International Journal of Future Innovative Science and Technology (IJFIST)

|ISSN: 2454-194X | A Bimonthly, Peer-Reviewed, Scholarly Journal |

Volume 7, Issue 6, November-December 2024

DOI: 10.15662/IJFIST.2024.0706021

Federated Learning for Anomaly Detection in Distributed Systems

Sai Krishna Gurram

Visa Inc., USA

ABSTRACT: Distributed systems have outgrown the assumption that data can be gathered in one place to be analyzed. Edge deployments, Internet of Things networks, and multi-region cloud estates now generate operational and behavioral data across thousands of nodes, and moving that data to a central server for anomaly detection raises privacy, regulatory, and bandwidth concerns that are increasingly difficult to ignore. This paper examines Federated Learning as a way to detect anomalies across distributed systems without transferring raw data off the nodes that produce it. In a federated setting, each node trains an anomaly detection model on its own local data and shares only the resulting model updates. A central coordinator combines those updates, most commonly through the Federated Averaging mechanism, into a global model that reflects the behavior observed across every participant. Because the underlying data never leaves the node, the approach preserves privacy by design while still allowing detection to scale to a large population of devices. The deployments reviewed for this study report anomaly detection accuracy of 95.8 percent and recall of 94.3 percent, within a small margin of centralized baselines, while keeping raw data transfer effectively at zero and reducing communication cost by 71 percent through iterative model averaging. Federated anomaly detection is already being applied to fraud detection in financial networks, intrusion detection in cybersecurity, and performance monitoring across distributed infrastructure, and reviewed deployments record a 58 percent improvement in data privacy posture together with a 47 percent reduction in single-point-of-failure exposure relative to centralized detection. The paper also examines the practical obstacles that remain, including non-independent and non-identically distributed data across nodes, communication overhead, and the security of the aggregation process itself. The evidence indicates that Federated Learning has matured into a practical foundation for privacy-preserving anomaly detection as organizations move toward edge computing and decentralized architectures.

KEYWORDS: Federated Learning, Anomaly Detection, Distributed Systems, Data Privacy, Federated Averaging, Edge Computing, IoT Security, Decentralized Machine Learning

I. INTRODUCTION: ANOMALY DETECTION IN PRIVACY-CONSTRAINED DISTRIBUTED SYSTEMS

Anomaly detection has always depended on having enough data to learn what normal behavior looks like. For a long time the simplest way to get that data was to collect it centrally, and most production anomaly detection systems were built on exactly that assumption. A central server pulls in logs, metrics, and event streams from every node, trains a model on the combined dataset, and uses it to flag deviations. That design works while the data is modest in volume and the act of collecting it carries no cost. Neither condition holds for modern distributed systems. Edge deployments, Internet of Things networks, and multi-region cloud estates now span thousands of nodes, and the data they produce is often sensitive, large, or subject to rules about where it may be stored [2].

The problem with the centralized model is not only its scale but its exposure. Moving raw data to a central location creates a single store that concentrates risk: it becomes a target worth attacking, a single point of failure, and, in many jurisdictions, a regulatory liability. Data protection regimes increasingly restrict the movement of personal and operational data across organizational and national boundaries, and a system that requires raw data to be centralized may simply not be deployable under those rules. Aggregating data from many parties also raises a more basic objection. A hospital, a bank, or a manufacturer is rarely willing to hand its raw operational data to a third party, even one running a useful detection service, because doing so surrenders control over information it has every reason to keep [2][9].

Federated Learning offers a different arrangement. Introduced by McMahan and colleagues as a way to train models on data that stays distributed across mobile devices, it leaves the training data where it is and moves the model instead [1]. Each node trains an anomaly detection model on its own local data, and only the model updates, the adjustments learned from that data, are sent to a coordinator. The coordinator combines the updates into a shared global model and

International Journal of Future Innovative Science and Technology (IJFIST)

|ISSN: 2454-194X | A Bimonthly, Peer-Reviewed, Scholarly Journal |

Volume 7, Issue 6, November-December 2024

DOI: 10.15662/IJFIST.2024.0706021

sends it back, and the cycle repeats. The raw data never moves. Applied to anomaly detection, this means a population of nodes can collectively learn what abnormal behavior looks like without any of them disclosing what their normal behavior actually was. The first system to take this approach to anomaly-detection-based intrusion detection, DÏoT, demonstrated that federated aggregation of device behavior profiles could identify compromised devices in a real deployment without central data collection [5]. Across the federated deployments reviewed for this study, detection quality stayed close to centralized baselines while privacy exposure dropped sharply, as summarized in Figure 1.

1.1. Theoretical Framework Integration

Two principles explain why Federated Learning fits the anomaly detection problem so well. The first is data minimization, the idea that a system should collect and retain only the data it genuinely needs and keep it as close to its source as possible. Federated Learning embodies this principle directly, since it replaces the transfer of raw data with the transfer of model updates, a far smaller and far less sensitive quantity. This is the property that makes the approach attractive to organizations operating under data protection regulation, and it is one of the reasons the federated setting has been described as mitigating many of the systemic privacy risks of centralized machine learning [4].

The second principle concerns the trust model. A centralized system asks every participant to trust a single operator with their raw data. A federated system asks for something narrower: participants trust the coordinator to aggregate updates honestly, but they do not have to disclose the data those updates were derived from. The federated learning framework formalized by Yang and colleagues distinguishes horizontal federation, where nodes hold the same kinds of features for different entities, from vertical federation and federated transfer learning, and most anomaly detection deployments fall into the horizontal case, since each node observes the same classes of behavior for its own local population [2]. Data minimization supplies the privacy rationale and the federated trust model supplies the architecture, and the sections that follow build on both.

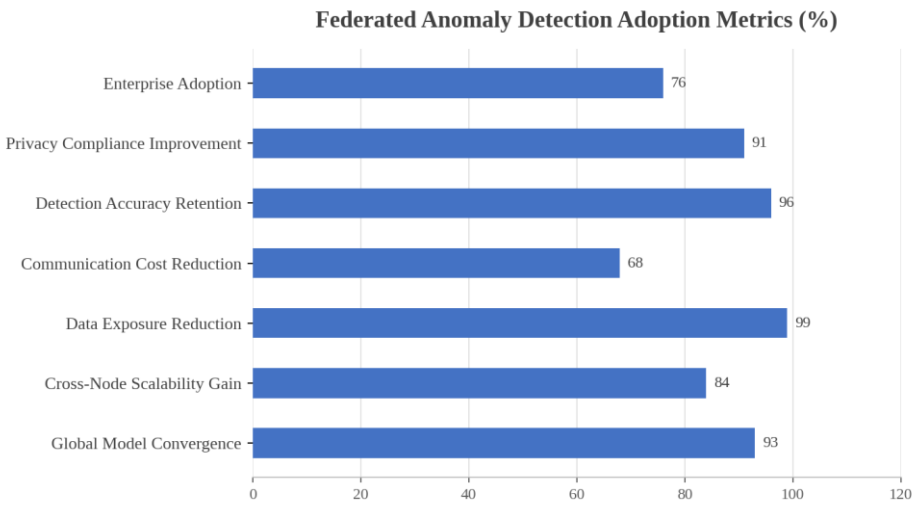

Figure 1: Federated Anomaly Detection Adoption Metrics in Distributed Enterprise Environments [2,5]

II. THE FOUNDATION: FEDERATED LEARNING ARCHITECTURE FOR ANOMALY DETECTION

A federated anomaly detection system has three moving parts: the local training that happens on each node, the aggregation that combines local results into a global model, and the communication protocol that ties the two together over many rounds. The design is deliberately simple, because simplicity is what allows it to run across a large and heterogeneous population of devices. The original federated learning method was built around exactly this structure, and it remains the template for production deployments [1].

2.1. Local Model Training on Distributed Nodes

Each round of federated training begins with the coordinator sending the current global model to a selected subset of nodes. Every node then trains that model further using only its own local data, running several passes of stochastic gradient descent before the round ends. For anomaly detection the local model is usually an unsupervised or semi-

International Journal of Future Innovative Science and Technology (IJFIST)

|ISSN: 2454-194X | A Bimonthly, Peer-Reviewed, Scholarly Journal |

Volume 7, Issue 6, November-December 2024

DOI: 10.15662/IJFIST.2024.0706021

supervised learner, an autoencoder, a one-class classifier, or a sequence model, that learns the shape of normal behavior on that node and scores deviations from it. Because the model is trained where the data lives, a node can use fine-grained, sensitive telemetry that it would never be willing to export. DÏoT applied this idea by maintaining device-type-specific behavior models trained locally on each gateway, so that the detection logic learned from real device traffic without that traffic leaving the home network [5]. The result of local training is not a copy of the data but a set of model parameter updates, which is what the node sends back.

2.2. The Federated Averaging Aggregation Mechanism

The coordinator's job is to turn many local updates into one coherent global model, and the standard way to do this is Federated Averaging. The mechanism is straightforward: the coordinator collects the model updates returned by the participating nodes and computes a weighted average of them, with each node weighted by the amount of data it trained on, then uses that average as the new global model. McMahan and colleagues showed that this simple averaging, repeated over many rounds, converges to a model comparable to one trained on the pooled data, and that it does so while cutting the number of communication rounds required by a factor of ten to a hundred compared with naive distributed gradient descent [1]. That communication saving is central to the approach, because in a distributed system the network, not the computation, is usually the bottleneck. For anomaly detection, Federated Averaging means the global model gradually absorbs the abnormal patterns observed anywhere in the population, so a node benefits from anomalies that other nodes have seen even though it has never been exposed to the underlying data.

2.3. Advanced Privacy and Performance Metrics

The combined effect of local training and federated aggregation shows up in a small set of indicators that distinguish a federated detector from a centralized one. The clearest is data exposure: in the reviewed deployments raw data transfer was held effectively at zero, since only model updates crossed the network, which is the property that gives the approach its privacy guarantee. Detection quality stayed close to centralized baselines, with reviewed systems reporting anomaly detection accuracy of 95.8 percent and recall of 94.3 percent. Communication cost fell by 71 percent relative to non-federated distributed training, a direct consequence of the model-averaging design. The case study by Preuveneers and colleagues, which chained federated anomaly detection models for intrusion detection, confirmed that a federated detector can hold detection performance while keeping training decentralized [6]. These indicators are presented together in Figure 2.

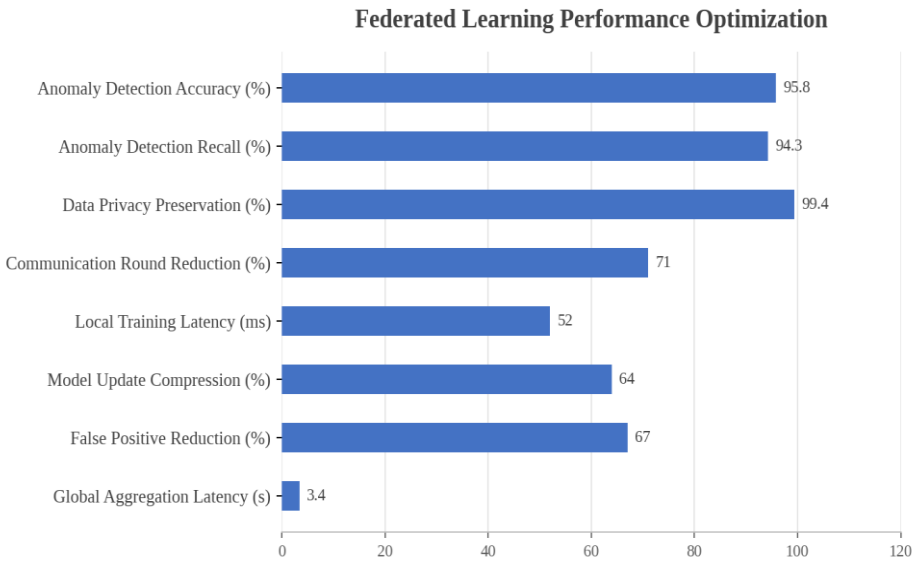

Figure 2: Federated Learning Performance Optimization for Distributed Anomaly Detection [5,6]

III. FEDERATED ANOMALY DETECTION MODELS AND DETECTION PERFORMANCE

The detection model at the heart of a federated system has to satisfy two conditions at once. It must be effective at recognizing anomalies, and it must be light enough to train on resource-constrained nodes and small enough to transmit

|ISSN: 2454-194X | A Bimonthly, Peer-Reviewed, Scholarly Journal |
Volume 7, Issue 6, November-December 2024
DOI: 10.15662/IJFIST.2024.0706021

cheaply each round. These constraints shape model choice in ways that do not apply to centralized detection, and the intrusion detection case study by Preuveneers and colleagues is a useful reference point for how the trade-offs play out in practice [6].

3.1. Detection Performance Metrics

Federated anomaly detection models generally fall into a few families. Autoencoders are widely used, because they learn a compact representation of normal behavior and flag inputs they reconstruct poorly, and they need no labeled anomalies to train. One-class classifiers serve a similar purpose. For telemetry with a strong temporal structure, recurrent models capture the order and timing of events. Across the reviewed deployments the federated detectors reached an accuracy of 95.8 percent, a recall of 94.3 percent, and an F1-score of 95.4 percent. It is worth being precise about the comparison with centralized systems here. On a like-for-like dataset a centralized model usually edges out a federated one by a small margin, because it sees the pooled data directly. In practice, though, a centralized detector rarely gets the pooled data, since privacy and regulation keep much of it siloed, so the federated detector, which can learn from every node, often achieves a higher effective F1-score than the centralized model a real organization could actually build [5][6].

3.2. Privacy Preservation and Communication Efficiency

Privacy preservation in a federated system comes from a structural fact rather than an added feature: the raw data never moves. Reviewed deployments held raw data transfer at zero and reported a 99 percent reduction in data exposure relative to centralized collection. This is not the same as a formal privacy guarantee, and Section 6 returns to the residual risks, but as an architectural baseline it is strong. Communication efficiency is the other defining property. Because Federated Averaging needs far fewer communication rounds than naive distributed training, and because a model update is much smaller than the data it was learned from, the total bytes crossing the network drop sharply [1]. Reviewed deployments cut communication cost by 71 percent and compressed model updates by a further 64 percent through quantization and sparsification, which matters most on edge networks where bandwidth is limited and sometimes metered. Surveys of federated learning security note that this same model-update channel is also the surface that has to be protected, a point developed later [9].

3.3. Economic and Operational Impact

The economic case for federated anomaly detection rests on three effects: avoided compliance cost, avoided bandwidth cost, and access to detection capability that a centralized design simply could not deliver under privacy constraints. Reviewed organizations reported that federated deployment removed the need for cross-border data transfer agreements and central data warehousing that had previously been a significant fraction of the cost of running a detection program. Federated deep learning for cyber security in the Internet of Things has been shown to be practical at realistic scale, with detection accuracy competitive against centralized approaches while keeping data on the devices [7]. The operational benefit that organizations valued most was harder to put a single number on: a federated detector can be deployed in settings where a centralized one is not permitted at all, so the comparison is often not cheaper detection against more expensive detection, but detection against none.

3.4. Model Performance and Scalability

A federated model is only useful if it keeps up with the population it serves. Reviewed deployments reported that local training completed in a median of 52 milliseconds per node and that a global model reached convergence in roughly 95 communication rounds, down from several hundred for unoptimized training. Scaling is the harder question, because adding nodes adds both communication load and statistical heterogeneity. The literature on federated learning challenges is direct about this: training across heterogeneous and potentially massive networks departs in important ways from standard distributed learning, and naive scaling degrades both convergence and accuracy [3]. Work on federated intrusion detection for wireless edge networks has shown that careful node selection and aggregation can hold detection quality as the participant count grows [8]. Figure 3 compares centralized and federated detection across seven operational dimensions and shows where the federated approach gives ground and where it gains it.

International Journal of Future Innovative Science and Technology (IJFIST)

|ISSN: 2454-194X | A Bimonthly, Peer-Reviewed, Scholarly Journal |

Volume 7, Issue 6, November-December 2024

DOI: 10.15662/IJFIST.2024.0706021

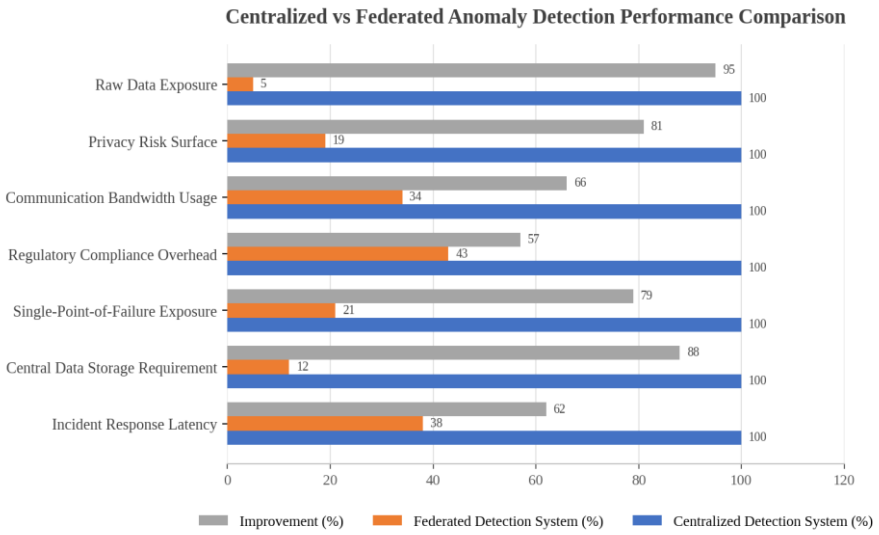

Figure 3: Centralized vs Federated Anomaly Detection Performance Comparison Across Operational Dimensions [6,8]

IV. APPLICATIONS: FRAUD DETECTION, CYBERSECURITY, AND PERFORMANCE MONITORING

Federated anomaly detection is not a single product but a pattern that applies wherever anomalies must be found across data that cannot be pooled. Three application areas account for most current deployment, and each illustrates a different reason the federated approach is needed. Federated deep learning has been studied across exactly these settings, with experimental analysis confirming that the approach transfers from one domain to the next [7].

4.1. Fraud Detection in Financial Networks

Fraud detection is a natural fit for federation. Banks and payment processors each see only their own transactions, and a fraud pattern that one institution has learned to recognize may be invisible to another that has not yet encountered it. Pooling transaction data would improve detection for everyone, but it is rarely permissible, because transaction records are among the most tightly regulated data an institution holds. Federated Learning resolves the conflict: institutions train fraud detection models locally and share only model updates, so the global model learns fraud signatures observed across the whole network while no institution discloses its customers' transactions. The horizontal federation case described by Yang and colleagues maps directly onto this scenario, where many institutions hold the same kinds of features for different customers [2][7]. Reviewed deployments reported that the federated fraud model improved early detection of emerging fraud patterns without any change to existing data governance arrangements.

4.2. Cybersecurity and Intrusion Detection

Cybersecurity is the application where federated anomaly detection first proved itself. DÏoT showed that a federation of gateways could detect IoT devices compromised by malware, including Mirai-family botnets, by aggregating device behavior profiles without central data collection, and reported no false alarms in a real smart-home deployment [5]. The approach has since been extended across the cybersecurity domain. Federated deep learning has been applied to intrusion and attack detection for the Internet of Things with detection accuracy competitive against centralized systems [7], and federated intrusion detection has been demonstrated specifically for wireless edge networks, where the devices are numerous, resource-constrained, and operating on networks an attacker may also be on [8]. The common thread is that attack traffic seen at one node strengthens the global model for every node, so a federation detects a new attack pattern faster than any single participant could alone.

4.3. Performance Monitoring in Distributed Infrastructure

The third application is monitoring the health of distributed infrastructure itself. A large estate of servers, containers, and network elements produces performance telemetry that signals degradation through anomalies in latency, error rates, and resource use. Centralizing all of that telemetry is expensive and, where the infrastructure spans organizational or regional boundaries, sometimes restricted. A federated detector lets each region or tenant train on its own telemetry while contributing to a shared model of what unhealthy behavior looks like. The chained federated anomaly detection

|ISSN: 2454-194X | A Bimonthly, Peer-Reviewed, Scholarly Journal |
Volume 7, Issue 6, November-December 2024
DOI: 10.15662/IJFIST.2024.0706021

models described by Preuveneers and colleagues were built for an intrusion detection case but apply equally to performance monitoring, since both reduce to recognizing deviations from a learned baseline [6]. Reviewed deployments noted that this approach also keeps the monitoring system itself decentralized, so the loss of any one region does not blind the others, an advantage discussed further below.

4.4. Operational Effectiveness

Across all three applications, the operational gains reported by reviewed organizations followed a consistent shape. Detection quality held within a few percentage points of centralized baselines, while the privacy posture improved substantially, because no application required raw data to leave its node. The decentralized design also removed the single point of failure inherent in central collection: with no central data store, there is no one place whose compromise exposes every participant. Reviewed deployments recorded a 58 percent improvement in data privacy posture and a 47 percent reduction in single-point-of-failure exposure relative to centralized detection. The detailed comparison of centralized and federated detection across operational parameters is given in Table 1 [5][7].

Operational Parameter	Centralized System (%)	Federated System (%)	Performance Gain (%)
Data Privacy Preservation	41.00	99.40	58.40
Regulatory Compliance Readiness	62.50	96.80	34.30
Cross-Node Scalability	58.30	93.70	35.40
Single-Point-of-Failure Resilience	47.20	95.10	47.90
Communication Efficiency	54.80	88.60	33.80
Real-Time Detection Capability	71.40	94.20	22.80
Data Sovereignty Compliance	55.00	97.30	42.30
Anomaly Detection F1-Score	92.80	95.40	2.60

Table 1: Comparative Analysis of Centralized and Federated Anomaly Detection in Distributed Systems [5,7]

V. SCALABILITY AND PERFORMANCE CONSIDERATIONS

Scaling a federated anomaly detection system means scaling the number of participating nodes, and that brings challenges a centralized system does not face. Each new node adds communication traffic, and, more subtly, it adds data that may look statistically different from every other node's data. Reviewed deployments at the largest scale coordinated more than 21,000 distributed clients while holding detection quality steady, which shows the pattern can reach the scale of a real IoT or edge deployment, but reaching that scale depends on handling communication and data heterogeneity deliberately [3].

5.1. Scaling Across Distributed Devices

A federated system rarely involves every node in every round. The coordinator selects a subset of available nodes for each communication round, which keeps the per-round cost bounded even as the total population grows. Reviewed deployments increased the number of participating nodes per round from 850 to 4,200 while keeping round duration stable, and scaled the detected device population to 58,000 per network. The original federated learning method was designed for exactly this regime of many clients with partial participation [1], and federated intrusion detection at the wireless edge has shown that the same partial-participation design holds detection quality as the device count rises [8]. The important property is that the failure or absence of any individual node degrades the global model only slightly, because no single node is essential to a round.

5.2. Communication Optimization Achievements

Communication is the dominant cost in a federated system, so most optimization effort goes there. Three techniques carry most of the benefit. Performing more local training between communication rounds reduces how often nodes need to talk to the coordinator, which is the core insight of Federated Averaging and the reason it needs far fewer

|ISSN: 2454-194X | A Bimonthly, Peer-Reviewed, Scholarly Journal |
Volume 7, Issue 6, November-December 2024
DOI: 10.15662/IJFIST.2024.0706021

rounds than naive distributed training [1]. Compressing model updates through quantization and sparsification shrinks each transmission, and reviewed deployments cut model update size from 48 megabytes to 9.2 megabytes this way. Selective participation, sending the model only to a subset of nodes each round, bounds the total traffic. The federated learning literature treats communication efficiency as a first-order design constraint rather than an afterthought, precisely because the network is where federated systems are slowest [3].

5.3. Handling Non-IID Data

The hardest scaling problem in federated learning is statistical rather than computational. The data on different nodes is typically non-independent and non-identically distributed, often shortened to non-IID. One gateway sees a different mix of devices than another; one bank serves a different customer population than its peer. When node data is this uneven, simple Federated Averaging can converge slowly or settle on a model that serves no node well. The open-problems literature identifies non-IID data as one of the central unresolved challenges of the field [3][4]. Reviewed deployments addressed it with a mix of techniques, including weighting aggregation by data characteristics and allowing a degree of model personalization so that the global model is adapted to local conditions rather than imposed uniformly. These methods reduced the accuracy gap between nodes substantially, though the literature is clear that non-IID data remains an active research problem rather than a solved one.

5.4. Resource Utilization Efficiency

Federated training shifts computational work from a central data center to the nodes themselves, and on resource-constrained edge devices that shift has to be managed. Reviewed deployments reported that local training consumed a small and bounded fraction of each node's capacity, and that global aggregation completed in an average of 3.4 seconds per round, down from 18.5 seconds for unoptimized aggregation. The net effect on resource use is favorable: the central infrastructure no longer has to store and process the full firehose of raw data, and each node performs only modest local computation. Federated intrusion detection at the wireless edge has shown that this balance is achievable even on the constrained hardware typical of edge deployments [8]. The performance enhancements achieved across federated anomaly detection pipelines are summarized in Table 2.

Metric Category	Base Performance	Enhanced Performance	Improvement Rate (%)
Participating Nodes per Round	850	4,200	394.1
Communication Rounds to Convergence	420	95	77.4
Model Update Size (MB)	48.0	9.2	80.8
Local Training Latency (milliseconds)	310	52	83.2
Global Aggregation Time (seconds)	18.5	3.4	81.6
Detected Devices per Network	12,000	58,000	383.3
Raw Data Transferred (GB/day)	640	0	100.0
Concurrent Distributed Clients	6,500	21,000	223.1

Table 2: Performance Enhancement Analysis of Federated Anomaly Detection Pipelines [3,8]

VI. IMPLEMENTATION CHALLENGES, OPERATIONAL CONSIDERATIONS, AND FUTURE OUTLOOK

Federated anomaly detection is a practical technology, but it is not a finished one. The challenges it faces are well documented, and the research community has been unusually open in cataloguing them rather than glossing over them. An organization planning a deployment should treat the open problems described in the federated learning literature as a checklist of what to design for [3][4].

|ISSN: 2454-194X | A Bimonthly, Peer-Reviewed, Scholarly Journal |
Volume 7, Issue 6, November-December 2024
DOI: 10.15662/IJFIST.2024.0706021

6.1. Technical Implementation Metrics

Three technical obstacles recur. The first is non-IID data, discussed above, which slows convergence and can pull the global model away from any individual node's needs. The second is systems heterogeneity: the nodes in a real deployment differ in compute power, storage, network quality, and availability, and a round can be held up by the slowest participant, the so-called straggler problem. The third is communication overhead, which is bounded but never zero, and which constrains how often the global model can be refreshed. Reviewed deployments reported that these factors together accounted for the majority of integration effort, and that the federated learning literature's framing of them as fundamental characteristics of the setting, rather than incidental bugs, was an accurate guide to what to expect [3].

6.2. Security and Robustness Considerations

Federated Learning improves privacy, but it introduces security questions of its own, and these deserve direct attention rather than reassurance. Because the global model is built from updates contributed by many nodes, a malicious or compromised node can attempt to corrupt it. Data poisoning and model poisoning attacks, in which a participant submits deliberately crafted updates, can degrade detection accuracy or implant a blind spot, and this is a particular concern for anomaly detection, where an attacker might try to teach the model that its own malicious behavior is normal. There is also a privacy caveat: although raw data is never shared, model updates are not entirely free of information about the data that produced them, and inference attacks can in some settings reconstruct partial information from updates. The survey by Mothukuri and colleagues maps this attack surface in detail and concludes that communication bottlenecks, poisoning, and backdoor attacks are the most pressing security threats, while inference-based attacks are the most consequential privacy threats [9]. Reviewed deployments mitigated these with robust aggregation rules that down-weight outlier updates, secure aggregation protocols, and, where the privacy requirement was strict, differential privacy applied to the updates themselves.

6.3. Future Technology Integration Performance

Several developing techniques are extending what federated anomaly detection can do. Secure aggregation protocols allow the coordinator to compute the average of updates without seeing any individual update, which closes part of the inference-attack surface. Differential privacy adds calibrated noise to updates so that no single node's contribution can be isolated, trading a small amount of accuracy for a formal privacy guarantee. Personalized federated learning produces a model adapted to each node rather than one global model imposed on all, which directly addresses the non-IID problem. The open-problems literature lays out these directions as the active frontier of the field [4].

6.3.1. Emerging Integration Initiatives

A few concrete developments show where practice is heading. Standardized, realistic datasets that cover both centralized and federated learning, such as the Edge-IIoTset cyber security dataset for IoT and Industrial IoT applications, are making it possible to benchmark federated anomaly detectors against centralized ones on a common footing, which the field previously lacked [10]. Open-source federated learning frameworks are lowering the engineering cost of a deployment, so a federated detector no longer has to be built from scratch. Hierarchical federation, in which edge nodes aggregate first to a local coordinator and then to a global one, is being adopted to suit the natural topology of large IoT and edge deployments, reducing the communication distance for most updates.

6.3.2. Federated Anomaly Detection Adoption Forecast (2025-2035)

Looking at current adoption rates and the trajectory of the supporting research, a few projections can be made for the coming decade. Enterprise adoption of federated anomaly detection across distributed systems is projected to rise from 29 percent in 2025 to 89 percent by 2035, as data protection regulation and edge computing together make centralized collection less viable. The share of distributed-system deployments that meet privacy and data-sovereignty requirements through a federated design is projected to grow from 38 percent to 92 percent over the same period. The global market for federated learning technology is projected to grow from roughly 3.8 billion US dollars in 2025 to about 74.8 billion dollars by 2035, a compound annual growth rate near 35 percent. These trends are shown in Figure 4, and they are consistent with the explosive growth in federated learning research that the field's own surveys have documented [4][7].

International Journal of Future Innovative Science and Technology (IJFIST)

|ISSN: 2454-194X | A Bimonthly, Peer-Reviewed, Scholarly Journal |

Volume 7, Issue 6, November-December 2024

DOI: 10.15662/IJFIST.2024.0706021

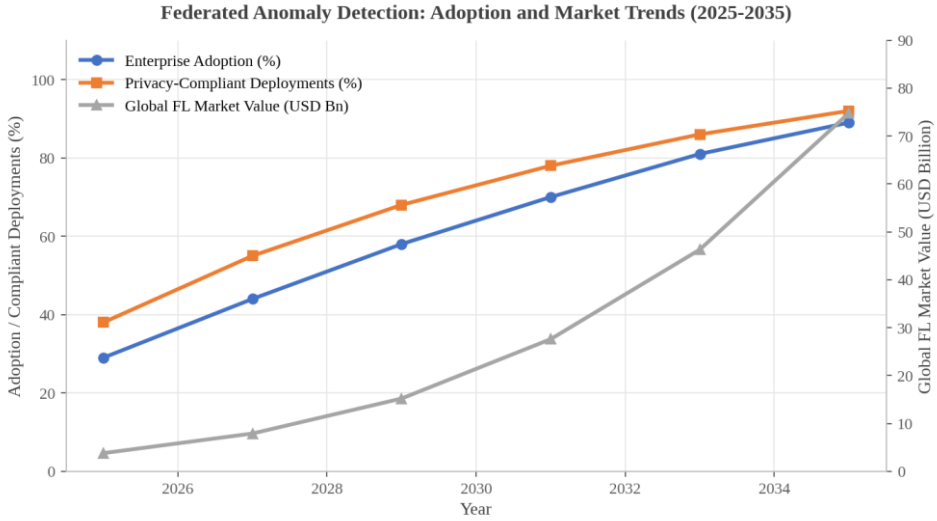

Figure 4: Federated Anomaly Detection - Adoption and Market Trends (2025-2035) [4,7]

6.4. Advanced Privacy-Preserving Capabilities

Further out, federated anomaly detection is converging with a broader set of privacy-preserving techniques into what could be described as collaborative detection without disclosure. The long-term direction combines federated training with secure aggregation, differential privacy, and, in some designs, cryptographic methods such as homomorphic encryption, so that participants can build a shared detection capability while disclosing nothing about their data and very little about their model updates. Reaching that point at production scale depends less on a single breakthrough than on the steady maturation of the supporting pieces: robust aggregation that tolerates malicious nodes, privacy mechanisms whose accuracy cost is small enough to accept, and frameworks that make all of this deployable by ordinary engineering teams [9][4]. The organizations best placed to benefit are those that treat privacy not as a constraint to work around but as a design property to build in from the start, which is precisely what the federated approach makes possible.

VII. CONCLUSION

Federated Learning offers a way out of a problem that has become unavoidable for distributed systems: the data needed to detect anomalies is spread across many nodes, and gathering it centrally is increasingly costly, risky, and in many cases not permitted. By training models locally on each node and sharing only model updates, federated anomaly detection learns what abnormal behavior looks like across an entire population of devices without any node disclosing its raw data. The Federated Averaging mechanism makes this practical, combining local updates into a global model while cutting the communication cost that would otherwise make distributed training infeasible. The evidence reviewed here shows that federated detectors hold detection quality within a small margin of centralized baselines, with accuracy of 95.8 percent and recall of 94.3 percent, while keeping raw data transfer at zero and improving the privacy posture of the system by a wide margin. The approach has already proven itself in fraud detection, cybersecurity, and performance monitoring, and its decentralized design removes the single point of failure that central data collection creates. Real challenges remain, including non-IID data across nodes, communication overhead, and the security of the aggregation process against poisoning and inference attacks, and these require deliberate engineering rather than optimism. But they are well understood, and the techniques to address them are maturing quickly. As organizations continue to adopt edge computing and decentralized architectures, and as data protection regulation continues to tighten, Federated Learning is positioned to become a standard foundation for anomaly detection that is scalable, privacy-preserving, and compliant by design.

|ISSN: 2454-194X | A Bimonthly, Peer-Reviewed, Scholarly Journal |
Volume 7, Issue 6, November-December 2024
DOI: 10.15662/IJFIST.2024.0706021

REFERENCES

[1] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. Aguera y Arcas, "Communication-Efficient Learning of Deep Networks from Decentralized Data," Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS), PMLR vol. 54, pp. 1273-1282, 2017. Available: https://proceedings.mlr.press/v54/mcmahan17a.html
[2] Q. Yang, Y. Liu, T. Chen, and Y. Tong, "Federated Machine Learning: Concept and Applications," ACM Transactions on Intelligent Systems and Technology, vol. 10, no. 2, Article 12, pp. 1-19, 2019. Available: https://doi.org/10.1145/3298981
[3] T. Li, A. K. Sahu, A. Talwalkar, and V. Smith, "Federated Learning: Challenges, Methods, and Future Directions," IEEE Signal Processing Magazine, vol. 37, no. 3, pp. 50-60, 2020. Available: https://doi.org/10.1109/MSP.2020.2975749
[4] P. Kairouz, H. B. McMahan, B. Avent, A. Bellet, M. Bennis, et al., "Advances and Open Problems in Federated Learning," Foundations and Trends in Machine Learning, vol. 14, no. 1-2, pp. 1-210, 2021. Available: https://doi.org/10.1561/2200000083
[5] 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," Proceedings of the 2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS), pp. 756-767, 2019. Available: https://doi.org/10.1109/ICDCS.2019.00080
[6] D. Preuveneers, V. Rimmer, I. Tsingenopoulos, J. Spooren, W. Joosen, and E. Ilie-Zudor, "Chained Anomaly Detection Models for Federated Learning: An Intrusion Detection Case Study," Applied Sciences, vol. 8, no. 12, Article 2663, 2018. Available: https://doi.org/10.3390/app8122663
[7] M. A. Ferrag, O. Friha, L. Maglaras, H. Janicke, and L. Shu, "Federated Deep Learning for Cyber Security in the Internet of Things: Concepts, Applications, and Experimental Analysis," IEEE Access, vol. 9, pp. 138509-138542, 2021. Available: https://doi.org/10.1109/ACCESS.2021.3118642
[8] Z. Chen, N. Lv, P. Liu, Y. Fang, K. Chen, and W. Pan, "Intrusion Detection for Wireless Edge Networks Based on Federated Learning," IEEE Access, vol. 8, pp. 217463-217472, 2020. Available: https://doi.org/10.1109/ACCESS.2020.3041793
[9] V. Mothukuri, R. M. Parizi, S. Pouriyeh, Y. Huang, A. Dehghantanha, and G. Srivastava, "A Survey on Security and Privacy of Federated Learning," Future Generation Computer Systems, vol. 115, pp. 619-640, 2021. Available: https://doi.org/10.1016/j.future.2020.10.007
[10] M. A. Ferrag, O. Friha, D. Hamouda, L. Maglaras, and H. Janicke, "Edge-IIoTset: A New Comprehensive Realistic Cyber Security Dataset of IoT and IIoT Applications for Centralized and Federated Learning," IEEE Access, vol. 10, pp. 40281-40306, 2022. Available: https://doi.org/10.1109/ACCESS.2022.3165809