International Conference on Science, Engineering & Management Trends

International Journal of Science, Engineering and Technology
ISSN: 2348-4098, P-ISSN: 2395-4752

Layered Privacy Protection with Communication-Aware Design for Federated Intrusion Detection in Resource-Constrained IoT Networks

Sujata Pattnaik*[1], Dr. Ridhika Sidhi*[2]

*[1]Research Scholar, Department of Computer Science, William Carey University, Meghalaya, India

*[2]Phd. Guide, Department of Computer Science, William Carey University, Meghalaya, India

Abstract- Federated learning (FL) enables Internet of Things (IoT) clients to train a shared intrusion detection model without centralising raw network traffic, but it neither guarantees privacy nor comes free of communication cost. Model updates can leak information about local traffic, and repeated update exchange can strain constrained IoT links. This paper presents a layered privacy design for federated intrusion detection together with a communication-aware analysis of its cost. Five layers—local data retention, update clipping with differential privacy, secure aggregation, encrypted transport, and authenticated participation—are each mapped to the threat they address, the parameters that must be reported and the overhead they add. An analytic cost model combines model size, numeric encoding, participating clients and rounds. For a lightweight 5,352-parameter detector, the per-client upload per round falls from 20.9 KB with 32-bit encoding to 5.2 KB with 8-bit quantisation, compared with 195.3 KB for a 50,000-parameter convolutional model, showing that model choice and encoding dominate communication cost. The paper also specifies an ablation plan for measuring the privacy–utility–cost balance on IoT intrusion datasets.

Keywords- Internet of Things, federated learning, intrusion detection, differential privacy, secure aggregation, communication efficiency, lightweight models

I. INTRODUCTION

IoT deployments span homes, industrial plants, healthcare facilities and city infrastructure, and their devices are frequent targets of botnets and denial-of-service attacks [1]. Machine-learning intrusion detection systems (IDS) can detect such activity [2], but training them centrally requires collecting sensitive traffic from all sites. Federated learning (FL) avoids this by exchanging model updates rather than data [3], and several federated IDS for IoT have been demonstrated [4], [5].

Two practical problems remain. First, privacy: shared updates can reveal properties of local training data [6], so data locality alone is not a privacy guarantee [7]. Second, cost: FL replaces a one-time data transfer with repeated update exchange, and privacy mechanisms add further computation and messages—a significant burden for constrained IoT clients [8]. These two problems interact, yet many studies report detection accuracy without the privacy parameters or communication volume needed to judge a deployment [9], [10].

International Journal of Science, Engineering and Technology
ISSN: 2348-4098, P-ISSN: 2395-4752

This paper contributes: (1) a five-layer privacy design for federated IDS in which each layer is linked to a threat, a reportable parameter and a cost; (2) an analytic communication model that includes numeric encoding; (3) a quantitative comparison of per-round upload size across model sizes and encodings; and (4) an ablation plan for measuring the privacy–utility–cost balance.

II. RELATED WORK

Differential privacy (DP) bounds the influence of any individual on a released result [11]. DP-SGD achieves this for deep learning by clipping and adding Gaussian noise [12], and client-level DP extends the idea to FL [13]; Wei *et al.* analysed the trade-off between privacy level and convergence in DP-based FL [14]. Secure aggregation lets a server compute the sum of client updates without seeing any individual update [15], while homomorphic encryption offers similar protection at higher computational cost [16]. For communication, structured and sketched updates [8] and sparsification for non-IID data [17] reduce upload size. Reviews of federated IDS note that these mechanisms are seldom combined and measured together in IoT settings [7], [18].

III. LAYERED PRIVACY DESIGN

1. System Overview

Fig. 1 shows the communication flow. Clients such as a smart-home gateway, an industrial IoT node and a healthcare edge device train a local IDS model. Each sends a protected model update over an encrypted channel to the federated server, which performs secure aggregation, updates the global model and returns it in encrypted form. Raw traffic is never transmitted.

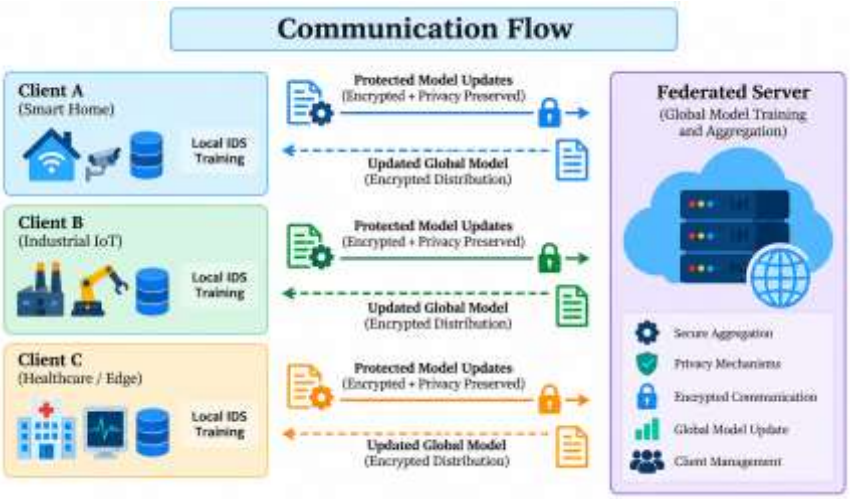


Fig. 1. Communication flow of the federated IDS. Clients send protected (privacy-preserved, encrypted) model updates and receive the encrypted global model; the server performs secure aggregation, applies privacy mechanisms and manages clients. Raw traffic records are not transferred.

2. Update Protection with Differential Privacy

Client k computes its update $\Delta_k = w_k^{t+1} - w^t$, clips it to norm C and adds Gaussian noise with multiplier σ [12], [13]:

$$\tilde{\Delta}_k = \Delta_k \cdot \min(1, C / \|\Delta_k\|_2) + N(0, \sigma^2 C^2 I) \quad (1)$$

The mechanism satisfies (ε, δ)-DP, meaning that for neighbouring inputs D and D' and any output set S [11],

International Journal of Science, Engineering and Technology
ISSN: 2348-4098, P-ISSN: 2395-4752

$$\Pr[M(D) \in S] \leq e^{\varepsilon} \Pr[M(D') \in S] + \delta \qquad (2)$$

The cumulative ε over all rounds depends on σ, the client sampling rate q and the number of rounds R, and is computed with a privacy accountant [12].

3. Secure Aggregation

Clients add pairwise masks agreed by key exchange, $y_k = \tilde{\Delta}_k + \Sigma_{j>k} s_{kj} - \Sigma_{j<k} s_{jk}$; the masks cancel in the sum, so the server learns only the aggregate [15]. Secret-shared mask seeds allow the sum to be completed when clients drop out, which is essential for intermittently connected IoT devices.

4. Mapping Layers to Threats and Costs

Table I links each layer to what it protects, what must be reported and what it costs. Clipping also bounds the effect of a poisoned update, so it supports robustness as well as privacy [19].

Table 1: Privacy Layers: Protection, Reported Parameters and Cost

Layer	Protects against	Parameters to report	Overhead
L1 Local data retention	Direct exposure of raw traffic	Client count, data per client	Repeated update exchange
L2 Clipping + DP	Inference of records or clients from updates [6]	C, σ, q, R, (ε, δ)	Negligible compute; utility loss
L3 Secure aggregation	Server inspection of individual updates	Threshold, dropout tolerance	Key-agreement messages per round
L4 Encrypted transport	Eavesdropping, tampering in transit	Protocol and cipher suite	Handshake and encryption
L5 Authentication	Unauthorised and Sybil clients	Enrolment method	Credential management

IV. COMMUNICATION-AWARE COST MODEL

Let P be the number of model parameters, b the bytes per transmitted value, K the clients participating per round, R the number of rounds and A the per-client secure-aggregation overhead per round. Total traffic is

$$C_{total} = R \cdot K \cdot (P \cdot b_{down} + P \cdot b_{up} + A) \qquad (3)$$

where downlink and uplink encodings may differ. Quantising the uplink to 8 bits gives $b_{up} = 1$ plus a small scale factor; top-r sparsification sends a fraction r of values with their indices, costing about $r \cdot P \cdot 8$ bytes with 32-bit values and indices. Note that quantisation and sparsification must be applied after noise addition so that the DP guarantee is preserved by post-processing [20].

Fig. 2 applies the model to two detectors: a lightweight multilayer perceptron (46 inputs, hidden layers of 64 and 32 units, 8 output classes; 5,352 parameters) and a compact 1-D convolutional network of about 50,000 parameters. The lightweight model uploads 20.9 KB per round at 32-bit precision, 10.5 KB at 16 bits, 5.2 KB with 8-bit quantisation and 4.2 KB with 10% sparsification. The convolutional model needs 195.3 KB at 32 bits—more than nine times the lightweight model—and still 39.1 KB when sparsified. With ten clients and 100 rounds, the 32-bit lightweight model moves about 43 MB in total, compared with about 400 MB for the convolutional model. Model size therefore has a larger effect than any single encoding choice, which supports selecting the smallest model that meets detection requirements. These values are analytic estimates that exclude protocol headers and secure-aggregation messages, which are to be measured directly.

International Journal of Science, Engineering and Technology
ISSN: 2348-4098, P-ISSN: 2395-4752

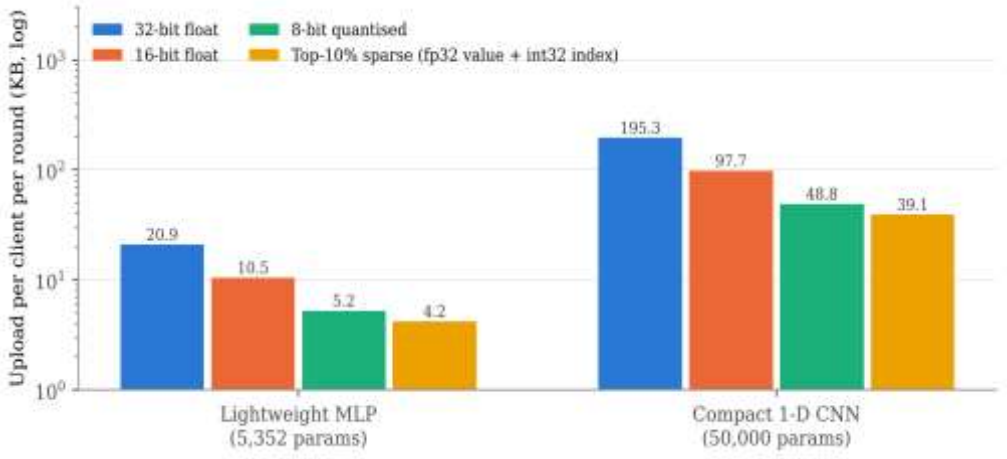


Fig. 2. Analytic per-client upload per round for two IDS model sizes under four encodings (log scale). Values follow Eq. (3) without protocol or secure-aggregation overhead; they are design estimates, not measurements.

V. EVALUATION PLAN

The design is evaluated on public IoT intrusion datasets such as CICIoT2023 [21] with logical client partitions. An ablation adds layers cumulatively, so that each change in utility and cost can be attributed to one mechanism (Table II). Each configuration is run with five seeds and evaluated on a common test set using macro-F1, per-class recall and false-positive rate, alongside measured bytes exchanged, rounds to convergence, client training time and the reported (ε, δ). Privacy–utility curves are produced for target $\varepsilon \in \{1, 4, 8, \infty\}$ at $\delta = 10^{-5}$.

Table 2: Ablation Configurations

ID	Enabled layers and encoding	Question addressed
C0	Centralised training (reference)	Upper reference for detection quality
C1	L1 + L4 + L5, 32-bit (standard FL)	Cost of federation without update protection
C2	C1 + clipping	Effect of bounding update norms on utility and robustness
C3	C2 + Gaussian noise (DP)	Privacy–utility trade-off across ε
C4	C3 + secure aggregation	Added message and time overhead
C5	C4 + 8-bit uplink quantisation	Communication saving and any utility loss

VI. DISCUSSION

The layered design makes explicit that privacy in federated IDS is a set of mechanisms with distinct guarantees, not a single property. DP is the only layer that offers a quantifiable guarantee against inference from the trained model, but its noise is likely to affect rare attack classes most, since their signal is weakest [14]; per-class recall is therefore reported rather than accuracy alone. Secure aggregation protects individual updates from the server but not the aggregate, and its message cost grows with the number of clients, which favours aggregating at gateways in large deployments. The cost model shows that a compact detector combined with uplink quantisation can reduce upload by roughly 37 times relative to a 32-bit, 50,000-parameter model, which may decide whether federated IDS is feasible on low-bandwidth IoT links. Limitations include the honest-but-curious server assumption and the use of analytic rather than measured traffic in this paper.

VII. CONCLUSION

This paper proposed a five-layer privacy design for federated intrusion detection in IoT and connected each layer to its protection, reportable parameters and overhead. An analytic communication model showed that model size and numeric encoding dominate per-round traffic: a lightweight detector with 8-bit quantisation needs about 5 KB per client per round, compared with about 195 KB for a modest convolutional model at full precision. The ablation plan will quantify the joint privacy–utility–cost balance on IoT intrusion datasets, supporting deployment choices grounded in measurement rather than assumption.

REFERENCES

1. C. Kolias, G. Kambourakis, A. Stavrou, and J. Voas, "DDoS in the IoT: Mirai and other botnets," *Computer*, vol. 50, no. 7, pp. 80–84, 2017, doi: 10.1109/MC.2017.201.
2. A. Khraisat, I. Gondal, P. Vamplew, and J. Kamruzzaman, "Survey of intrusion detection systems: Techniques, datasets and challenges," *Cybersecurity*, vol. 2, Art. no. 20, 2019, doi: 10.1186/s42400-019-0038-7.
3. H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Proc. 20th Int. Conf. Artif. Intell. Statist. (AISTATS)*, Fort Lauderdale, FL, USA, 2017, pp. 1273–1282, doi: 10.48550/arXiv.1602.05629.
4. T. D. Nguyen, S. Marchal, M. Miettinen, H. Fereidooni, N. Asokan, and A.-R. Sadeghi, "DÏoT: A federated self-learning anomaly detection system for IoT," in *Proc. IEEE 39th Int. Conf. Distrib. Comput. Syst. (ICDCS)*, Dallas, TX, USA, 2019, pp. 756–767, doi: 10.1109/ICDCS.2019.00080.
5. V. Mothukuri, P. Khare, R. M. Parizi, S. Pouriyeh, A. Dehghantanha, and G. Srivastava, "Federated-learning-based anomaly detection for IoT security attacks," *IEEE Internet Things J.*, vol. 9, no. 4, pp. 2545–2554, 2022, doi: 10.1109/JIOT.2021.3077803.
6. L. Melis, C. Song, E. De Cristofaro, and V. Shmatikov, "Exploiting unintended feature leakage in collaborative learning," in *Proc. IEEE Symp. Secur. Privacy (SP)*, San Francisco, CA, USA, 2019, pp. 691–706, doi: 10.1109/SP.2019.00029.
7. A. Vyas, P.-C. Lin, R.-H. Hwang, and M. Tripathi, "Privacy-preserving federated learning for intrusion detection in IoT environments: A survey," *IEEE Access*, vol. 12, pp. 127018–127050, 2024, doi: 10.1109/ACCESS.2024.3454211.
8. J. Konečný, H. B. McMahan, F. X. Yu, P. Richtárik, A. T. Suresh, and D. Bacon, "Federated learning: Strategies for improving communication efficiency," 2016, *arXiv:1610.05492*, doi: 10.48550/arXiv.1610.05492.
9. S. Agrawal *et al.*, "Federated learning for intrusion detection system: Concepts, challenges and future directions," *Comput. Commun.*, vol. 195, pp. 346–361, 2022, doi: 10.1016/j.comcom.2022.09.012.
10. A. Belenguer, J. A. Pascual, and J. Navaridas, "A review of federated learning applications in intrusion detection systems," *Comput. Netw.*, vol. 258, Art. no. 111023, 2025, doi: 10.1016/j.comnet.2024.111023.
11. [11] C. Dwork, "Differential privacy," in *Proc. 33rd Int. Colloq. Automata, Lang. Program. (ICALP)*, Venice, Italy, 2006, pp. 1–12, doi: 10.1007/11787006_1.
12. [12] M. Abadi *et al.*, "Deep learning with differential privacy," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur. (CCS)*, Vienna, Austria, 2016, pp. 308–318, doi: 10.1145/2976749.2978318.
13. [13] R. C. Geyer, T. Klein, and M. Nabi, "Differentially private federated learning: A client level perspective," 2017, *arXiv:1712.07557*, doi: 10.48550/arXiv.1712.07557.
14. [14] K. Wei *et al.*, "Federated learning with differential privacy: Algorithms and performance analysis," *IEEE Trans. Inf. Forensics Security*, vol. 15, pp. 3454–3469, 2020, doi: 10.1109/TIFS.2020.2988575.

15. [15] K. Bonawitz *et al.*, "Practical secure aggregation for privacy-preserving machine learning," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur. (CCS)*, Dallas, TX, USA, 2017, pp. 1175–1191, doi: 10.1145/3133956.3133982.
16. [16] L. T. Phong, Y. Aono, T. Hayashi, L. Wang, and S. Moriai, "Privacy-preserving deep learning via additively homomorphic encryption," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 5, pp. 1333–1345, 2018, doi: 10.1109/TIFS.2017.2787987.
17. [17] F. Sattler, S. Wiedemann, K.-R. Müller, and W. Samek, "Robust and communication-efficient federated learning from non-i.i.d. data," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 9, pp. 3400–3413, 2020, doi: 10.1109/TNNLS.2019.2944481.
18. [18] A. Khraisat *et al.*, "Federated learning for intrusion detection in IoT environments: A privacy-preserving strategy," *Discover Internet Things*, vol. 5, Art. no. 72, 2025, doi: 10.1007/s43926-025-00169-7.
19. [19] L. Lyu *et al.*, "Privacy and robustness in federated learning: Attacks and defenses," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 7, pp. 8726–8746, 2024, doi: 10.1109/TNNLS.2022.3216981.
20. [20] C. Dwork and A. Roth, "The algorithmic foundations of differential privacy," *Found. Trends Theor. Comput. Sci.*, vol. 9, no. 3–4, pp. 211–407, 2014, doi: 10.1561/0400000042.
21. [21] E. C. P. Neto, S. Dadkhah, R. Ferreira, A. Zohourian, R. Lu, and A. A. Ghorbani, "CICIoT2023: A real-time dataset and benchmark for large-scale attacks in IoT environment," *Sensors*, vol. 23, no. 13, Art. no. 5941, 2023, doi: 10.3390/s23135941.