big Data Management, 1(1), 1. https://doi.org/10.1504/ijbdm.2020.10032871

Mittal, P., Kaur, A., & Jain, R. (2022). Online Learning for Enhancing Employability Skills in Higher Education Students: The Mediating Role Of Learning Analytics. TEM Journal, 11(4), 1469–1476. https://doi.org/10.18421/TEM114-06

Valkenburg, P. M., Meier, A., & Beyens, I. (2022). Social media use and its impact on adolescent mental health: An umbrella review of the evidence. Current Opinion in Psychology, 44, 58–68. https://doi.org/10.1016/j.copsyc.2021.08.017


DOI:10.48001/978-81-966500-7-0-8 | ISBN: 978-81-966500-7-0 