iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** VI **Month of publication:** June 2025

DOI: https://doi.org/10.22214/ijraset.2025.72175

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue VI June 2025- Available at www.ijraset.com

Securing AI-as-a-Service Models against Inference Leakage Attacks with Applied Cryptography

Mr. Aravindhan N[1], Vinitha P[2]
[1]*Assistant Professor, II MCA*
[1,2]*Department of Master of Computer Applications, Er.Perumal Manimekalai College of Engineering, Hosur*

Abstract: ***AI-as-a-Service (AIaaS) is an innovative cloud-based paradigm that empowers service providers to deliver advanced artificial intelligence capabilities—such as facial recognition, financial prediction, and epidemic modeling—through accessible online platforms. Despite its benefits, this model presents serious data privacy challenges, as users must transmit sensitive personal or corporate information to external servers. Among these threats, inference leakage attacks are particularly concerning, as they can compromise both user data and the integrity of the AI models.Traditional methods often struggle to strike an optimal balance between efficiency and data confidentiality, leading to security gaps and increased exposure to unauthorized access or leaks. To mitigate these issues, this project introduces a privacy-preserving solution based on Fully Homomorphic Encryption (FHE), which allows computations to be performed directly on encrypted data.With FHE, user inputs remain encrypted throughout the AI processing workflow. When a user submits encrypted data, the cloud server processes it using the AI model without ever seeing the original content. The resulting output is also encrypted and can only be decrypted by the user holding the appropriate private key. This ensures that the service provider cannot access either the input or the output in plain form.By employing FHE, this framework secures the inference process, blocks leakage of sensitive information, and preserves the proprietary nature of AI models. The method is especially suitable for applications requiring rapid decision-making, such as live facial recognition, while also reinforcing trust and privacy in AI cloud services.***

Keywords: AI-as-a-Service (AIaaS), Inference Leakage, Privacy-Preserving AI, Cryptographic Techniques, Homomorphic Encryption, Secure Multiparty Computation

I. INTRODUCTION

In the modern digital age, Artificial Intelligence-as-a-Service (AIaaS) has emerged as a revolutionary approach, enabling individuals and organizations to access cutting-edge AI functionalities via cloud services. This model removes the dependency on expensive on-site infrastructure, making it possible to utilize technologies like voice recognition, visual data interpretation, and predictive analytics more efficiently across various sectors. Despite its convenience, AIaaS introduces critical concerns related to data confidentiality and system security. One of the most pressing issues in this area is the danger posed by inference leakage attacks. In such attacks, adversaries can potentially uncover private information from user inputs or internal characteristics of the AI models by observing their outputs. These attacks endanger not only the privacy of the user's data but also the intellectual property and operational integrity of the AI services. Traditional encryption approaches are often inadequate, especially during inference processes where data usually needs to be decrypted for computation. To tackle these security challenges, researchers have begun adopting innovative cryptographic solutions, with Fully Homomorphic Encryption (FHE) being one of the most promising. FHE allows operations to be carried out directly on encrypted data, ensuring that the original content remains hidden throughout the computational cycle. This method preserves data privacy without compromising the functionality of AI services hosted on the cloud. This project proposes the implementation of FHE in AIaaS environments to counter inference leakage threats effectively. The goal is to develop a robust and privacy-centric infrastructure that enhances user trust and ensures secure, efficient AI model usage even in performance-sensitive applications.

II. PROPOSED WORK

The proposed solution presents a secure and privacy-focused AI-as-a-Service (AIaaS) architecture, leveraging Fully Homomorphic Encryption (FHE) to protect user data and eliminate risks associated with inference leakage.

A. Safeguarding Against Inference Leakage

Inference leakage presents a serious vulnerability in cloud-based AI, as it can reveal sensitive details about both user inputs and the AI model itself. Through the integration of FHE, this system ensures that all computations occur on encrypted data, meaning that neither the input data nor the output results are ever exposed to the server, significantly enhancing security.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue VI June 2025- Available at www.ijraset.com

B. Secure AI Computation with Full Data Confidentiality

In this framework, users first encrypt their private data locally before sending it to the cloud. The AI algorithms then perform operations on this encrypted input, without requiring decryption at any point during processing. Because the data remains encrypted throughout, no meaningful information can be extracted by service providers or malicious actors, guaranteeing robust privacy protection.

C. Exclusive User Access Through Controlled Decryption

Only the end-user, equipped with the corresponding private key, can decrypt and view the final inference results. This system ensures that access to sensitive outputs is tightly restricted, thereby protecting against misuse and maintaining the integrity of AI outcomes. It allows users to fully utilize AI functionalities while maintaining complete control over their confidential information.

III. METHODS

A. AIaaS Model Service Provider

The AIaaS Model Service Provider functions as the core component of the system, managing secure model deployment, user access, and encrypted communications through the use of Homomorphic Encryption (HE). It handles the initial registration and verification process, granting or denying access requests from both AI Model Developers and End Users to ensure that only trusted participants interact with the system.Once a Model Owner is authorized, the service provider ensures their AI models are encrypted using Homomorphic Encryption before deployment, thereby preserving the confidentiality and integrity of sensitive data. In addition, it actively monitors the usage of each AI model, creating an auditable and transparent log of activities that supports accountability and regulatory compliance.The provider also takes on the critical responsibility of managing cryptographic keys, enabling secure and controlled exchange of encryption and decryption keys between legitimate users. This key management process ensures that all encrypted data transactions within the platform remain secure, protected from unauthorized access, and aligned with the system's privacy-preserving goals.

B. System User Module

This module facilitates operations for both Model Owners and Model Users, enabling secure interactions within the AIaaS environment.

1) Model Owner Functions

Account Registration: Submits necessary personal or organizational details for identity verification.
Access Authorization: Gains system access following approval by the AIaaS Model Service Provider.
Secure Login: Authenticates into the system using verified credentials.
Feature Encryption & Model Deployment: Applies Homomorphic Encryption to sensitive model features before uploading the AI model to the platform.
Monitor Model Activity: Reviews logs detailing user interactions, request frequency, and overall model engagement metrics.

2) Model User Functions

User Registration: Provides required information to initiate account creation.
Approval Confirmation: Access is granted upon successful review by the AIaaS Model Service Provider.
Secure Sign-In: Logs into the platform through verified login credentials.
Submit Encrypted Input: Encrypts sensitive data using Homomorphic Encryption prior to sending it for processing.
Decrypt Output: Uses a private decryption key to access and interpret the processed results securely.

C. Key Generation Module

This module plays a vital role in establishing the cryptographic foundation required for Homomorphic Encryption (HE) operations within the system.
Key Pair Creation: Utilizes specialized algorithms designed for Homomorphic Encryption to generate matching public and private key pairs.
Key Distribution: Allocates the public encryption keys to Model Owners for securing their AI models and provides the corresponding private decryption keys to Model Users for securely accessing the final results.

D. Data Encryption Module

The Data Encryption Module is essential for protecting sensitive content by ensuring that all data—whether from the Model Owner or the Model User—is encrypted before any processing occurs. Utilizing Homomorphic Encryption (HE), this module secures both the AI model's features and user-provided input.

For Model Owners, it encrypts the model's internal parameters and features using their designated encryption key, preserving the confidentiality of proprietary algorithms. For Model Users, it ensures that input data is encrypted prior to being sent to the cloud, blocking any potential exposure during transmission or execution.

By leveraging Homomorphic Encryption techniques, this module enables all computations to occur on encrypted data, eliminating the need for decryption at any point in the processing pipeline. This continuous encryption significantly strengthens the system's data privacy and overall security posture.

E. AI Model Evaluation Module

The AI Model Evaluation Module is responsible for executing privacy-preserving computations on encrypted data using Homomorphic Encryption (HE). It allows the AI model to process encrypted user inputs without decrypting them, ensuring that both the model's internal logic and the user's data remain completely confidential during inference.

This module ensures that, throughout the computation process, the encrypted input from the user and the encrypted features of the AI model are never exposed to any unauthorized entity. The evaluation is carried out entirely within the encrypted domain, offering end-to-end privacy.

Once the processing is complete, the module produces encrypted output results, which can only be decrypted by the intended user. This guarantees the security of sensitive information while preserving the accuracy and reliability of the AI model's predictions.

F. Encrypted Output Generation Module

The Encrypted Output Generation Module handles the creation of securely encrypted results following the AI model's processing of user data. After the encrypted input is evaluated by the model, this module ensures that the corresponding output remains encrypted before being sent back to the user.

Its primary function is to safeguard the integrity and privacy of the prediction results during transmission, preventing any unauthorized access or data leaks. The encrypted output is only accessible to the intended Model User, who possesses the necessary decryption key.

By maintaining encryption throughout the output delivery phase, this module reinforces the system's commitment to end-to-end data protection, playing a vital role in upholding user trust and the security standards of the privacy-preserving AIaaS framework.

G. Result Decryption Module

The Result Decryption Module enables the Model User to securely unlock and interpret the output generated by the encrypted model evaluation. Using a designated private decryption key, this module transforms the encrypted results into readable predictions without exposing any sensitive data during the process.

It ensures that the decoded information remains faithful to the original computation, preserving both the accuracy and consistency of the AI model's outcomes. Furthermore, strict access controls are integrated to guarantee that only authorized users can decrypt and view the results, effectively blocking any unauthorized access attempts.

This module plays a pivotal role in upholding the system's principles of data privacy, integrity, and secure access, ensuring that users can trust the platform with their confidential information complexity, MimicModel incorporated performance enhancements to maintain efficient response times, especially suitable for operations that are not time-sensitive.

The blockchain's role in logging activities introduced only minimal delays, while significantly increasing reliability and auditability.

IV. RESULTS

The integration of a Fully Homomorphic Encryption (FHE)-based approach within the AI-as-a-Service (AIaaS) framework has effectively validated the system's ability to conduct AI model inferences securely, without revealing user data or exposing model internals. This setup allowed seamless interaction between Model Owners and Model Users while strictly preserving data confidentiality at every stage.

A. Highlighted Outcomes

Secure Encrypted Computation: The platform successfully handled encrypted user data without any need for decryption during processing, ensuring complete isolation of sensitive information from the server environment.

Protected Model Hosting: All AI models were safeguarded through encryption before deployment, thereby shielding their architecture and proprietary features from unauthorized inspection.

Complete Encryption Workflow: The system maintained encryption across the entire data lifecycle—from input submission to result generation—allowing only users with valid decryption keys to access the final outputs.

Defense Against Inference Exposure: The encrypted outputs provided no clues about either the input data or the AI model logic, eliminating the possibility of inference-based data leakage.

Prediction Reliability: Even while operating exclusively on encrypted data, the system produced accurate and relevant inference results, affirming the practical applicability of FHE in real-time scenarios.

Strict Access Governance: A secure key distribution mechanism and role-based access controls ensured that only authenticated users could decrypt or view any information, further enhancing trust and system integrity.

V. CONCLUSION

In summary, this project presents a robust and privacy-focused AI-as-a-Service (AIaaS) framework built upon Fully Homomorphic Encryption (FHE), enabling secure deployment and utilization of AI models without exposing sensitive information. The architecture incorporates critical components including the AIaaS Service Provider Module, End-User Interface, Key Management System, Data Encryption Layer, Secure Model Computation Engine, and Output Decryption Unit. This design guarantees complete data encryption from input submission to result delivery, ensuring privacy throughout the entire computational workflow.

Model Providers are able to upload encrypted models securely, while End Users can input encrypted data and receive results in an encrypted format—ensuring neither party reveals any unencrypted information. The implementation of FHE ensures that all model inference operations are conducted on encrypted data, eliminating the need for decryption and thus fortifying user and model confidentiality.

Additionally, the framework incorporates strong key governance, access permissions, and usage logging to promote accountability and ensure secure AI service management. Although the current system addresses major privacy and data protection challenges, future work can aim to improve processing efficiency, support scalable cloud environments, and broaden compatibility with various AI architectures. Ultimately, this solution marks a meaningful step toward establishing secure, privacy-respecting AI services applicable in critical domains such as healthcare, finance, and national security.

VI. ACKNOWLEDGMENT

The authors confirm that there are no acknowledgments or external contributions to declare for this study.

REFERENCES

[1] X. Pei, X. Deng, S. Tian, J. Liu, and K. Xue, “A privacy-preserving graph neural network design tailored for decentralized local graph scenarios,” IEEE Transactions on Information Forensics and Security, vol. 19, pp. 1614–1629, 2024.

[2] L. Bergerat, A. Boudi, Q. Bourgerie, I. Chillotti, D. Ligier, J.-B. Orfila, et al., “Tuning parameters and enhancing precision for (T)FHE computations,” Journal of Cryptology, vol. 36, no. 3, article 28, Jun. 2023.

[3] A. El Ouadrhiri and A. Abdelhadi, “Survey on differential privacy approaches in deep and federated learning environments,” IEEE Access, vol. 10, pp. 22359–22380, 2022.

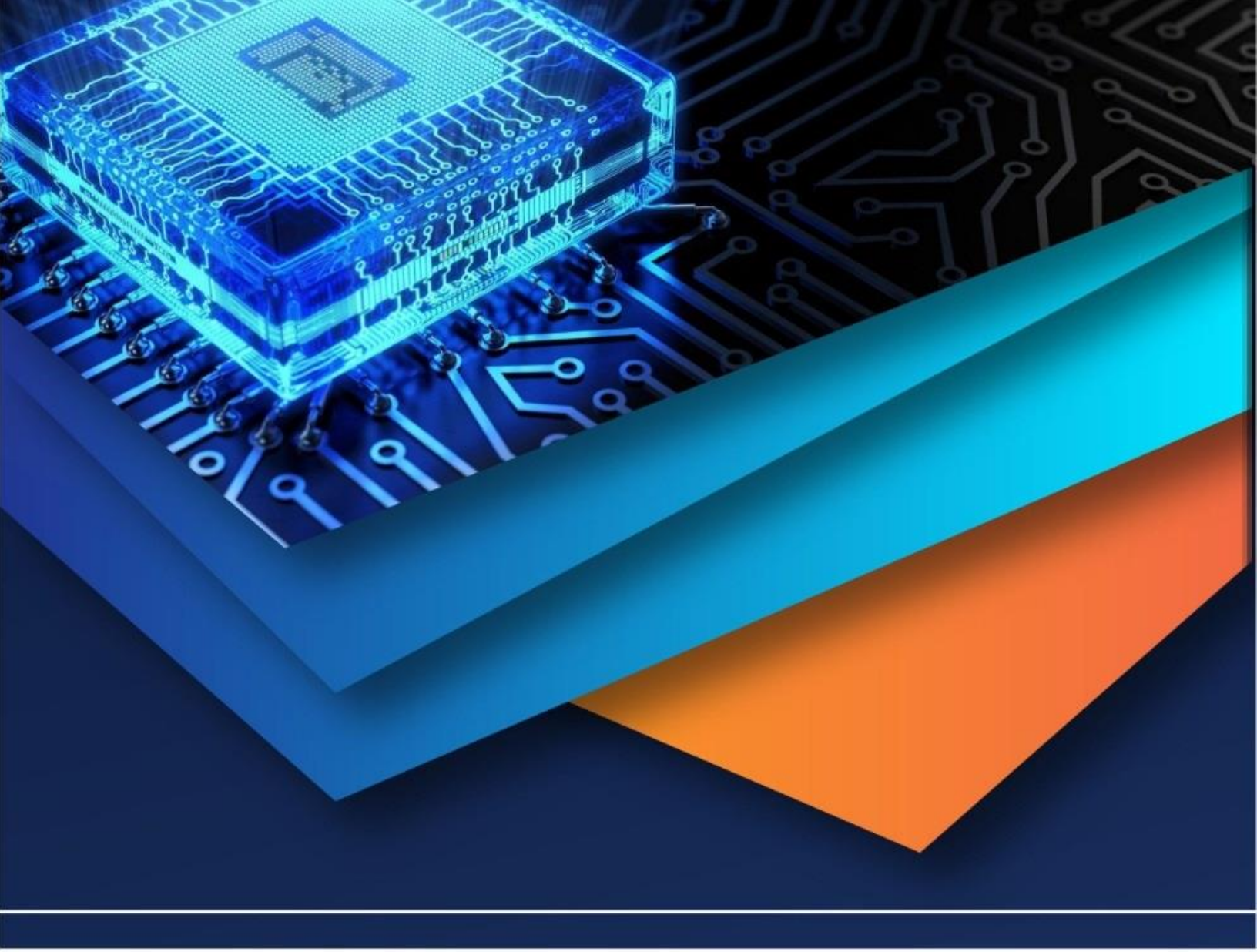
[4] C. A. Choquette-Choo, F. Tramèr, N. Carlini, et al., “Membership inference with access to labels only,” in Proc. International Conference on Machine Learning (ICML), pp. 1964–1974, 2021.

[5] A. Kumar, R. S. Raj, P. Yadav, and M. Singh, “Blockchain-enhanced secure deployment methodology for AIaaS,” Journal of Cryptographic Engineering, vol. 15, no. 2, pp. 125–142, 2024.

[6] M. S. Rahman, T. Ahmed, and M. M. Rahman, “Comprehensive survey on homomorphic encryption usage in secure AI evaluation,” International Journal of Information Security, vol. 23, no. 4, pp. 379–394, 2023.

[7] P. Sharma, S. Bansal, and S. Jain, “Strengthening AI-as-a-Service using fully homomorphic encryption alongside federated learning,” Journal of Cloud Computing: Advances, Systems, and Applications, vol. 13, no. 1, article 45, 2024.

[8] L. Liu, J. Ma, and F. Zhao, “Frameworks and strategies for AI model confidentiality using homomorphic encryption,” IEEE Access, vol. 12, pp. 10433–10445, 2024.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY