Blockchain-Enabled Federated Learning with Reputation for Secure and Trustworthy WBAN Systems

Monirul I Mahmud, Jan Bieniek and Mohamed Rahouti

Abstract

Wearable Body Area Networks (WBANs) and Internet-of-Things (IoT) devices generate sensitive health and activity data that can be harnessed to build predictive models for clinical monitoring and wellness applications. Centralizing such data is impractical due to privacy risks, communication overhead, and regulatory constraints. Federated Learning (FL) addresses this challenge by enabling distributed model training across edge devices, but remains vulnerable to data/model poisoning, Sybil attacks, inference risks, and unreliable clients. This proposal introduces a blockchain-enabled FL architecture tailored to WBAN/IoT deployments, combining cross-device learning with permissioned blockchain services for identity, access control, auditability, and incentive management. We design a novel *reputation-based mechanism* that evaluates contributions using validation improvement, consistency, and plausibility checks, updates participant reputations, and enforces access control, incentives, and slashing. Robust aggregation, differential privacy, secure aggregation, and communication compression further mitigate adversarial risks and resource constraints. The proposed methodology provides defense-in-depth while keeping heavy model updates off-chain, thus ensuring scalability, transparency, and trustworthiness. A comprehensive evaluation plan on representative WBAN tasks and datasets validates the feasibility and security of the approach, positioning this work as a realistic step toward secure, privacy-preserving, and auditable learning in next-generation IoT healthcare systems.

Index Terms

Blockchain, PoR, Federated Learning, WBAN, IoT

I. Blockchain-Enabled Federated Learning for WBAN/IoT: System and Methodology

A. Problem and Objectives

We aim to learn clinical and wellness models (e.g., arrhythmia detection, activity recognition, vitals forecasting) from distributed *wearable body-area network* (WBAN) devices without centralizing raw data. The solution must: (i) respect strict privacy constraints, (ii) operate under heterogeneous and resource-limited phones/sensors, (iii) resist data/model poisoning and Sybil attacks, and (iv) provide an auditable life-cycle for training rounds and incentives. We therefore combine *cross-device Federated Learning (FL)* with a *permissioned blockchain* that provides identity, access control, audit, and incentive accounting, while keeping model payloads off-chain for scalability.

B. Three-Tier Architecture

Tier 1 (WBAN sensors). On/near-body sensors (ECG, PPG, accelerometer, SpO_2, etc.) stream windows/features to a local *phone/fog gateway* via BLE/IEEE 802.15.6. No blockchain logic runs on sensors.
Tier 2 (Phone/Fog gateway = FL client). The smartphone (or a home fog node) caches windows, performs local training/inference, applies privacy filters (gradient clipping and DP noise), compresses updates (sparsification/quantization), and interacts with smart contracts for enrollment, reputation, and payouts.
Tier 3 (Services). An *off-chain* FL *aggregator* coordinates rounds, runs robust aggregation and secure aggregation protocols; a *permissioned EVM chain* maintains participant registry, round logs (hashes/metadata), reputation/staking, and incentives. Only identifiers, hashes, and events are committed on-chain; model parameters remain off-chain.

C. Federated Learning Workflow

Let the global model at round t be $\mathbf{w}^{(t)}$. Client i with local data D_i trains for E epochs to obtain $\mathbf{w}_i^{(t)}$ (or $\Delta\mathbf{w}_i^{(t)} = \mathbf{w}_i^{(t)} - \mathbf{w}^{(t)}$). We allow partial participation each round.

a) Reputation-aware robust aggregation.: Define a per-client reputation $\rho_i^{(t)} \in [0,1]$. Aggregation uses both data size and reputation:

$$\mathbf{w}^{(t+1)} = \mathrm{RobustAgg}\Big(\{\mathbf{w}_i^{(t)}\}_{i\in\mathcal{S}_t},\ \{\alpha_i^{(t)}\}_{i\in\mathcal{S}_t}\Big), \quad \alpha_i^{(t)} \propto |D_i|\,\rho_i^{(t)}, \tag{1}$$

where $\mathrm{RobustAgg}$ is a Byzantine-resilient rule (coordinate-wise median, trimmed-mean, geometric median) applied after per-update ℓ_2-clipping and outlier filtering.

All authors are with the Department of Computer and Information Science, Fordham University, New York, NY, USA.
Project proposal received at 26th September, 2025.

b) Personalization (optional).: To mitigate non-IID data, gateways may adopt FedProx (proximal term), FedBN (local batchnorm), or lightweight head personalization while the backbone remains global.

D. Privacy, Security, and Threat Model

Adversaries. (i) data/model poisoning (incl. backdoors), (ii) Sybil identities, (iii) gradient leakage/model inversion, (iv) network tampering/DoS at the phone/fog link.

Controls. (i) *Secure Aggregation*: pairwise/additive masking so the server only sees masked $\sum_i \Delta \mathbf{w}_i$. (ii) *Differential Privacy*: per-example gradient clipping to C and Gaussian noise $\mathcal{N}(0, \sigma^2 C^2)$ on the gateway, tracking (ε, δ). (iii) *Robust aggregation & consistency filters* (cosine similarity to robust center, IQR-based norm filters). (iv) *Sybil resistance*: permissioned enrollment, minimal stake, and reputation thresholds. (v) *Secure transport*: TLS to aggregator; local storage encrypted on phone; on-chain only hashes/metadata. (vi) *Operational hardening*: fog-side rate-limits, allow-lists, BLE pairing.

E. Blockchain Design (On-Chain Minimalism)

We deploy four EVM contracts: *Registry* (membership, roles, ACLs), *Rounds* (append-only events: round index, model hash, stats hash), *Reputation* (mapping $\text{addr} \to \rho$ in basis points, updated by an oracle after server-side evaluation), and *Treasury* (stake/payouts/slashing). This keeps state small and auditable; all heavy ML math stays off-chain.

F. Contribution Quality and Reputation

Let $A(\cdot)$ be a validation metric (e.g., F1/AUC) on a server-held validation slate $\mathcal{V}$. Client i's round quality is

$$q_i^{(t)} = A\big(\mathbf{w}^{(t)} \oplus \Delta \mathbf{w}_i^{(t)}; \mathcal{V}\big) - A\big(\mathbf{w}^{(t)}; \mathcal{V}\big), \tag{2}$$

combined with (a) *consistency* $c_i^{(t)} = \cos\angle(\Delta \mathbf{w}_i^{(t)}, \widetilde{\Delta \mathbf{w}}_{\text{robust}}^{(t)})$ and (b) *plausibility* via clipped norm within the IQR. We aggregate into a signed score $s_i^{(t)} \in [-1, 1]$ and update reputation via an EMA with sigmoid squashing:

$$\rho_i^{(t+1)} = \text{clip}_{[0,1]}\Big((1-\lambda)\rho_i^{(t)} + \lambda\, \sigma\big(s_i^{(t)}\big)\Big), \quad \sigma(x) = \frac{1}{1+e^{-kx}}. \tag{3}$$

Low ρ reduces sampling probability and weight; very low ρ quarantines a client (temporary exclusion).

a) Incentives & slashing.: Payout $p_i^{(t)} = \tau \cdot \max(0, s_i^{(t)})$ is disbursed by the Treasury after the `Rounds` contract logs finalization for round t. Detectable misbehavior (duplicate/replayed/random updates; extreme outliers beyond robust envelope) triggers a slash $\kappa \cdot \text{stake}_i$ and sharp ρ reduction.

G. Communication & Systems Optimizations

To fit WBAN/phone constraints we use: (i) top-k sparsification with error feedback, (ii) 8-bit quantization for selected layers, (iii) multi-rate participation (train when charging + Wi-Fi), (iv) energy-aware early-exit heads, and (v) client selection based on battery/connectivity/freshness/reputation.

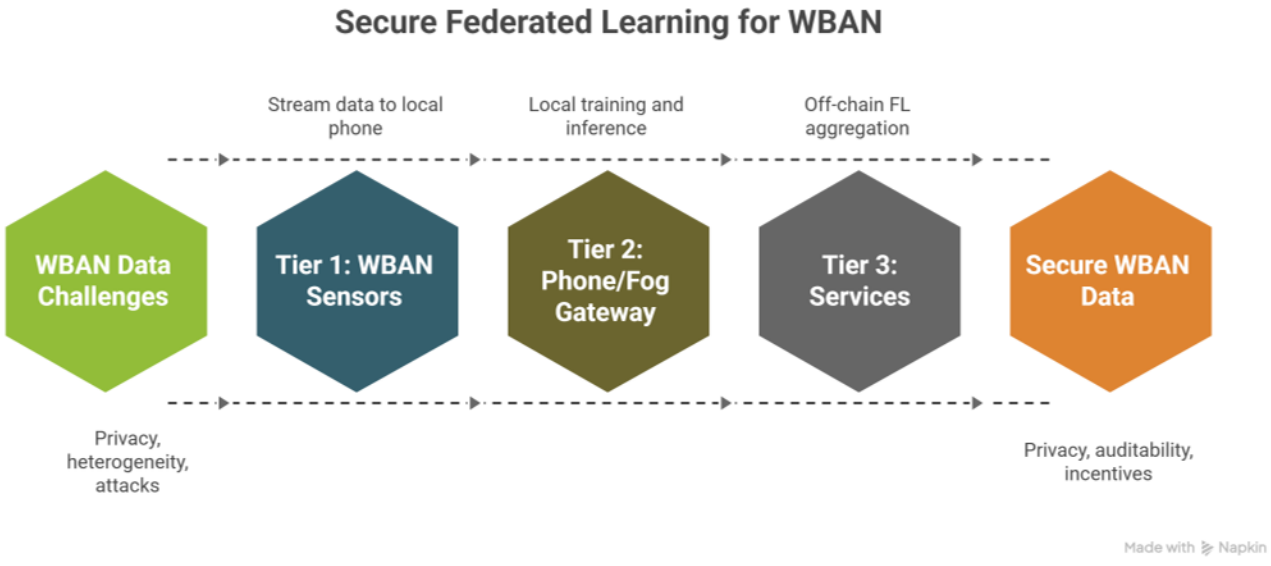

Fig. 1. SecureCare overall framework diagram.

Algorithm 1 Server-Side Reputation-Aware RobustFedAvg

Require: Initial model $\mathbf{w}^{(0)}$, rounds T, client pool $\mathcal{C}$, reputations $\{\rho_i^{(0)}\}$, clip C

for $t = 0$ to $T-1$ **do**
  Sample $\mathcal{S}_t \subset \mathcal{C}$ with $P(i \in \mathcal{S}_t) \propto \rho_i^{(t)}$ and availability
  Broadcast $\mathbf{w}^{(t)}$, round seed; receive masked updates $\{\Delta\mathbf{w}_i^{(t)}\}_{i\in\mathcal{S}_t}$
  Unmask aggregate via SecureAgg; clip each $\Delta\mathbf{w}_i^{(t)} \leftarrow \mathrm{clip}_C(\Delta\mathbf{w}_i^{(t)})$
  Filter outliers by similarity and IQR
  Compute robust center $\widehat{\mathbf{w}}^{(t)} \leftarrow \mathrm{RobustAgg}(\{\mathbf{w}_i^{(t)}\})$
  Set $\alpha_i^{(t)} \propto |D_i|\rho_i^{(t)}$ over inliers; update $\mathbf{w}^{(t+1)} \leftarrow \sum_i \alpha_i^{(t)}\mathbf{w}_i^{(t)}$ **or** use $\widehat{\mathbf{w}}^{(t)}$
  for all $i \in \mathcal{S}_t$ **do**
    Compute $s_i^{(t)}$ using (2) + consistency/plausibility; update $\rho_i^{(t+1)}$
  Emit on-chain round event: $\langle t,\ \mathrm{hash}(\mathbf{w}^{(t+1)}),\ \mathrm{statsHash}\rangle$

Algorithm 2 Client-Side (Phone/Fog Gateway)

Require: Global model $\mathbf{w}^{(t)}$, local data D_i, epochs E, batch size B, DP noise σ

Initialize optimizer (SGD/Adam); set proximal/scaffold terms if enabled
for $e = 1$ to E **do**
  for batches $(x, y) \subset D_i$ **do**
    Update $\mathbf{w} \leftarrow \mathbf{w} - \eta\nabla\ell(\mathbf{w}; x, y)$
Form update $\Delta\mathbf{w}_i^{(t)} \leftarrow \mathbf{w} - \mathbf{w}^{(t)}$; per-example clip to C; add DP noise
Compress (top-k, quantize); mask for SecureAgg; upload
(Optional) keep a personalized head for local inference

H. Algorithms

Algorithms 1 and 2 formalize the end-to-end workflow of our blockchain-enabled federated learning framework with a reputation system for WBAN/IoT devices.

a) Server-side workflow (Algorithm 1).: The server (or aggregation service) is responsible for coordinating training rounds, collecting masked model updates, and performing robust aggregation. A key difference from standard FedAvg is the introduction of *reputation-aware sampling and weighting*. Clients with higher reputation scores are more likely to be selected and receive greater weight in the global update, which naturally incentivizes consistent, high-quality contributions. Robust aggregation functions, such as coordinate-wise median or trimmed mean, are applied to protect against poisoning and Byzantine faults. The algorithm also integrates *secure aggregation*, ensuring that only the aggregate update is visible, thereby preserving client privacy. Finally, the server logs round metadata (model hash, statistics, participants) on the blockchain, providing an immutable and auditable record of training progress.

b) Client-side workflow (Algorithm 2).: On the gateway side, the client receives the current global model and trains it locally using its private WBAN data. Several safeguards are embedded: (i) per-example gradient clipping to bound sensitivity; (ii) Gaussian noise addition for differential privacy guarantees; and (iii) update compression through sparsification or quantization to reduce communication overhead. Before sending updates to the server, clients engage in secure aggregation protocols, masking their contributions. Clients may also retain a lightweight personalized model head for local inference, which is useful in heterogeneous WBAN environments where sensor data distributions vary across users.

c) Blockchain interaction.: While the algorithms themselves focus on FL training, each round is tightly coupled with smart contracts. The server, acting as an oracle, updates on-chain reputation values and disburses incentives. Reputation updates are derived from contribution quality scores (Equation 2) and consistency checks, ensuring that only honest and beneficial contributions are rewarded. Malicious or low-quality clients can be penalized via on-chain slashing, which directly affects their reputation and stake. This integration of algorithmic steps with blockchain primitives distinguishes our framework from prior FL approaches.

d) Novelty and practicality.: Compared to the state-of-the-art, the novelty lies in combining (i) defense-in-depth mechanisms—differential privacy, secure aggregation, and robust aggregation—with (ii) a blockchain-backed reputation and incentive layer tailored for resource-constrained WBAN devices. The proposed algorithms are lightweight enough for execution on smartphones acting as gateways, while maintaining rigorous protections against poisoning and Sybil attacks. Moreover, the auditable and tamper-proof logging of training rounds ensures accountability and compliance in sensitive healthcare deployments.

In summary, these algorithms define a practical, secure, and incentive-compatible foundation for federated learning across WBAN/IoT devices, bridging the gap between cryptographic security, statistical robustness, and blockchain-enabled trust management.

```
// Minimal on-chain interfaces (sketch)
contract FLRegistry {
  enum Role { Patient, Clinician, Operator }
  struct Member { bool active; Role role; uint64 joined; }
  mapping(address=>Member) public members;
  function enroll(address a, Role r) external;
  function deactivate(address a) external;
}
contract FLReputation {
  mapping(address=>uint16) public rhoBps; // 0..10000
  function set(address a, uint16 bps) external; // server-oracle
}
contract FLRounds {
  event RoundFinalized(uint64 t, bytes32 modelHash, bytes32 statsHash, uint32 n);
}
contract FLTreasury {
  mapping(address=>uint256) public stakes;
  function stake() external payable;
  function slash(address a, uint256 amt) external;
  function payout(address a, uint256 amt) external;
}
```

Fig. 2. Smart-contract suite: registry, rounds log, reputation, and treasury.

I. Evaluation Plan

Tasks/Datasets. Smartphone/patch ECG for arrhythmia (e.g., PhysioNet subsets), wearable HAR datasets (e.g., WISDM/UCI-like), and vitals anomaly detection from synthetic WBAN streams respecting device sampling rates.
Baselines. Centralized (upper bound), local-only, FedAvg/FedProx, robust aggregators (median/trimmed-mean/Krum), with-/without reputation, with/without DP and compression.
Metrics. Utility (AUC/F1), poisoning resilience (AUC under $p\%$ malicious clients), privacy (ε, δ) at target utility, systems KPIs (round time, uplink MB/round, phone energy).
Ablations. DP/no-DP; compression on/off; robust rule variants; reputation on/off; permissioned vs public EVM (simulate fee/latency); effect of stake and reputation thresholds.

J. Ethics, Privacy, and Compliance

No raw PHI leaves the phone. Consent and participation are governed by Registry/ACLs; right-to-withdraw is enforced by deactivating keys and ceasing future participation. We track and report DP budgets; on-chain logs provide an immutable audit trail limited to metadata.

K. Risks and Mitigations

On-chain leakage risk is avoided by keeping weights and features off-chain; only hashes/IDs/events are recorded. Latency/fee volatility is handled via permissioned IBFT/PoA or L2 sidechains; events may be batched. Heterogeneous devices are accommodated with adaptive local epochs and energy-aware scheduling.

L. Novelty and Contribution of this Proposal

Our design positions blockchain in its *strength lane* (identity, audit, incentives, minimal on-chain state) and augments cross-device FL with *defense-in-depth*: secure aggregation, DP, robust aggregation, and a principled reputation/stake mechanism tailored for WBAN constraints. The methodology is practical (phone/fog executors), auditable (round events), and resilient to poisoning/Sybil attacks.