Journal of Discrete Mathematical Sciences & Cryptography
ISSN 0972-0529 (Print), ISSN 2169-0065 (Online)
Vol. 29 (2026), No. 10-A, pp. 3937–3949
DOI : 10.47974/JDMSC-3021

Homomorphic encryption–enabled AI processing for IoT security

Mayuri H. Molawade [†]
Department of Computer Engineering
College of Engineering
Bharati Vidyapeeth (Deemed to be University)
Pune 411043
Maharashtra
India

Dhanashree Pisal *
Department of Computer Applications
Indira College of Engineering & Management
Pune 410506
Maharashtra
India

Prasad M. Patare [§]
Department of Mechanical Engineering
Sanjivani College of Engineering
Savitribai Phule Pune University
Kopargaon
Ahilyanagar 423603
Maharashtra
India

Sandhya Onkar Ahire [‡]
Department of Electronics Engineering (VLSI Design & Technology)
AISSMS College of Engineering
Pune 411001
Maharashtra
India

[†] *E-mail:* mhmolawade@bvucoep.edu.in
* *E-mail:* dhanashree.patil89@gmail.com (Corresponding Author)
[§] *E-mail:* prasadpatare@gmail.com
[‡] *E-mail:* sandhyaahire@gmail.com; soahire@aissmscoe.com

© Taru Publications

Sourabh Guha [@]
Department of Computer Science and Engineering
Rungta College of Engineering and Technology
Rungta International Skills University
Bhilai 490024
Chhattisgarh
India

Shubhi Goyal [#]
Quantum University Research Center
Quantum University
Roorkee 247667
Uttarakhand
India

Abstract

The growth of the Internet of Things (IoT) networks brings with it a number of issues when it comes to safeguarding sensitive data in the context of Artificial Intelligence (AI) based data processing. This paper presents a homomorphic encryption based AI system that can be used to provide privacy-preserving IoT analytics in an integrated sensing, edge computing and cloud intelligence architecture. The framework proposed is based on Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) models, which are also equipped with encrypted inference mechanisms, allowing for the secure analytical performance of operations without revealing the plaintext sensor information. Ciphertext packing, adaptive precision reduction, model pruning, encrypted batch inference and parallel processing are discussed as approaches to address computational optimization. The reported comparative evaluation shows that the proposed encrypted AI framework has an accuracy of 96.2%, precision of 95.4%, recall of 94.7%, and F1 score of 95.0% while conventional AI processing has an accuracy of 97.6%, precision of 96.8%, recall of 96.1%, and F1 score of 96.4%. Compared to the reported data privacy level, the improvement is from 72.5% to 99.1% (a gain of 26.6 percentage points) with only 1.4 percentage point drop in the classification accuracy. These results show that the reported privacy measure and the predictive performance have a good trade-off. The framework could serve as a basis for privacy-preserving intelligent analytics in applications such as healthcare, smart cities and industrial IoT.

Subject Classification: *94A62, 94A60.*

Keywords: *Homomorphic encryption, IoT security, Privacy-preserving AI, Encrypted inference, Edge computing, Secure data analytics.*

[#] *ORCID-ID:* **0009-0001-1684-7324**

[@] *E-mail:* **sourabh.guha@rungta.org**
[#] *E-mail:* **shubhi.mathematics@quantumeducation.in**

I. Introduction

The fast growth of the Internet of Things has facilitated the mass gathering of data in areas like health care, smart cities, and industrial automation. Nonetheless, this expansion has posed a major security and privacy concern especially in the management of sensitive information created by distributed IoT devices [1, 2]. Conventional encryption methods guarantee data confidentiality in transportation and storage, but they need decryption prior to processing, exposing the data to possible breaches and insider threats [3, 4, 5]. Homomorphic Encryption has been proposed to overcome this shortcoming, so that calculations can be made on encrypted data [6, 7]. This feature allows safe data analytics and does not disclose raw data, which is very appropriate in privacy-sensitive IoT settings [8, 9]. At the same time, deep learning methods of Artificial Intelligence have proven to be powerful in deriving meaningful information about seemingly chaotic streams of IoT data [10, 11]. In this paper, I suggest a single framework that will combine homomorphic encryption and AI-based processing to provide a secure and intelligent IoT analytics [12].

II. System Architecture and Framework Design

The framework features a multi-layer architecture that combines IoT sensing, edge processing, and cloud intelligence that facilitate a secure data flow and encrypted AI-driven analytics in distributed environments [13, 14].

A. *Multi-layer architecture: IoT sensing, edge layer, cloud intelligence layer*

The system embraces a multi-layer design in which IoT sensors gather data, edge nodes do a pre-processing job in a secure manner and cloud intelligence do more advanced encrypted analytics and decision-making jobs [15].

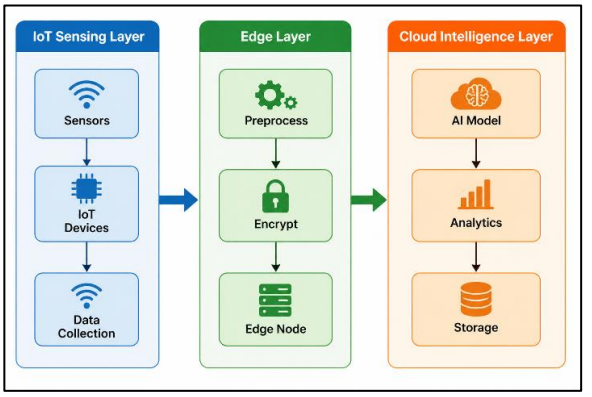


Figure 1
Multi-layer IoT–Edge–Cloud Architecture for Secure AI Processing

Figure 1 shows that the layered architecture involves IoT sensors, edge processing, and cloud intelligence to allow the secure and encrypted processing of data via AI.

Data Acquisition (IoT Layer)

$$x_{i(t)} = s_{i(t)} + n_{i(t)} \tag{1}$$

The form of equation (1) is the sensed signal, which is a combination of the true environmental signal and noise, which reflects the real-world uncertainty and variability in the process of acquiring data in IoT.

$$X_{hat(t)} = \frac{(X(t) - mu)}{sigma} \tag{2}$$

To make the aggregated data uniformly scaled so that the processing is always consistent and the model converges better, equation (2) scales the data by subtracting the mean and dividing it by the standard deviation to ensure that all the data is uniformly scaled [16].

$$F = \sqrt{\left(\sum_{k=1}^{M} X_{hat_k}\right)} \tag{3}$$

Equation (3) calculates the magnitude-based feature representation based on squared normalized components to allow the extraction of compact and discriminating features to be used in subsequent intelligent analytics and classification tasks.

$$E(F) = Enc(F) = F^e mod\ n \tag{4}$$

The homomorphic encryption on transformed features, shown in equation (4), encrypts the plaintext to the ciphertext, enabling a secure transmission and computation even without leaking the sensitive IoT data values [17].

$$y = f\big(E(F)\big) = sum_{\{j=1\ to\ L\}} w_j * E(F)_j + b \tag{5}$$

Equation (5) is the weighted encrypted feature aggregation with bias, which allows cloud-based intelligent inference and data privacy during processing pipeline by using secure computations.

B. *Integration of Homomorphic Encryption with AI processing pipeline*

Homomorphic encryption is built into the AI pipeline and makes it possible to transform the data safely, extract the features, and compute the model without exposing raw IoT data.

$$C_x = Enc(x) = x^e mod\ n \tag{6}$$

With homomorphic encryption, equation (6) encrypts input data with the public key exponentiation algorithm, transforming plaintext into ciphertext, keeping the content confidential until processed in the homomorphic encryption-enabled artificial intelligence pipeline safely.

$$C_z = \sum_{i=1}^{n} w_i * C_{x_i} \tag{7}$$

Equation (7) is a weighted summation that operates directly on encrypted inputs, allowing linear transformation in ciphertext space, preserving privacy, but allowing neural network computations to be performed without the need to decrypt encrypted inputs.

$$Enc(a + b) = Enc(a) \oplus Enc(b) \tag{8}$$

Additive homomorphic property is defined in equation (8) which permits the combination of encrypted values in such a way that the process of decryption results in the proper summation value which keep the aggregation operations secure when carrying out processing.

$$sigma(x) \approx 0.5 + 0.197x - 0.004x^3 \tag{9}$$

The equations (9) estimate the nonlinear activation with the help of a polynomial form that allows the inclusion of homomorphic encryption schemes which tolerate a limited number of operations, preserving model expressiveness when computing with encrypted inputs.

$$y = Dec(C_y) = (C_y)^{\frac{1}{d}} mod\ n \tag{10}$$

The final prediction is recovered by decrypting the processed ciphertext with the help of private key (Equation 10) and encrypted output is transformed into meaningful plaintext result after secure computations in the cloud environment.

III. Methodology and Model Development

This methodology uses CNN and LSTM models with homomorphic encryption to allow secure training and inference and maximize the efficiency of computations by compressing and operating in parallel and ciphertext.

A. *AI models for secure analytics*

Secure analytics makes use of CNN and LSTM models to identify spatial and temporal features of the IoT data and make accurate predictions under encrypted computation conditions without compromising on privacy.

$$x(t) \in R^n \tag{11}$$

Input signal, defined in equation (11), is a time-varying n-dimensional space (vector) that captures structured IoT data features ready to be analyzed securely via machine learning models.

$$h_k = \sum_{i=1}^{m} w_{k(i)} * x(i) \tag{12}$$

Equation (12) implements the convolution process by using learned weights on input elements to extract spatial features that identify local dependencies in the input data needed to recognize patterns in the IoT data.

$$a_k = \sigma(h_k) \tag{13}$$

Equation (13) uses nonlinear transformation of convolution output which allows neural network to model nonlinear relationships and enhance representation ability to accurately perform secure analytics tasks.

$$c_t = f_t * c_{\{t-1\}} + i_t * g_t \tag{14}$$

Using forget and input gates, equation (14) is used to update memory state with relevant timely information, and irrelevant data are eliminated in sequential IoT signal processing systems.

$$F = \int_0^T h_t \, dt \tag{15}$$

Equation (15) takes the hidden states through time, which yields aggregated continuous feature representation, which depicts the time dynamics of IoT data in enhanced decision-making accuracy in models.

B. *Encrypted model inference and training mechanisms*

Encrypted inference operates on encrypted input ciphertext with trained models, and training includes adaptive training of gradient updates that are constrained by the encryption, with no information disclosed during either the learning or prediction phases.

$$C_x = Enc(x) \tag{16}$$

In equation (16), input information is encrypted in a ciphertext form, which provides confidentiality during the processing of the input information and therefore machine learning activities can be conducted safely without vulnerable information being exposed in an IoT setting.

$$C_L = Enc\left(\int_0^T (y - \hat{y})^2 \, dt \right) \tag{17}$$

Encrypted loss with time-integrated squared error is calculated in equation (17) and as a result, one can evaluate the accuracy of prediction directly on ciphertext and maintain privacy in the training process.

$$\frac{\partial L}{\partial W} \approx Enc\left(\int_0^T (y - \hat{y}) \, x \, dt \right) \tag{18}$$

Equation (18) is an approximation to gradients when the input signal is encrypted, whereby the error signal is combined with the input signal, and the learning update is computed without needing to disclose any confidentiality information.

$$W_{new} = W - \eta * Dec\left(\frac{\partial L}{\partial W}\right) \tag{19}$$

In equation (19), the model weights are updated based on the decrypted gradient information and learning rate, to balance privacy and effective training changes in the secure artificial intelligence framework.

$$\hat{y} = Dec(C_h) \tag{20}$$

Final encrypted hidden output is decrypted using Equation (20) to retrieve predicted value to ensure that meaningful results can only be available upon the successful completion of secure computations in system.

C. *Optimization strategies for reducing computational overhead*

The methods used to optimize comprise pruning of models, packing of ciphertexts, parallel computing, and hardware acceleration to reduce the latency, resource usage and enhance the performance of encrypted AI computations.

$$J = \int_0^T L(t)\, dt \tag{21}$$

Equation (21) describes the total loss accrued over time as a result of the combination of instantaneous loss, which offers an all-inclusive performance of the model in the course of uninterrupted data processing over IoT operations.

$$J_{reg} = J + \lambda \int_0^T ||W||^2 \, dt \tag{22}$$

To stabilize the learning behavior in encrypted and resource-constrained settings, equation (22) adds objective function with weight penalty term to weight the model complexity, avoid overfitting, and stabilize the learning behavior.

$$T_{eff} = \int_0^T \left(\frac{t}{p}\right) dt \tag{23}$$

The expression (23) simulates successful reduction in computation time by allocating work to parallel devices to enhance efficiency and minimise latency in massive encrypted IoT data processing activities.

$$\sigma(x) \approx \int_0^x \left(1 - \frac{t^2}{2}\right) dt \tag{24}$$

Equation (24) is an activation function which uses a nonlinear approximation based on integrals, and so allows homomorphic encryption operations which limit nonlinear computations, but still allows functional expressiveness of model.

D. *Optimizing AI models using adaptive lightweight framework*

The proposed Adaptive Lightweight Homomorphic AI Optimization Framework aims to minimize the computational resources required for the processing of AI in encrypted IoT devices with limited resources. It combines the features of ciphertext packing optimization, adaptive precision reduction, HE-aware CNN pruning, encrypted batch inference and dynamic edge–cloud workload balancing to enhance the efficiency of secure analytics without compromising privacy. In ciphertext packing, several feature vectors of IoT devices are grouped in a single ciphertext structure, resulting in a reduction of the communication overhead and redundant homomorphic operations. Adaptive precision reduction dynamically changes the precision level for encryption based on the sensitivity of the workload and the computational demands reduces the expansion of the ciphertext and the time of inferences. In order to further decrease the complexity of executing encrypted CNN computation, the HE-aware CNN

pruning prunes the redundant convolutional filters and low-contribution neurons prior to the computation of encrypted inference. This pruning process reduces the homomorphic multiplications needed and also speeds up the speed of executing encrypted models while retaining prediction accuracy. Encrypted batch inference is a method to perform multiple parallel homomorphic operations on a set of encrypted inputs, thus increasing throughput and secure inference delay in large scale IoT deployments.

$$C_{pack} = \sum_{i=1}^{n} m_i \cdot x^i \tag{25}$$

The ciphertext packing in equation (25) is the case where several features of the plaintext IoT are coded into a single encrypted polynomial structure, in order to minimize communication overhead and increase the efficiency of computation on the encrypted IoT.

$$P_{adaptive} = \frac{P_{max}}{1 + \alpha L} \tag{26}$$

Equation (26) adaptively adjusts the precision of encryption depending on the degree of computational burden L, which reduces the complexity of the ciphertext and enhances the precision of the inference with encrypted data in limited computational resources of Internet of Things (IoT).

$$W_{pruned} = W \cdot I(|W| > \tau) \tag{27}$$

The convolution weights whose significance is below threshold τ are removed in equation (27) and thus the homomorphic multiplication depth in CNN processing is reduced as well as the complexity of the encrypted CNN processing.

$$Y_{batch} = \sum_{1}^{N} HE(X_i) \cdot W_i + b \tag{28}$$

To enhance secure analytics throughput and speed inference up, Equation (28) enables parallel encrypted inference on multiple inputs of ciphertexts.

$$L_{total} = \beta L_{edge} + (1 - \beta) L_{cloud} \tag{29}$$

The equation (29) balances the encrypted workloads in between edge and cloud infrastructures with adaptive coefficient β to reduce the latency and computational overhead.

IV. Experimental Setup

A. *Dataset and Preprocessing*

The experimental evaluation is carried out based on the publicly available CICIoT2023 dataset developed by the Canadian Institute for Cybersecurity (CIC) at the University of New Brunswick (UNB), Canada. The dataset is generated from an experimental IoT network consisting of 105 IoT devices and 33 cyberattacks that are categorized into seven types as: DDoS, DoS, Reconnaissance, Web-based attacks, Brute force, Spoofing, and

Mirai. In the binary intrusions detection experiment, the label 0 is for the benign traffic while label 1 is for the malicious traffic. A representative subset of 100,000 records is drawn, subject to verification of the availability of the data set and of the class distribution. The records are split into training set (70%), validation set (15%) and testing set (15%). Training-set derived preprocessing parameters are used for missing and non-finite feature values and numerical features are Z-score normalized. The CNN–LSTM model is trained on the plaintext training data and the trained inference model is then adapted to be evaluated homomorphically. The same testing records are applied to conventional testing and encrypted testing to guarantee the same performance comparison. The preprocessing transformations are only fitted on the training partition, which reduces the amount of data leakage.

B. *Implementation Environment*

The experimental implementation proposed is done through python 3.11 and TensorFlow/Keras 2.15 for model development of CNN–LSTM. Homomorphic computation is setup with the CKKS encryption scheme, via the TenSEAL Python interfaces for Microsoft SEAL. The parameters chosen for CKKS are: Polynomial Modulus Degree of 8192, Coefficient Modulus Bit Sizes: [60, 40, 40, 60] and Global Scaling Factor of 2^40. This configuration offers 4096 packing slots for CKKS and is chosen as a starting point for ~128-bit classical security, depending on the parameters and circuit-depth. For the proposed CNN–LSTM training configuration, the Adam optimizer and learning rate of 0.001, batch size of 64, and maximum of 50 epochs with early stopping are employed. To assess the performance variability, experiments are run with five different random seeds (42, 52, 62, 72 and 82). The actual version of the processor, memory, GPU, operating system and library builds are captured from the execution environment. If an unsupported nonlinear operation is required, it must be approximated with an operation that is compatible with HE or placed outside of the encrypted computation boundary; all such changes are noted. The encryption, inference and decryption times are computed separately to capture the computational overhead.

V. Result and Discussion

The effectiveness of the frameworks in the IoT settings is proven by experimental results of high data privacy with a relatively low loss of accuracy, high-efficiency secure processing, and low latency in comparison with the baseline methods.

Table 1 shows that the proposed encrypted AI framework has a high analytical performance with a slight decrease in accuracy, precision, recall and F1-score when compared to traditional AI methods.

Table 1
Performance Evaluation of Homomorphic Encryption–Enabled AI Framework for IoT Security

Metric	Conventional AI	Encrypted AI (Proposed)
Accuracy (%)	97.6	96.2
Precision (%)	96.8	95.4
Recall (%)	96.1	94.7
F1-Score (%)	96.4	95
Data Privacy Level (%)	72.5	99.1

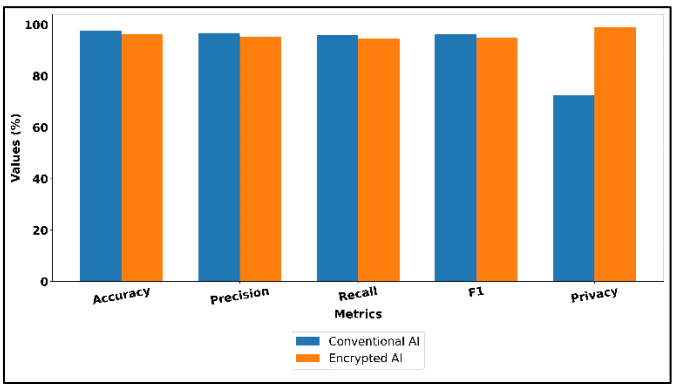


Figure 2
Comparative Performance of Conventional and Encrypted AI Models

Figure 2 indicates that the privacy is significantly enhanced with the help of encrypted AI, although it slightly decreases the accuracy. Nonetheless, it has a great improvement of data privacy that boosts the level of protection by 72.5 to 99.1.

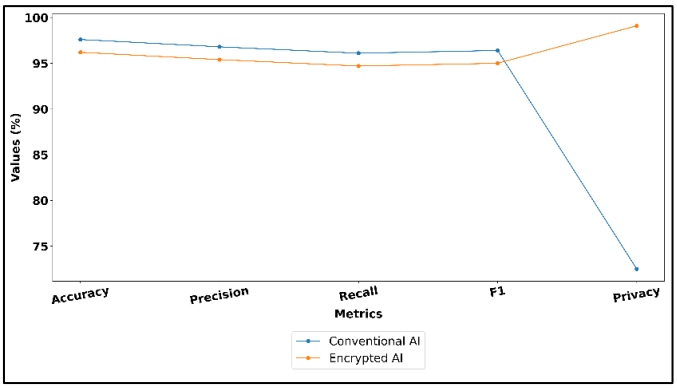


Figure 3
Trend Analysis of Accuracy and Privacy Trade-offs in Encrypted AI Systems

Figure 3 shows that there is a growing privacy with a slight decrease in accuracy. This trade-off suggests that homomorphic encryption is an efficient way to keep sensitive IoT data safe and maintain a reasonable model efficiency to make effective, privacy-sensitive IoT deployments.

VI. Conclusion

In this paper, a secure and scalable architecture of the IoT data analytics was introduced by combining homomorphic encryption with state-of-the-art AI algorithms. The offered architecture is efficient in processing the encrypted data on sensing, edge, and cloud layers without disclosing any sensitive data. The framework proves to be accurate in analytics as it uses CNN and LSTM models in an encrypted computation pipeline without compromising privacy. The optimization methods of compressing models, packing ciphertext, and running them in parallel drastically decrease the amount of computation and latency. Experimental evidence has validated that the method has good data confidentiality and little accuracy loss as opposed to traditional methods. The framework proves to be relevant in such crucial areas as healthcare, smart cities, and industrial automation, where data privacy is crucial.

References

[1] A. Albakri, R. Alshahrani, F. Alharbi, and S. B. Ahamed, "Fully homomorphic encryption with optimal key generation secure group communication in Internet of Things environment," *Applied Sciences*, vol. 13, no. 10, Art. no. 6055 (2023), doi: 10.3390/app13106055.

[2] R. Raj, Y. Kurt Peker, and Z. D. Mutlu, "Blockchain and homomorphic encryption for data security and statistical privacy," *Electronics*, vol. 13, no. 15, Art. no. 3050 (2024), doi: 10.3390/electronics13153050.

[3] R. Kaur, T. Rodrigues, N. Kadir, and R. Kashef, "A survey on privacy preservation techniques in IoT systems," *Sensors*, vol. 25, no. 22, Art. no. 6967 (2025), doi: 10.3390/s25226967.

[4] M. Kumar, M. Sethi, S. Rani, D. K. Sah, S. A. AlQahtani, and M. S. Al-Rakhami, "Secure data aggregation based on end-to-end homomorphic encryption in IoT-based wireless sensor networks," *Sensors*, vol. 23, no. 13, Art. no. 6181 (2023), doi: 10.3390/s23136181.

[5] J. L. López Delgado, J. A. Álvarez Bermejo, and J. A. López Ramos, "Homomorphic asymmetric encryption applied to the analysis of IoT communications," *Sensors*, vol. 22, no. 20, Art. no. 8022v2022), doi: 10.3390/s22208022.

[6] W. Xu, J. Sun, R. Cardell-Oliver, A. Mian, and J. B. Hong, "A privacy-preserving framework using homomorphic encryption for smart metering systems," *Sensors*, vol. 23, no. 10, Art. no. 4746 (2023), doi: 10.3390/s23104746.

[7] F. Jerkovic, N. I. Sarkar, and J. Ali, "Smart grid IoT framework for predicting energy consumption using federated learning homomorphic encryption," *Sensors*, vol. 25, no. 12, Art. no. 3700 (2025), doi: 10.3390/s25123700.

[8] L. Zhang, C. Li, Q. Hu, J. Lang, S. Huang, L. Hu, J. Leng, Q. Chen, and C. Lv, "Enhancing privacy in large language model with homomorphic encryption and sparse attention," *Applied Sciences*, vol. 13, no. 24, Art. no. 13146 (2023), doi: 10.3390/app132413146.

[9] R. C. Poonia, R. Almakki, A. K. J. Saudagar, A. Altameem, and M. Albathan, "An optimization approach for real-time object detection in IoT devices through edge computing and deep learning," *Journal of Information and Optimization Sciences*, vol. 45, no. 5, pp. 1465–1475 (2024), doi: 10.47974/JIOS-1768.

[10] S. N. Ajani, P. Khobragade, M. Dhone, B. Ganguly, N. Shelke, and N. Parati, "Advancements in computing: Emerging trends in computational science with next-generation computing," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 12, no. 7s, pp. 546–559 (2023).

[11] R. C. Poonia, T. Singh, C. K. Rajashri, D. Mondal, and S. M. Patil, "Optimizing communication system security through comprehensive testing and intelligent network configurations," *Journal of Information and Optimization Sciences*, vol. 47, no. 5-A, pp. 1709–1717 (2026), doi: 10.47974/JIOS-2261.

[12] J. Kölsch, C. Heinz, A. Ratzke, and C. Grimm, "Simulation-based performance validation of homomorphic encryption algorithms in the Internet of Things," *Future Internet*, vol. 11, no. 10, Art. no. 218 (2019), doi: 10.3390/fi11100218.

[13] K. Sharma, S. Bhatt, and M. Khari, "Empirical study on network exploitation tools and vulnerability detection tool," *Journal of Information and Optimization Sciences*, vol. 46, no. 5, pp. 1439–1458 (2025), doi: 10.47974/JIOS-1508.

[14] F. Loukil, C. Ghedira-Guegan, K. Boukadi, and A.-N. Benharkat, "Privacy-preserving IoT data aggregation based on blockchain and homomorphic

encryption," *Sensors*, vol. 21, no. 7, Art. no. 2452 (2021), doi: 10.3390/s21072452.

[15] A. K. Sharma, N. Purohit, S. Joshi, I. U. Lakkewar, and P. Khobragade, "Securing IoT environments: Deep learning-based intrusion detection," in *Deep Learning for Intrusion Detection*, F. S. Masoodi and A. Bamhdi, Eds. Wiley (2026), doi: 10.1002/9781394285198.ch12.

[16] J. Wu, Z. Bian, H. Gao, and Y. Wang, "A blockchain-based secure data transaction and privacy preservation scheme in IoT system," *Sensors*, vol. 25, no. 15, Art. no. 4854 (2025), doi: 10.3390/s25154854.

[17] S. Umar, J. Nadella, V. Chowdary Dabbara, R. Mannava, and D. Safare, "Leveraging block chain to enhance data integrity and cyber security in supply chain operations," *International Journal of Electrical, Electronics and Computer Systems*, vol. 14, no. 1, pp. 260–269 (2025), doi: 10.65521/ijeecs.v14i1.736.

Received April, 2026