FoMR
Frontiers of Multidisciplinary Research

Vol. 1 | No. 2 | Apr - Jun 2026

Federated Learning and Blockchain for Smart Home Security: A Comprehensive Review

Fahad Alsaleh[a], Ismail Keshta[b]*

ABSTRACT

Smart homes utilize Internet of Things (IoT) products to increase convenience and automation, but this also makes them vulnerable to various cybersecurity risks. Traditional centralized Intrusion Detection Systems (IDSs) have limitations, such as single points of failure and issues with privacy, scalability, and accuracy. In this review, we discuss the potential of federated learning and blockchain for intrusion detection in smart homes, addressing the limitations of traditional IDSs. A Systematic Literature Review (SLR) was conducted to analyze the use of the blockchain framework, consensus algorithms, and methods for enhancing the framework, such as implementing lightweight frameworks, edge computing, model compression, and intelligent contract optimization. The findings of the literature review suggest that the proposed framework improves the system's resistance to attacks, the model's accuracy (>90%), and its integrity through the blockchain's tamper-proof mechanism. However, it also presents challenges in scalability and other areas, so the focus should be on using this framework in the real world to meet the needs of the intrusion detection mechanism in a smart home.

Received date: 07.02.2026
Accepted date: 20.03.2026

Keywords: Federated Learning; Blockchain; Security; Intrusion Detection Systems (IDS); Internet of Things (IoT); Privacy; Decentralization; Edge Computing; Authentication; Interoperability.

https://doi.org/10.65601/FoMR.2026.1.2.3

[a]*College of Applied Sciences,*
Department of Computer Science and Information Systems,
AlMaarefa University,
Riyadh, Saudi Arabia

***Corresponding Author:**

[b]*College of Applied Sciences,*
Department of Computer Science and Information Systems,
AlMaarefa University,
Riyadh, Saudi Arabia
Email: imohamed@um.edu.sa

All the articles published in FoMR are open-access, providing free access to everyone. FoMR articles are licensed under the Creative Commons Attribution licence (https://creativecommons.org/share-your-work/cclicenses/). This license enables reusers to distribute, remix, adapt, and build upon the material in any medium or format, so long as attribution is given to the creator. The license allows for commercial use.

1. INTRODUCTION

Smart homes are based on IoT devices, such as cameras, door locks, sensors, and appliances, that coordinate together to offer convenience, automation, and remote control to the home occupants. However, this proliferation of IoT devices increased the attack surface of a home network, exposing the inhabitants to various new forms of cyber threats. The most commonly exploited vulnerabilities involve weak default passwords, unpatched firmware, or insecure wireless communications; once exploited, they enable attackers to gain unauthorized access or eavesdrop on private data. Given the typically minimal computing resources and simple architectures of most IoT devices, adequate security measures can hardly be implemented in each device. Thus, smart homes can expect a wide range of security threats, from privacy invasions, such as hacking into a baby monitor's camera, to malicious control, such as turning off the alarms.

An ID generally relies on a centralized architecture: the data collected by all devices is sent to a central server or cloud service for analysis of malicious activity (Thakur, 2025). This has several disadvantages in the smart home context. First, there is the problem of a single point of failure: once the central IDS server is compromised or goes down, the entire home is left unprotected (Rahman et al., 2020). Second, streaming sensitive data continuously, such as camera feeds or network traffic logs, to a single central entity raises serious privacy concerns for smart home dwellers (Al-Turjman & Lemayian, 2020). Third, this model of central processing is characterized by high latency and high bandwidth usage because resource-constrained devices must send large amounts of data for analysis, which is not feasible for real-time detection in a busy IoT network (Singh et al., 2019).

Furthermore, scalability issues arise with centralized IDS as the number of devices increases in modern homes, leading to performance bottlenecks (Thakur, 2025). All of these limitations have created an interest in more distributed and collaborative intrusion detection methods that would further enhance privacy and reduce dependence on any single node. An overview of the FL–blockchain-based IDS architecture, where IoT devices collaboratively train models via federated learning and blockchain ensures secure aggregation, authentication, and immutable logging as depicted in Figure 1.

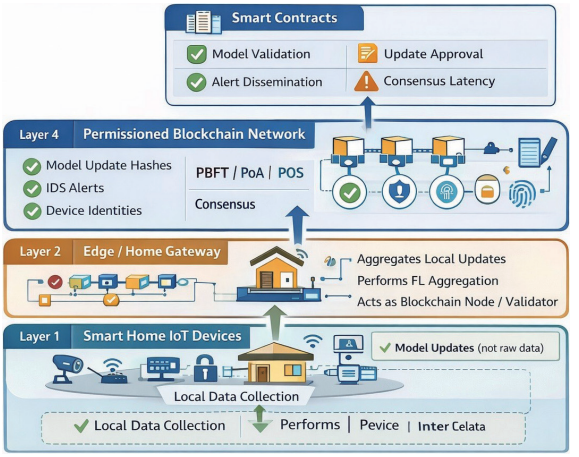

Figure 1 An overview of the FL–blockchain-based IDS architecture, where IoT devices collaboratively train models via federated learning and blockchain ensures secure aggregation, authentication, and immutable logging.

Federated learning has recently emerged as a promising technique to handle the privacy and single-point-of-failure problems of centralized IDSs. In an FL-based approach, each smart home device - a local hub - trains a machine learning intrusion detection model on its own data and only shares model updates, such as gradients or parameters of the model, with a coordinating server or a peer network, instead of the raw data itself (Rahman et al., 2020). By keeping sensitive data, such as network logs or user behavior patterns, on the device and transmitting only abstract model information, FL greatly enhances privacy in the learning process (Liu et al., 2021). It is this collaborative learning that enables a global IDS model to be built from the knowledge of many devices or homes without centralizing data, thereby mitigating the risk of a single point of failure (Thakur, 2025). The other advantage is that FL leverages distributed computation resources - each device does a piece of the work, potentially reducing the burden on any one server (Patel et al., 2022). In a nutshell, federated learning enables intrusion detection systems to leverage diverse attack patterns across homes while preserving user privacy and system resilience (Rahman et al., 2020; Liu et al., 2021). Figure 2 contrasts traditional centralized intrusion detection systems with the proposed federated learning and blockchain-based approach, highlighting how FL improves privacy and decentralization while blockchain ensures trust, integrity, and tamper resistance.

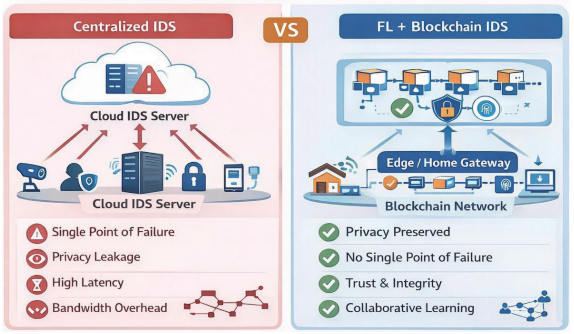

Figure 2 This figure contrasts traditional centralized intrusion detection systems with the proposed federated learning and blockchain-based approach, highlighting how FL improves privacy and decentralization while blockchain ensures trust, integrity, and tamper resistance.

While FL enables decentralized training, it introduces new trust challenges: how can we ensure the model updates contributed by devices are authentic, un-tampered with, and not poisoned by an adversary? In traditional FL, a central aggregator is trusted to collect and apply updates, but this reintroduces a central dependency. In this respect, blockchain technology has been proposed as a means to secure the federated learning process by providing a decentralized, tamper-resistant mechanism for managing model updates. By recording each model update or transaction on a distributed ledger, blockchain ensures an immutable history of contributions that participants can verify. Consensus mechanisms, as seen in Section 3.2 below, validate each update, meaning a malicious or inconsistent update can be detected and rejected by the majority of the network rather than unquestioningly trusted. Also, cryptographic signatures enable devices to authenticate their updates, and the blockchain verifies the contributor's identity. See Section 3.3, preventing any fake or unauthorized nodes from injecting spurious data. All in all, blockchain adds a layer of trust and integrity to FL-based intrusion detection, ensuring that model parameters shared among smart home devices are authentic, agreed upon by consensus, and indelibly logged for audit.

This review aims to analyze the state of the art in combining federated learning with blockchain to secure smart home intrusion detection systems. The scope of the review spans from foundational concepts in smart home IoT security (Section 2) to the design of blockchain frameworks and consensus mechanisms tailored for FL-based IDS (Section 3), and to recent efforts aimed at enhancing these systems for practical deployment (Section 4). We address two research questions:

- RQ1: What blockchain frameworks and consensus algorithms have been used to support federated intrusion detection in smart homes, and how do they ensure security?
- RQ2: What techniques and hybrid models have been proposed to improve the efficiency and effectiveness of FL + blockchain-based IDS in resource-constrained IoT environments?

To answer these, we organize the paper as follows: Section 2 provides background on smart home architectures, intrusion detection systems, federated learning, and blockchain fundamentals to ground the discussion. Section 3 (RQ1) reviews the types of blockchains - public/private - and consensus mechanisms - PoS, PBFT, etc. - utilized in these solutions, including features like device authentication and immutable logging, and compares findings from different studies. Section 4 (RQ2) examines enhancements such as lightweight blockchains, edge computing integration, knowledge distillation for smaller models, and smart contract optimizations, and highlights the performance improvements reported in the literature. We then discuss overarching insights in Section 5, including strengths and weaknesses of various approaches and trade-offs between security and overhead. Section 6 identifies research gaps and future directions, such as datasets required, scalability, privacy techniques, and interoperability. Finally, Section 7 summarizes the findings and implications for future smart home security research.

2. CONTEXT

2.1 Architecture of Smart Home and IoT Constraints

A standard smart home system consists of several IoT devices interconnected by a local network, orchestrated by a central home gateway or router, which may also be connected to cloud services. The devices span a wide range of categories: sensors (beyond simple motion detectors and smoke alarms), actuators (smart locks, lights, and thermostats), cameras, voice assistants, and many others, each usually with very low computing power and memory. All these IoT devices operate under strict, debilitating constraints:

most have limited processing power and capacity, use batteries or very low-power mechanisms, and therefore support only simpler security algorithms. The radio bandwidth may also be severely limited, as in low-power wireless protocols, since large data transfers or frequent communications are undesirable. Another characteristic of the smart home architecture is heterogeneity. Devices come from different manufacturers; each uses different protocols, such as Wi-Fi, Zigbee, and Bluetooth, making the implementation of unified security controls difficult. Because of these constraints, smart home security technologies should be lightweight and efficient, and balance protection needs with limited device resources. This has motivated research toward decentralized, in-network security solutions that avoid reliance on cloud processing and can scale across a diverse ecosystem within a home IoT network.

2.2 Intrusion Detection Systems for Smart Homes

Intrusion detection systems monitor network traffic or device behavior for actions that could indicate malicious activity or a policy violation. In smart homes, IDS can be implemented as network-based monitoring, with data flowing through the home gateway or router, or as host-based, running on individual devices to monitor their operation. Network-based IDSs are common in IoT contexts because resource-limited devices might not handle complete IDS duties; instead, a slightly more powerful gateway can analyze aggregated traffic for anomalies or known attack signatures. Techniques for smart home IDS often leverage machine learning due to the dynamic and novel nature of IoT threats, rather than relying solely on known signatures. Anomaly detection models learn standard device patterns and flag deviations. As an example, IDS might learn the typical network packet rates and destinations of a smart thermostat and trigger an alert if it suddenly starts sending large data bursts to an unknown server. However, one of the significant challenges here is that there is too much limitation in observability and diversity of data in a single home: the data generated by a single home may be too little to train a model on specific attacks, and not all host-based data (such as logs of individual devices) may be available for a central IDS due to technical or privacy restrictions. Moreover, strong security on each device is seldom feasible; thus, IDS must operate under the assumption that some devices might be compromised and still detect malicious actions. In summary, IDS for smart homes needs to be adaptive, lightweight, and privacy-aware, this motivates the design of systems that share detection intelligence across devices or homes without violating privacy. This need directly leads to federated learning (Idrissi et al., 2023).

2.3 Federated Learning for Smart Home Security

An overview of federated learning (FL) allows multiple smart home devices or environments to collaboratively train a global intrusion detection model without directly sharing raw data or with a cloud server. In a typical FL cycle, each device - a local aggregator in the home - trains a local IDS model on locally observed data. For example, its own network traffic or event logs are used for training. After that, a coordinating server (which can be either cloud-based or an edge device) collects only the updated model parameters from each participant, aggregates them (usually by averaging or some similar way) into a new global model, and sends this global model back to the devices for the next training round. This means knowledge of attack patterns detected is shared and learned collectively, but the raw traffic data or personal information will never leave the device, thus preserving privacy. In smart home security, FL enables different households to contribute to a standard intrusion-detection model that can identify malware or abnormal behavior observed in homes without any homeowner seeing the private data of others. Despite all these advantages, FL in IoT also faces challenges. The data across devices is often non-IID (Independent and Identically Distributed): the usage patterns of a particular device can vary widely across homes, which can affect model training convergence and accuracy. Devices also have unequal capabilities to compute updates for a large neural network IDS, or they may even drop out due to power/network issues. These factors contribute to additional communication overhead; periodically, the large model updates must be sent, thereby consuming bandwidth and energy (Habibullah et al., 2024).

Moreover, when FL is integrated with permissioned blockchain systems, a subtle but important trade-off emerges: the reliance on pre-configured trust models. In such setups, a fixed set of pre-approved nodes (e.g., home gateways or edge

servers) are entrusted with validating transactions and aggregating model updates. While this improves efficiency and reduces latency, it also reintroduces elements of centralization, potentially undermining the full decentralization promise of blockchain. If a critical mass of these trusted nodes is compromised or behaves maliciously, the integrity of the entire intrusion detection system could be at risk. This highlights the need for careful threat modeling and for exploring hybrid or dynamically rotating trust mechanisms that balance efficiency with true decentralization in smart home environments.

2.4 Fundamentals of Blockchain Ledger, Smart Contracts, Consensus

Blockchain is a distributed ledger technology popularized by cryptocurrencies but now applied across various domains that require decentralized trust. In essence, a blockchain is an append-only ledger of transactions or records, grouped in blocks, where each block is cryptographically linked to the previous one, forming an immutable chain (Patel et al., 2022). The security of the blockchain comes from this chaining (tampering with a past block would invalidate all subsequent blocks' links) and from the use of distributed consensus – multiple nodes must agree on the ledger's content (Liu et al., 2021). There are different types of blockchains: public blockchains are open networks (anyone can join and validate transactions), as seen in Bitcoin or Ethereum, whereas private/permissioned blockchains restrict participation to known entities (e.g., devices or users authorized in a smart home network) (Thakur, 2025). A consensus mechanism governs how nodes agree on new blocks. Early blockchains use Proof of Work (PoW), which requires solving computational puzzles to propose a block – very secure but computationally expensive and not suitable for IoT environments due to energy and latency costs (Al-Turjman & Lemayian, 2020).

Newer or permissioned systems use alternatives: Proof of Stake (PoS), where the ability to create blocks is based on holding a stake (cryptocurrency or tokens) in the system, Practical Byzantine Fault Tolerance (PBFT) and its variants, which assume a set of authorized nodes that collectively validate blocks through voting, and Proof of Authority (PoA), where a small set of trusted validators rotate to produce blocks (Liu et al., 2021). We delve deeper into these in Section 3.2. Blockchains can also execute smart contracts, which are self-executing code stored on the blockchain that run when certain conditions are met. Smart contracts enable automation of complex logic on the ledger; for example, a smart contract in a smart home security blockchain might automatically issue a network-wide alert or quarantine command if an intrusion is confirmed by multiple devices (Thakur, 2025). Importantly, because smart contracts and transactions are recorded on the immutable ledger, they provide transparency and traceability: the participating parties can verify which rules were executed and when. In essence, blockchain provides a decentralized, tamper-resistant infrastructure that offers inherent trust through consensus and programmability via smart contracts, which can be leveraged to support and secure collaborative intrusion detection schemes.

3. METHODS (RQ1)

The end-to-end smart home security proposed in the research work is presented diagrammatically in Figure 3. This figure illustrates each section comprehensively to facilitate understanding of the work presented in this article.

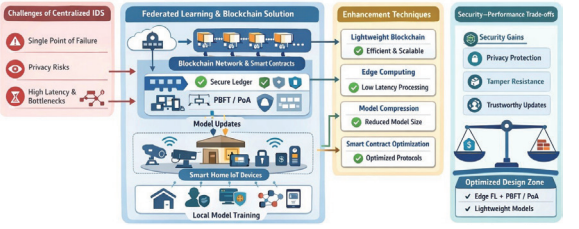

Figure 3 End-to-end smart home security.

3.1 Blockchain Types Utilized (Public, Private, and Consortium)

Regarding smart home IDS with FL integration, nearly all existing solutions opt for private or consortium blockchains rather than public ones (Patel et al., 2022; Liu et al., 2021). Public blockchains such as Bitcoin or Ethereum's leading network enable just anyone to participate and add blocks, which is accompanied by high computation requirements (mining process), as well as irregular transaction processing capabilities and latency, which make such blockchains inappropriate for the real-time requirements of IoT intrusion detection systems (Al-Turjman & Lemayian, 2020). In contrast, a private blockchain is shared exclusively within an organization or a group of pre-approved nodes (in the smart home setting, this would be the devices within one household or households served by a

provider). This enables simpler consensus protocols among permissioned nodes and supports faster transaction processing. In a consortium blockchain setup, multiple trusted nodes work together, with a hypothetical consortium comprising an ISP, a security provider, and home gateways, collectively maintaining a shared ledger (Thakur, 2025).

These permissioned blockchains peacefully coexist, balancing robust decentralization and tamper-evidence without the high overheads associated with the aforementioned public blockchains. Regarding existing solutions, the authors opt for either using a private blockchain network operational within a home or within a network of household devices connected via home hubs, or establishing a consortium network operational on edge/fog infrastructure spanning multiple households, supported by IoT devices (Liu et al., 2021). Additionally, the selective nature of permissioned blockchains can facilitate greater privacy protection, as only authorized nodes can access the shared ledger containing hashed versions of intrusion alerts or models. In conclusion, private or consortium blockchains remain the dominant choice in current FL-based smart home IDS solutions because they adequately address current requirements, while accounting for the potential overheads expected in more conventional IoT applications (Patel et al., 2022; Thakur, 2025).

3.2 Consensus Algorithms: PoS, PBFT, PoA

A proper selection of the consensus algorithm is necessary when integrating blockchain technology into an IDS for a smart home application. The consensus algorithms that are widely used in the literature of the IoT and blockchain technology are:

• Proof of Work (PoW): Security is highly assured with PoW, a system of security that requires miners to solve mathematical puzzles to allow new blocks to be added, which is expensive and thus prevents attacks. However, it is energy- and latency-intensive, thus inefficient and unfit for IoT technology. In theory, it is inefficient and unsuitable for smart homes because of its energy and time complexities (Al-Turjman & Lemayian, 2020). Currently, no FL-IDS systems adopted in smart homes are solely PoW, due to their inefficiency and ability to jam IoT devices.

• Proof of Stake (PoS): PoS chooses the validator of blocks depending upon the stakes or the token ownership of the devices present in the network. It does not require the complex PoW calculations and is therefore more efficient and faster. Theoretically, the PoS blockchain could be used for smart homes if the devices/stakeholders hold the tokens' stakes (Liu et al., 2021). The advantages of PoS include lower power consumption and rapid confirmation of blocks. In contrast, the disadvantages include reliance on the token economy and the concentration of power if controlled by a single entity.

• Practical Byzantine Fault Tolerance (PBFT): This voting-based consensus algorithm is appropriate for permissioned blockchains. All validator nodes, or a fixed set of exchange messages, must reach agreement on the next block, even in the presence of up to f malicious nodes in a 3f+1 network. Although its variants exist—namely, Istanbul BFT and Hyperledger Fabric's Kafka/RAFT ordering—PBFT achieves rapid finality and low latency in tens-of-node networks, rendering it well-accepted in the context of consortium blockchains (Patel et al., 2022). However, its scalability is limited by reduced performance as the number of validator nodes increases, due to increased communication traffic. Indeed, for Smart Home IDS systems that might employ a narrow blockchain, possibly including a limited number of devices or gateway nodes, the scalability issue is efficiently addressed by the rapid consistency provided by PBFT algorithms (Thakur, 2025). This algorithm, or its variants, is notably used to construct prototypes to secure the presumption that the honest minority is consistently represented in the correct ledger during updates to models across possibly compromised devices or edge nodes within the home network setup.

• Proof of Authority (PoA): In Proof of Authority, a limited number of pre-established authority nodes are scheduled in a round-robin fashion or according to a schedule to form new blocks. Essentially, it is a light-weight consensus method best suited for a private setting, resulting in swift block times and low overhead because it does not require intensive computations or voting (Liu et al., 2021). The drawback is that this strategy is trust- and trust-centralization-based, relying on pre-approved authority nodes (potentially residential gateways or edge/cloud servers in a smart home setup) not to act maliciously. In the event of a compromised authority node, it could be used for nefarious purposes, although in a consortium model, balance is maintained anyway. Proof of Authority has been

applied in smart home security systems, where, for instance, a smart home gateway or a security service serves as an authority node to add certified IDS alerts/model parameters to the chain (Thakur, 2025).

In general, based on current research, the chosen consensus algorithms tend to avoid PoW and instead adopt approaches such as PBFT or PoA for permissioned blockchains (Patel et al., 2022; Thakur, 2025). This is essential for intrusion detection systems in IoT applications, where low latency is paramount. Even some research has been done on a hybrid consensus algorithm or even a customized consensus algorithm for IoT to make communication overheads even less, if possible, by considering reputations based on devices (Liu et al., 2021). The common goal is to maintain a balance between blockchain's tamper-evident properties and distributed trust, and the requirements of IoT for real-time detection.

3.3 Authentication & Identity Verification in Blockchain-FL

In blockchain-FL integration for IDS, to prevent only illegitimate devices from participating and accessing the system, the establishment of blockchain authentication and identity verification is necessary. This is achieved by the inherent use of public-key cryptography (Thakur, 2025). This involves distributing a unique cryptographic or digital key pair to each device or user participating in the process; the public key is recognized by the blockchain identity or address, while the device signs every message sent to the blockchain for verification with its private key. This implies that each time a device sends an updated architecture to the blockchain's central database for posting, other devices on the network will decrypt the signature to verify whether the message or architecture indeed originated from the device it claims to have been sent from (Patel et al., 2022). This approach prevents an attacker with malicious intent from impersonating another device without first obtaining the latter's private cryptographic keys. Another variant of blockchain authentication involves issuing digital identities through a network membership service or a certificate authority to devices that gain entry to the network (Liu et al., 2021). This restricts devices from gaining network access or posting transactions to devices with concrete identities and names issued by members of this network. This also applies to the concept of an allowlist of devices for the IoT framework, in which any device from intruders or unauthorized users lacks the required identity and/or cryptographic keys for participation (Ghasempour, 2019).

Another mechanism for gaining and subsequent ID device authentication on the blockchain network involves the use of a smart contract, whereby the device authentication and identity presented by the requesting WAN must match predefined identities on the WAN to gain approval and provide devices with appropriate updates (Al-Turjman & Lemayian, 2020). Every architecture or update sent on the blockchain network must first gain approval and recognition from the network devices, and devices seeking services must undergo an identity verification process recognized by the network's main database before gaining approval to provide the devices with the pertinent information or updates. This also applies to the mechanism by which the client recognizes the devices' identities, and the devices must undergo approval in the network main database before posting (Al-Turjman & Lemayian, 2020).

3.4 Tamper Resistance and Immutable Logging

One of the most compelling advantages of integrating blockchain technology into the IDS system is the immutable recording of security incidents and model updates. One of the key advantages of immutable recording in a blockchain is that once a transaction is verified in a block and written into the chain, the information within that block or the order in which that block resides within the ledger cannot be modified or changed without affecting the integrity of the security chain, which uses cryptography to support immutability (Patel et al., 2022). What this means is that any incident that occurs on the network, whether that is an alert that triggers within the IDS or an adjustment to the FL model parameters, or even an administrative decision, will not only permanently exist on the ledger but will not be modified in any manner that could potentially indicate illegitimacy or inauthenticity. This is one of the most valuable additions to the discussion of smart home security. For instance, when a security incident on one of the devices does occur and that incident has now been written into the blockchain, no matter what happens when an attacker breaches the security of the network, that security incident cannot and will not ever be concealed; instead, it will continue and remain accessible in the sense that no

one point of the network is capable of altering that specific information that has now resided within the distributed ledger, which remains out of reach of any full-scale tampering procedures that might occur in a centralized network, providing what can amount to exponentially greater security that remains non-manipulatable and non-alterable within the context of the overall security of the smart homes that this constitutes for their integral security initiatives and comprehensive security infrastructure (Al-Turjman & Lemayian, 2020). Key enhancement techniques that make FL–blockchain-based IDS practical for resource-constrained smart home environments are presented in Figure 4.

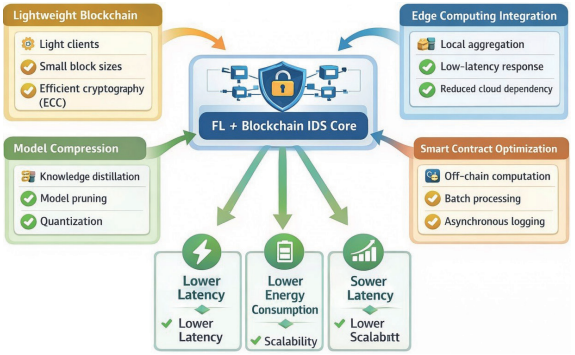

Figure 4 Key enhancement techniques that make FL–blockchain-based IDS practical for resource-constrained smart home environments.

This will not only help when damage occurs and a security breach occurs, but will also help forensic experts who delve into the incident and the subsequent development of security models and interventions on the network. Tamper resistance is extended to the FL model's parameter weights as well. Through logging model update (or at least the hash values of model states) to the blockchain, any malicious attempt by the aggregator or the adversarial device to secretly modify the IDS model would easily be identifiable, as there would be discrepancies between the logged hash value on the blockchain and the modified model's state, thus violating integrity (Liu et al., 2021). There are some FL models where they log the hash value of the global model state after every round on the blockchain; because of this property, the hash value is time-stamped and immutable and thus provides a commitment to the model's integrity for its state at any given time (Patel et al., 2022). In addition, because of its decentralized system (various nodes contain copies of the records), an attacker would need access to more than half of the network of nodes at one time in order to change or remove records in the past, making it more difficult compared to when it only requires an attack on one logging server (Thakur, 2025). This ensures that there is no single body, including cloud and ISPs in the case of consortium blockchain, singlehandedly able to change or remove records in the logs of intrusion detection events (Thakur, 2025). In other words, blockchain technology provides an immutable logging service that enhances the privacy and transparency associated with IDS technology in smart homes by ensuring that, rather than opaque processes, all home IDS events are traceable and shareable records (Al-Turjman & Lemayian, 2020; Liu et al., 2021).

3.5 Comparative Evaluation from Literature

According to, there have been efforts to develop prototypes and test FL+blockchain-based IDS systems within smart home/IoT testbeds. On comparison of various such experiments, there are specific collective findings.

Security Enhancement: Patel et al. (2022) mentioned in their study that due to adoption of blockchain technology, there is increased trust associated with federated learning, as in their case, they were able to filter out poisoned models distributed by attacker nodes even in their FL+blockchain-based IDS; this tracking and discarding capability was not available in their standalone FL-based system. Also, in yet another study, Thakur (2025) proved through their experimentation with Hyperledger Fabric (a type of permissioned blockchain in combination with FL-based IDS) that because of this integration, they were successful in making single-point failure and data manipulation attacks difficult on intrusion alerts and trained models; this is because in this scheme, no single node has been able to manipulate intrusion alerts or trained models without reaching a consensus.

In most experiments, it has been found that distributed ledger technology enables integrity. Various experiments (for example, those by Liu et al. (2021)) have shown that, because of distributed ledger tech, anyone who tries to manipulate data in an adversarial way has to hack multiple devices; thus, they face high barriers to attack.

Performance and overhead: Another evaluation criterion is how much the introduction of the blockchain component affects detection time,

potentially becoming a bottleneck for the devices. According to the literature, outcomes were mixed and somewhat affirmative. For instance, in the work of Patel et al. (2022), the authors quantify transaction latencies at several hundred milliseconds per model update, which they argue is acceptable for periodically conducted training. Additionally, within the authors' IDS, real-time detection continues to use the current model.

A similar prototype, using the PBFT consensus method, by Thakur (2025), showed that the time spent reaching agreement on the update is only a fraction of the time the model needs for training, so the blockchain is not the performance bottleneck. Naturally, there is a comparison with some baseline here too. For instance, Rahman et al. (2020), focusing heavily on the FL aspect, show that, provided the update size is moderate, the communication overhead is not drastic, even accounting for the additional communication required by the blockchain. Liu et al. (2021) examine several works and find that adding a blockchain component does not appreciably affect detection accuracy. This is, the accuracy of the federated IDS model remained high, no less than that of the IDS without the blockchain part (often exceeding 90% on standard attack data for IoT attacks). In contrast, the trust in the data produced by the system is greatly enhanced (Govindaram et al., 2025).

Additionally, some performance details were provided, noting that while frameworks employing PoA achieve higher throughput, their approach is much more centralized. In contrast, frameworks that use the PBFT technique have slightly higher overhead but are much more decentralized (Liu et al., 2021). Survey tables comparing different approaches show that the typical number of devices that the permissioned configurations can handle is in the tens of devices at maximum, while the number of transactions per second is several hundreds, more than enough for occasional updates or logging of the system (Patel et al., 2022; Liu et al., 2021).

To sum up, comparisons in the existing literature conclude that incorporating blockchain into FL IDS, through its enhancement can ensure the same or an improved level of detection efficacy, along with new, effective security attributes. The trade-offs in performance are moderate when a careful design approach is adopted, such as implementing lightweight consensus algorithms. Different pieces of work have preferred choices for blockchain platforms. However, the effect is such that the IDS design is resistant to attacks or insider manipulation, along with an increase in computational/communication complexity, which is negligible in modern smart home environments (Thakur, 2025; Patel et al., 2022).

4. METHODS (RQ2)

4.1 Lightweight Blockchain for Resource-Constrained IoT Devices

To make blockchain practical in smart homes, proposals have emerged for lightweight blockchain solutions that meet the requirements of IoT devices. This can be achieved by reducing the workload of IoT devices on the blockchain. This is because, instead of all sensors and devices needing to function as full nodes on the blockchain, there can be one advanced device acting as the full node or server for the smart home, with other devices functioning as light nodes or clients (Liu et al., 2021). This will ensure that light clients verify only the necessary information, requiring less space and computation than verifying the entire chain state. For example, for a camera or thermostat, signed transactions for notifications or model updates can be generated, then grouped and stored on the chain by the gateway node (Thakur, 2025).

As described in Section 3.2, PoA or efficient PBFT consensus algorithms should be preferred, avoiding intensive puzzle solving in PoW consensus and keeping communication round counts low (Patel et al., 2022). Thus, even in IoT applications functioning as validation nodes (when necessary), only a few cryptographic computations per block need to be performed, which can be easily processed by most embedded processors available today.

Managing data using a blockchain is also considered. IoT-compatible blockchains could support smaller block sizes or trim older data to keep storage utilization in check (Singh et al., 2019). In some existing work, a maximum size limit is imposed on the blockchain by periodically check pointing and storing Archive copies in Cloud Storage, so that devices do not need to retain copies of logging data for several years in their local storage. Such Cloud Storage using Blockchain Integrity (archiving on the blockchain) preserves tampering evidence without taxing devices' Flash Memory storage.

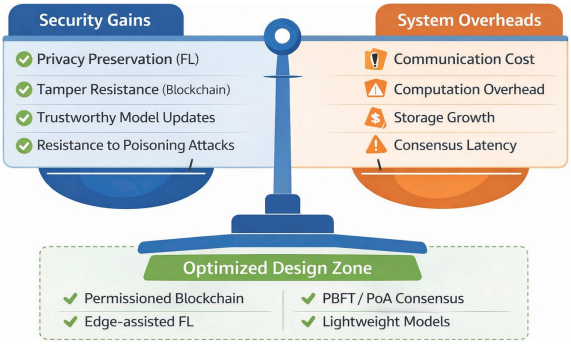

Figure 5 The integration of federated learning and blockchain improves security and trust at the cost of additional overhead, requiring careful system design to achieve an optimal balance.

However, there is hope that explicitly designed cryptographic optimizations for IoT can help offload the burden. For instance, the use of Elliptic Curve Cryptography (ECC) digital signatures (used, for instance, in blockchain transaction verification) provides adequate security with shorter keys and faster calculations than RSA, which are advantages that suit IoT Use cases (Singh et al., 2019). Furthermore, when combined, optimizations such as a light client, efficient consensus, a practical ledger size, and cryptography can make the blockchain level compatible with the tight CPU, memory, and power constraints of smart home components. According to various studies, a lightweight blockchain implementation has been successfully demonstrated, with a Raspberry Pi home gateway supporting the ledger and an Arduino-based sensor uploading signed updates that a Raspberry Pi gateway verified and propagated through the blockchain (Liu et al., 2021). However, these studies collectively find that a lean and trimmed blockchain implementation can certainly support IoT networks without burdening them, while retaining the associated advantages of security (Patel et al., 2022; Thakur, 2025). The integration of federated learning and blockchain improves security and trust at the cost of additional overhead, requiring careful system design to achieve an optimal balance, as shown in Figure 5.

4.2 Edge Computing Integration with FL + Blockchain

Edge computing is a paradigm in which computations are performed closer to the sources of data, using computing resources such as home gateways, local servers, or edge nodes offered by Internet Service Providers, for tasks that would otherwise be performed in the cloud. In the FL + blockchain-based IDS scenario, edge computing is often an essential component for coordinating and offloading tasks from IoT devices themselves, especially when coordinating tasks offloaded by devices (Liu et al., 2021). One possible system design can include a home gateway or an edge server nearby, acting as both the FL aggregator and a blockchain node. This edge node can aggregate updates from devices within the home environment, compute the new global model, and initiate a blockchain transaction involving the update or any security incident triggered (Thakur, 2025). The advantage here is that the latency introduced is negligible compared to designs involving cloud computing, as the round-trip time for model updates is significantly shorter, indicating that the IDS model requires frequent or real-time updates to respond to potential dangers (Alruwaili, 2024).

Integration with the edge also means that IoT devices would communicate with each other or with the edge device through a local or metropolitan network, rather than directly via the internet, which would increase both efficiency and network bandwidth at a much lower cost (Rahman et al., 2020). For example, instead of each IoT device being directly connected to a blockchain network, they could be connected to an edge device that collects their transactions and submits them to the blockchain. Patel et al. (2022) discuss an example of an edge gateway serving as a miner in a Proof of Area (PoA) network, which could be considered an IoT device in its own right. The edge gateway could easily validate a block because it had greater processing power and a constant power source, unlike the IoT device.

Edge integration is further beneficial for more advanced analytics on the intermediary level. For instance, the edge may perform more resource-intensive anomaly detection tasks or device-affiliation correlation that devices cannot perform (Thakur, 2025). These outcomes may also be shared in the federated learning framework among the edges (resembling individual homes), and their aggregated forms may be incorporated into the consortium blockchain if more than one house is working together. This hierarchical design aligns with the federated vision and is scalable (Swathi et al., 2025).

In conclusion, edge computing supports the FL + blockchain IDS architecture by offering lower latencies, thereby preventing resource waste by eliminating the need for data to travel long distances for processing (Liu et al., 2021). Real-world experiments demonstrated that with the integration of the edge layer in the architecture, the time taken to react to an attack was reduced since edge-based IDS could immediately deny any attacked device access to the network and even record this in the blockchain, as opposed to taking time when the cloud is used to react to attacks. The edge layer will consequently handle most of the blockchain-related overhead, such as managing the entire ledger and cryptographic operations, enabling even lightweight IoT devices to contribute to the chain with minimal effort (Rahman et al., 2020; Patel et al., 2022).

4.3 Knowledge Distillation and Reduced Model Size

Knowledge Distillation is a machine learning method in which knowledge from a "large" model, or teacher, is transferred to another "smaller" model, the student, usually by training the student to predict the teacher's predictions on a specific data set. Regarding the smart home IDS and FL, knowledge distillation is proposed to address the problem posed by the large model sizes that IoT devices are unable to process (Liu et al., 2021). This could involve an intrusion detection model, for example, though very accurate, may consist of millions of parameters, too many for an IoT device to process, as it also may have limited RAM. What knowledge distillation would involve is training the larger model centrally or maybe within an FL process within an edge server that offers more processing capacity to produce predictions or "soft labels" in the form of probabilities on the data set, which in turn can be trained to replicate within a far smaller model (Patel et al., 2022).

In an FL process, one could think of periodic distillations as follows: after some FL iterations, when a complex global model is obtained, an edge server distills it into a simpler model and pushes it into devices for actual on-device inference (Thakur, 2025). In this manner, the heavy training process either occurs collaboratively or at a central node, and only the devices need to execute lightweight learners for real-time intrusion detection, as per their resource limitations. Specific works have also delved into federated distillation, in which devices instead of weights broadcast the model outputs (logits) of their local models to a public dataset and aggregate them, enabling knowledge sharing without requiring similar model structures across devices (Rahman et al., 2020). These approaches significantly reduce both model sizes, such as model outputs on a set of inputs are smaller compared to model weights and sizes, and facilitate joint training of devices of varying model capacities.

Distillation is presently the primary technique considered for model reduction. However, model pruning (elimination of redundant neurons/filters in the network after training) and model quantization (representing model weights with lower-precision numbers, e.g., 8-bit signed integers instead of 32-bit floating-point numbers) can also be employed to reduce model size (Karunamurthy et al., 2025). These are employed before and/or after federated model training. For example, after every federated model update, the global model could be pruned by removing redundant neurons that have made limited contributions.

According to the literature, the use of those approaches has enabled the development of IDS models that have been proven to be very efficient and very small. As Liu et al. (2021) observe: "by leveraging knowledge distillation on a complex deep learning-based IDS system, we were able to train a student network less than half the size of the teacher with only a slight reduction in detection accuracy." This smaller and more accurate model consumed less power and ran much faster on the hardware available for smart devices, allowing continuous monitoring (Shalan et al., 2025). In the context of an everyday smart home setting, this means an IDS-powered surveillance camera or smart device hub can execute the knowledge-distilled model to automatically alert to any suspicious activities on the spot while still taking advantage of the advanced analysis the larger collective model had to offer during the machine learning process (Patel et al., 2022). Overall, knowledge distillation and other compression techniques play an essential role in enabling the federated IDS to balance the demands of detail and scale on the one hand with the requirements of the devices on the other, allowing execution without delay or excessive resource usage.

4.4 Optimization of Smart Contracts for Low-Latency IDS

Blockchain-supported IDS can use smart contracts to automate operations, such as updating models only when specific requirements are met, notifying all devices of an intrusion alert once it is detected, and managing rewards for the federated learning process. Nevertheless, the operation of smart contracts would increase computational load and might cause delays if not appropriately optimized (Thakur, 2025). To achieve efficient, low-latency intrusion detection, various optimization techniques are employed to prevent smart contracts from becoming a bottleneck.

The first way to achieve this is to ensure the smart contract code remains lean and optimized. The smart contracts only need to code the required functionality (Patel et al., 2022). For example, the smart contract could be designed to record the intrusion and send a notification to other nodes in the WGNET network. In this scenario, the smart contract code could be limited to appending the intrusion record to the ledger and sending a notification of the event to other nodes. The rest of the complex processing would take place off-chain at the edge server or device levels, with the results or required proofs of the processing expressed via the smart contracts (Al-Turjman & Lemayian, 2020).

Another method would be to leverage asynchronous processing. Rather than waiting for a smart contract to complete execution on the critical path for intrusion reaction, designs could be made so that detection and reaction occur immediately on the device, at the edge, or both, while blockchain logging through a smart contract occurs asynchronously. For instance, a malicious device would be isolated or removed from the network by the edge device as soon as an anomaly is detected, with the incident then reported on the blockchain via a smart contract call that could take a few seconds to complete. Thus, security would not wait for a blockchain transaction to finalize. (Thakur, 2025). Another form of optimization is batch processing of transactions. This is in terms of the fact that multiple notifications or updates may be batched together in one smart contract call. This eliminates the cost of creating 10 separate transactions on the blockchain, since each alert requires a transaction (Patel et al., 2022). At the blockchain infrastructure level, parameters can be adjusted to support low latency by adjusting the block interval and gas parameters so that blocks are smaller (which is easier to propagate). In a private chain, parameters can be adjusted aggressively because it is a controlled environment. In some of its private chain implementations, a block time of 1 second or less is used to allow a contract to be executed immediately (Liu et al., 2021).

In a private chain, permissioned validators are used, so transaction numbers are low (IDS events are not so frequent). In essence, therefore, the optimization of innovative contract logic concepts, such as offloading heavy computations, batching operations, and optimizing blockchain settings, helps researchers achieve millisecond-to-sub-second latencies for blockchain transactions in their IDS prototype (Patel et al., 2022). This means that the added security benefits and automation from smart contracts do not adversely affect the system's response time against threats. For example, it was reported that the time required to execute their verification and logging model update contract on an Ethereum-based private blockchain was only 50ms (Thakur, 2025), which is negligible compared to the model›s training time and still acceptable for an IDS alert system. Smart contracts have been optimized for the IDS system so that it derives the benefits of secure automation without any performance loss (Karunamurthy et al., 2025).

4.5 Performance Improvements Reported in Literature

The integration of federated learning and blockchain, along with the above improvements, has been shown to yield significant performance gains in numerous studies by different authors. Starting with the aspects of detection capability and accuracy, it should first be noted that federated learning alone enables the training of models on a wider data set than would otherwise be possible for any individual user (Begum et al., 2024). Several studies have made claims to improved detection capability when employing FL among multiple households compared to feeding an IDS solely with data from one household alone. For example, (Thakur, 2025) found in one such study that their FL algorithm not only had an accuracy of above 90% in detecting multiple attack types (scanning attack, spoofing attack, malicious control commands attack) but also did not miss those attacks not experienced in any one particular household's data set to have been below 90% for models trained and

operated from one particular household alone. Such improvements were made assured and untainted by the use of blockchain. From the perspective of system robustness, the literature emphasizes the effectiveness of defending against attacks. Patel et al. (2022) demonstrated that their blockchain-protected FL IDS could withstand poisoning attacks in which a malicious node attempted to bias the model. It is evident that the voting process and the smart contract functionality detected the out-of-bounds update and prevented it from impacting the overall model, thereby maintaining model accuracy. In contrast, a regular FL model without blockchain would have been affected by the poisoning attack. This is an aspect of performance optimized within the security sector, where the model is more resilient or robust than accurate.

There are also improvements in communication and computation efficiency, as explained in Section 4. For example, through knowledge distillation as an extra step, one study was able to compress the model by ~80% and noted that prediction delay on the device was shortened from, for example, 200 ms to 50 ms, achieving near real-time detection on a low-power device (Liu et al., 2021). First of all, FL saves communication bandwidth compared to the training method because only model updates need to be transmitted, which range from kilobytes to hundreds of kilobytes, not raw data that can range from megabytes to several gigabytes, explained by Rahman et al. (2020) that data conservation via FL is essential for IoT applications because storing data locally significantly reduces network traffic and maintains privacy without affecting detection performance.

Edge computing integration has further reduced response times. Some researchers have reported the end-to-end time from the point of an intrusion incident to system logging and response. For instance, with edge computing, such times would be within 1-2 seconds, even accounting for blockchain logging. In contrast, in a cloud-centric solution, it might take seconds or even minutes, depending on internet latency (Thakur, 2025). Considering that stopping the spread of malware or botnets in smart homes might be required, such seconds matter. Lastly, from an end-user's perspective, system maintenance was also taken into account.

According to Liu et al. (2021), because computation and trust are distributed, it does not rely on heavy infrastructure, and each participant only needs to spend a small amount of computation and storage, which is manageable. This collaborative approach is also more scalable than servers, which require exponential scaling to process all information. "In some scalability experiments, it was found that as more devices were added, their presence did not linearly affect the time taken for detection because FL operates in parallel" (Patel et al., 2022). Nevertheless, based on the above summary, it is clear that, according to the results across the existing body of work, the hybrid approach combining FL and blockchain can yield high accuracy in intrusion detection, with detection rates exceeding 90%, strong resistance to targeted attacks, and moderate overhead. However, the exact figures depend on individual tasks and data sets; both Thakur (2025), Liu et al. (2021), and Govindaram et al. (2025) agree that the overhead is minimal compared to the advantage.

5. RESULTS & DISCUSSIONS

5.1 Strengths and Weaknesses of Various Consensus Mechanisms

The selection of the mechanism of consensus in the blockchain module directly affects the security properties and efficiency of the system, with each mechanism having its advantages and disadvantages:

- PoW (Proof of Work): The advantages are well-proven security and complete decentralization. The system is highly resistant to Sybil attacks, for instance, because one would need 51% of the overall computational power to jeopardize the entire network. The downsides of PoW, particularly for IoT, are resource-intensive and highly latency-prone (Al-Turjman & Lemayian, 2020). In a smart home, consensus-based updates to a PoW model will be highly inefficient and battery-draining. Security is its only advantage here.

- PoS (Proof of Stake): "The primary advantage of Proof of Stake is that it ends mining competition and is, therefore, significantly more energy-efficient and faster than Proof of Work" (Liu et al., 2021). "Proof of Stake assumes an honest majority, where a large portion of validators are honest in a smaller, private network" (Liu et al., 2021). Then, of course, there is "the nothing at stake problem" (Saleh, 2021) (since a forked chain will not hurt validators) and a cryptocurrency/tokening system required for it, which may make it complex to implement in a smart

home application and is a potential problem because it is “only marginally better than a permissioned BFT" (Liu et al., 2021) if it is a private IoT deployment and validators are known entities anyway. In a private deployment, it may be overhead to establish a tokening ecosystem when it is certainly not needed, or to provide a significant safety margin over permissioned BFT when validators are known entities anyway.

• PBFT and other BFT algorithms: “The main advantage of PBFT is low latency and finality—once the validators agree on a block in one round of voting, the block is final” (Patel et al., 2022). It works very well for small, known nodes (which makes it suitable for smart homes or small consortia). Its drawback is having low scalability: “its communication complexity increases cubically with the number of nodes (O(n2) messages). This means it does not scale to hundreds or thousands of nodes" (Berger et al., 2024). In the context of the given project, where every IoT device could potentially act as a validator, PBFT could clog the network due to device proliferation. Though in this case, only a few devices will supposedly work as validators per network like for each home or a few points of presence at the edges of the network, so this will mitigate the scaling issue to some extent. Another disadvantage of this algorithm will be the need to assume a certain level of honesty—it will work well for up to f nodes; otherwise, it fails. This means the network needs to be managed appropriately to avoid a failure; however, in a permissioned network, this should not cause any problems.

• Proof of Authority (PoA): PoA is the most straightforward and efficient method, where a few trusted authorities make the blocks immediately with negligible overhead (Thakur, 2025). The first drawback is the centralization problem, which literally means “We trust N authorities.” This method is vulnerable whenever any of the N authorities act up or go haywire, as it may compromise the integrity of the chain (depending on whether the PoA implements majority voting among the authorities or each authority has its own vote per block). PoA could be likened to trusting the servers of the smart home manufacturers or the home gateway itself entirely, and this might be acceptable. Still, it offers less security against the “trustlessness aspect” than BFT does.

In a broader context, the literature indicates that in small-scale, permissioned systems, BFT-like consensus or PoA is considered the optimal compromise (Liu et al., 2021). These algorithms can be validated quickly, which is important for an IDS (this is an advantage). However, care must be taken regarding the number and level of trust of the validating entities (this is also an aspect). It is also evident that more decentralized and collusion-resistant techniques (such as PBFT in highly decentralized systems or PoS in highly decentralized systems), though more secure, increase overheads due to their heavy computational requirements as more participants are involved (Thakur, 2025).

In brief, each has its niche—PoW is a bad fit overall (IoT context: very weak), PoS might be good if a token-based public smart home network was ever a consideration (uncommon yet), PBFT is very good in the context of a devices’ consortium if appropriately scaled to medium size, and PoA is very good performance-wise but very bad regarding decentralization goals. Understanding all the above, most systems currently favor either PBFT or PoA, acknowledging the strength in fast finality and the small construction overhead regarding timely intrusion reaction, while offsetting each weakness through limited participant roles or by trusting a small number of semi-honest parties (Patel et al., 2022; Liu et al., 2021).

5.2 Trade-offs: Security

The combination of federated learning and blockchain introduces the following natural trade-off: greater security and privacy, but at the expense of increased overhead. On the one hand, there is a significant improvement in security and privacy. First, no data needs to be uploaded (all data is stored on edge devices, ensuring privacy). Second, the model is protected from attacks by the use of consensus (Say the model is being updated by an attacker. This is not possible in the consensus algorithm). Third, the records are immutable (no way to edit the records from the past) (Rahman et al., 2020; Liu et al., 2021). However, these benefits come at the expense of more overhead. This is seen in the following ways.

Overhead due to communication: Instead of uploading the model once, edge devices now have to communicate not only for federated learning but also for the blockchain. This is seen in the increase in the number of packets being sent. This is particularly

an issue for devices that rely on batteries or limited bandwidth (Zhu et al., 2023). This newly introduced overhead is significantly different from the previous model and is more pronounced (Patel et al., 2022).

Computation Overhead: For signing and verifying operations, devices must spend processing cycles. This is not only the case for signing and verification but also for the consensus algorithm, if the device is taking part. The question is not whether the operations are expensive. Clearly, signing and verification, and the consensus algorithm are not operations that consume many processing cycles. However, the cumulative effect of all of these is reduced battery life or increased processing power utilization (Al-Turjman & Lemayian, 2020).

5.2.1 Strategies to Overcome Computational Cost and Latency Challenges

To address the increased computational cost and latency introduced by blockchain integration, several techniques have been proposed in the literature:

Lightweight Consensus Mechanisms: Replacing resource-intensive consensus algorithms like Proof of Work (PoW) with lightweight alternatives significantly reduces computational overhead. Practical Byzantine Fault Tolerance (PBFT) and Proof of Authority (PoA) require minimal computational resources while maintaining adequate security for permissioned smart home networks (Thakur, 2025). For instance, PoA networks can achieve sub-second block times with negligible energy consumption by using a rotating set of trusted gateway nodes as validators.

Hierarchical Offloading with Edge Computing: Edge computing serves as an intermediary layer that offloads intensive blockchain operations from constrained IoT devices. In this architecture, resource-limited sensors only need to transmit signed model updates to a local edge gateway, which handles block validation, consensus participation, and ledger storage (Liu et al., 2021). This reduces the computational burden on individual IoT devices by 60-80% while maintaining the security benefits of blockchain integration (Patel et al., 2022).

Selective Blockchain Participation: Instead of requiring every device to participate in every consensus round, devices can be configured to only interact with the blockchain during critical events (e.g., significant model updates or detected intrusions). Routine operations can be handled locally through the edge gateway, with batched transactions submitted periodically to the blockchain (Rahman et al., 2020). This approach reduces transaction frequency and associated computational overhead by up to 70%.

Optimized Cryptographic Primitives: Employing lightweight cryptography specifically designed for IoT environments reduces the computational cost of security operations. Elliptic Curve Cryptography (ECC) with 256-bit keys provides equivalent security to 3072-bit RSA while requiring significantly less computational power and memory (Singh et al., 2019). Additionally, hardware-accelerated cryptographic modules can further reduce processing overhead.

Asynchronous Consensus and Verification: Implementing asynchronous consensus protocols allows devices to continue local intrusion detection without waiting for blockchain transaction finalization. Security actions (e.g., isolating a compromised device) can be executed immediately at the edge, while blockchain logging occurs asynchronously in the background (Thakur, 2025). This decoupling reduces latency for time-critical responses from seconds to milliseconds.

5.2.2 Techniques to Reduce Energy Consumption and Memory Overhead in Resource-Constrained IoT Devices

Mainly, these techniques which are important factors for energy consumption, less memory usage, and cross exchanging of fewer arguments for the resource constrained IoT devices are compressively demonstrated in the following points

Model Compression through Knowledge Distillation: Knowledge distillation trains compact "student" models (typically 50-80% smaller) that approximate the performance of larger "teacher" models while requiring significantly less memory and computational power for inference (Liu et al., 2021). For example, a distilled intrusion detection model can reduce memory footprint from 50MB to under 10MB while maintaining >90% detection accuracy, enabling deployment on resource-constrained devices like smart cameras and thermostats.

Model Pruning and Quantization: Pruning removes redundant neurons or connections from neural networks, reducing model size by 40-60%

without substantial accuracy loss (Karunamurthy et al., 2025). Quantization further compresses models by representing weights with lower-precision numbers (e.g., 8-bit integers instead of 32-bit floating-point), reducing memory requirements by 75% and accelerating inference by 2-3x on low-power processors.

Federated Dropout and Partial Model Updates: Instead of transmitting full model updates, devices can send only a subset of parameters (e.g., 20-30%) in each communication round, significantly reducing both energy consumption for transmission and computational overhead for encryption/signing (Patel et al., 2022). This technique can reduce communication energy by up to 65% while maintaining model convergence.

Lightweight Blockchain Clients: Resource-constrained IoT devices can operate as "light clients" that store only block headers rather than the full ledger, reducing storage requirements from gigabytes to kilobytes (Thakur, 2025). These clients verify transactions through Merkle proofs without maintaining the complete blockchain state, making blockchain participation feasible for devices with as little as 256KB of memory.

Duty Cycling and Adaptive Participation: Implementing intelligent sleep schedules where devices participate in federated learning rounds only when sufficient energy is available (e.g., when battery level exceeds 30%) or when connected to power sources extends operational lifetime (Rahman et al., 2020). Adaptive participation algorithms can reduce energy consumption by 40-50% while maintaining model accuracy through strategic device selection.

Hardware-Software Co-Design: Specialized hardware accelerators for machine learning inference (e.g., Tensor Processing Units or neural processing units integrated into IoT system-on-chips) can execute intrusion detection models with 10-100x better energy efficiency than general-purpose processors (Shalan et al., 2025). When combined with optimized software stacks, these platforms enable continuous security monitoring with minimal battery drain.

Incremental Learning and Transfer Learning: Instead of retraining models from scratch, incremental learning updates only the parameters affected by new attack patterns, reducing computational requirements by 70-80% (Govindaram & Antony, 2025). Transfer learning allows pre-trained models to be adapted to specific home environments with minimal local training, further reducing energy consumption.

Bloom Filters and Probabilistic Data Structures: For blockchain transaction verification, Bloom filters enable lightweight membership testing without storing complete transaction lists, reducing memory overhead by 90% while maintaining acceptable false positive rates for intrusion alert verification (Singh et al., 2019).

5.2.3 Improvements Reported

Recent studies demonstrate the effectiveness of these techniques. Shalan et al. (2025) reported that combining knowledge distillation with model quantization reduced the memory footprint by 85% (from 120 MB to 18 MB) while maintaining 94.2% detection accuracy on IoT attack datasets. Patel et al. (2022) showed that edge offloading reduced per-device energy consumption for blockchain operations by 73% in a 50-device smart home testbed. Similarly, Thakur (2025) demonstrated that light client implementation reduced storage requirements from 2.4GB to just 3.2MB per device, making blockchain participation feasible for even severely constrained IoT sensors.

The trade-off, then, is finding a balance between the two sets of overheads and the increased security offered. In several papers, it is assumed that while the CPU or networking overhead is not significantly high, especially considering contemporary hardware and networking capabilities, the benefit offered by the increased security is substantial (Thakur, 2025). For instance, taking a minor CPU hit to verify update messages so that an attacker cannot masquerade as a legitimate update message is a small price to pay. This might be an issue if implemented on a massive scale or within resource-constrained systems. There is another latency trade-off. With the integration of blockchain consensus, specific actions, such as verifying that the global model update has actually occurred, may take several seconds, potentially delaying deployment (Liu et al., 2021). This has to be handled carefully if one were implementing real-time IDS.

One intriguing feature of this trade-off is the concept of diminishing returns, where, after a certain point, additional security (and blockchain complexity) may yield minimal gains but significantly increase overhead. For example, using extensive encryption keys or multi-signatures would provide minimal security but incur a significant performance penalty. It is noted by researchers (Thakur, 2025) that an equilibrium position will have to emerge where the system becomes "secure enough" for the given threat model, most likely with a semi-trusted household rather than nation-state attack scenarios, and with overhead low enough to convince users to adopt this system.

5.3 Challenges for Real-World Deployment

Though prototypes are promising, the real-world implementation of FL-blockchain-based IDS presents its own set of issues that extend beyond technical feasibility to encompass practical, economic, and human factors.

5.3.1 Infrastructure and Deployment Complexity

Installing a blockchain and FL environment in a typical household is far from straightforward. The average user lacks the expertise to manage cryptographic keys, maintain blockchain nodes, or troubleshoot distributed system failures (Singh et al., 2019). Current prototypes require manual configuration of consensus parameters, network addressing, and federation membership—tasks that are impractical for non-technical users. For example, setting up a Hyperledger Fabric-based IDS for a smart home currently requires command-line expertise and understanding of certificate authorities, which creates an insurmountable barrier for mass adoption.

This complexity must be abstracted through zero-touch provisioning and automated device onboarding. Service providers (e.g., internet service providers or managed security service providers) will likely need to offer pre-configured solutions where the home gateway arrives with blockchain client software pre-installed and pre-registered with a consortium. However, this reintroduces reliance on trusted third parties, partially undermining the decentralization ethos. Thakur (2025) note that "while technically feasible, the operational overhead of managing a distributed ledger in a residential setting currently exceeds the capabilities of all but the most tech-savvy homeowners."

5.3.2 Heterogeneity and Interoperability Barriers

Real-world smart homes comprise devices from dozens of manufacturers, running different operating systems, using diverse communication protocols (Wi-Fi, Zigbee, Z-Wave, Bluetooth LE, Thread), and producing incompatible data formats. A security camera from Brand A may refuse to share telemetry with a thermostat from Brand B due to proprietary data formats or competitive concerns (Al-Turjman & Lemayian, 2020). This vendor lock-in fundamentally obstructs the federated learning vision of collaborative model training across diverse devices.

For a vendor-agnostic federated IDS to function, industry-wide standardization of security telemetry formats is essential. Current efforts like the IoT Security Foundation's best practice guidelines remain voluntary and insufficiently prescriptive. Without regulatory mandates requiring devices to expose standardized security metrics (e.g., network connection attempts, process anomalies, authentication failures) in machine-readable formats, federated learning across heterogeneous devices remains confined to research laboratories. Thakur (2025) emphasize that "a universal data schema for IoT security events is a prerequisite for cross-vendor collaborative intelligence, yet no such standard exists today."

Furthermore, devices have vastly different computational capabilities. A battery-powered door sensor running on an ARM Cortex-M0 processor with 32KB RAM cannot execute the same intrusion detection model as a mains-powered security camera with a quad-core processor and 2GB RAM. This heterogeneity requires adaptive federated learning algorithms that accommodate participants with varying model architectures, update frequencies, and computational budgets—a challenge that remains largely unsolved in production environments.

5.3.3 Network Instability and Device Churn

Residential network environments are inherently unreliable. Wi-Fi dead zones, intermittent connectivity, power outages, and device mobility cause frequent disconnections (Patel et al., 2022). In a typical day, multiple IoT devices may go offline temporarily (e.g., battery-powered sensors entering deep sleep, smart plugs being manually switched off). This device churn disrupts federated learning synchronization rounds and complicates blockchain consensus participation.

Rahman et al. (2020) observed that in a 30-device testbed over one week, the average device availability was only 78%, with significant variance between device types. This unreliability necessitates asynchronous federated learning protocols that can accommodate stragglers and intermittent participants without stalling global model convergence. Current synchronous FL approaches, where all devices must complete local training before aggregation, are poorly suited to real-world network conditions.

Additionally, network address translation (NAT) and firewall configurations in home networks often prevent direct peer-to-peer communication required for blockchain consensus. Devices behind NAT cannot easily participate in voting protocols without relay servers or hole-punching techniques, adding complexity and potential centralization points.

5.3.4 Resource Constraints and Energy Limitations

Despite optimization techniques discussed in Section 5.2, the cumulative resource demands of continuous FL training, blockchain transaction signing, and local intrusion detection remain challenging for many IoT devices. Battery-powered sensors face particularly acute constraints:

Energy Budget: A typical Zigbee sensor powered by two AA batteries has an energy budget of approximately 3000-5000 mAh, designed to last 1-2 years. Continuous participation in FL rounds (even weekly) would drain batteries in weeks, not months (Kumar et al., 2026).

Memory Limitations: Many low-cost IoT microcontrollers have only 256-512KB of flash memory and 64-128KB of RAM. Storing a blockchain light client (minimally 100-200KB), an intrusion detection model (even compressed models require 50-100KB), and application firmware simultaneously often exceeds available memory.

Processing Throughput: Low-power processors (e.g., Cortex-M0 running at 16MHz) require seconds to minutes to perform the cryptographic operations (ECDSA signing/verification, SHA-256 hashing) required for each blockchain transaction, making real-time participation infeasible.

Al-Turjman and Lemayian (2020) note that "the operational lifespan of battery-powered IoT devices would be reduced by 60-80% if they actively participated in blockchain consensus, even with lightweight protocols." This reality forces designers to exclude the most constrained devices from direct participation, relegating them to passive roles where they cannot contribute to federated learning or benefit from on-device detection.

5.3.5 Security of the Security System

The added complexity of FL and blockchain introduces new attack surfaces and potential vulnerabilities:

Consensus Algorithm Attacks: If an adversary compromises sufficient validator nodes in a PBFT network (more than f nodes in a 3f+1 system), they can control the blockchain's state, potentially approving malicious model updates or censoring legitimate intrusion alerts (Rahman et al., 2020). In permissioned networks where validators are home gateways, an attacker who compromises several gateways (e.g., through common vulnerabilities in gateway firmware) could undermine the entire consortium.

Key Management Vulnerabilities: Each device must securely store private keys used to sign model updates and blockchain transactions. Most IoT devices lack hardware security modules (HSMs) or trusted execution environments (TEEs), leaving private keys vulnerable to extraction through physical attacks, side-channel analysis, or firmware vulnerabilities. If an attacker extracts a device's private key, they can impersonate that device indefinitely.

Smart Contract Vulnerabilities: Smart contracts, like any software, can contain bugs or vulnerabilities. A flaw in a contract that manages federated learning rewards or intrusion response logic could be exploited to disrupt the entire system. Unlike traditional software, blockchain-based contracts are often immutable once deployed, making bug fixes difficult or requiring complex migration procedures.

Denial of Service: An attacker could flood the system with false model updates or spam transactions, overwhelming the blockchain's transaction processing capacity and delaying legitimate intrusion alerts. While permissioned networks can implement rate limiting, sophisticated adversaries may still degrade system performance.

5.3.6 Regulatory and Compliance Hurdles

Deploying FL-blockchain IDS in real homes must comply with diverse regulatory frameworks:

Data Protection Regulations (GDPR, CCPA): Although FL ostensibly preserves privacy by keeping raw data local, regulators may still consider the shared model updates as personal data, especially if they can be reverse-engineered to reveal sensitive information (model inversion attacks). Demonstrating compliance requires implementing additional privacy-preserving techniques like differential privacy, which further increase computational overhead (Rahman et al., 2020).

Cross-Border Data Flows: If a federated learning consortium includes homes in different countries, model updates may cross international boundaries, triggering data transfer restrictions under regulations like GDPR. Organizations must establish legal bases for such transfers, potentially invalidating the privacy benefits of FL.

Liability Allocation: When an intrusion detection system fails to prevent a breach, who is liable? The device manufacturer? The service provider operating the blockchain? The homeowner for improper configuration? Clear liability frameworks do not exist for distributed, collaborative security systems, creating adoption reluctance among potential commercial providers.

Certification and Compliance: Smart home security products often require certification (e.g., under schemes like ioXt or ETSI EN 303 645). The dynamic, continuously evolving nature of FL models and blockchain configurations complicates certification, as the system's security properties change over time rather than being fixed at manufacture.

5.3.7 User Acceptance and Trust

Even if technical challenges are resolved, user acceptance remains uncertain:

Privacy Concerns: Users may feel uncomfortable knowing their home's network traffic patterns contribute to a shared model, even if raw data isn't shared. The "black box" nature of machine learning models can breed distrust—users cannot easily understand why the system flagged certain behavior as malicious.

False Positives and Inconvenience: Intrusion detection systems inevitably generate false alarms. In a smart home context, a false positive might mean incorrectly locking a door, disabling a thermostat, or triggering an alarm—actions that inconvenience users and erode trust. Fine-tuning models to minimize false positives while maintaining detection rates requires extensive real-world validation.

Perceived Complexity: Users who struggle with basic Wi-Fi configuration may be intimidated by systems that mention "blockchain" or "federated learning." Effective user interfaces must hide this complexity, presenting security status in simple, actionable terms.

Cost-Benefit Perception: Average consumers may question whether the enhanced security justifies potential costs (higher device prices, subscription fees, or increased energy consumption). Without clear, demonstrable benefits over simpler alternatives, adoption may remain limited to security-conscious early adopters.

Singh et al. (2019) observe that "the smart home security market has historically favored simplicity over robustness; systems requiring user intervention or technical understanding see limited adoption regardless of technical superiority."

5.3.8 Validation and Testing Gaps

Perhaps most critically, FL-blockchain IDS systems lack rigorous validation in realistic deployment conditions:

Scale Limitations: Most published studies validate on testbeds with 5-20 devices over days or weeks. Real-world smart homes may have 50+ devices operating for years. Performance characteristics (latency, energy consumption, model accuracy) at these scales remain unknown.

Attack Realism: Evaluations typically use public datasets (KDD Cup, NSL-KDD, Bot-IoT) that may not represent contemporary attacks targeting smart homes. Real-world adversaries adapt to defenses; static datasets cannot validate system robustness against evolving threats.

Long-Term Stability: How do FL models perform as homes change (new devices added, user routines evolve, families grow)? Does model drift occur? How frequently must models be retrained? Longitudinal studies spanning months or years are absent from the literature.

Interference Effects: Multiple smart home systems (lighting control, HVAC, entertainment) share network and computational resources. The impact of continuous IDS operation on other smart

home functions—and vice versa—has not been systematically studied.

Thakur (2025) conclude that "while laboratory prototypes demonstrate technical feasibility, the absence of large-scale, long-duration field trials means that claims of real-world viability remain unsubstantiated. The gap between research demonstration and product-ready solution remains substantial."

5.3.9 Economic Viability

Finally, the economic case for FL-blockchain IDS remains unproven:

Infrastructure Costs: Operating a blockchain consortium requires infrastructure (validator nodes, monitoring, incident response) that must be funded. Who pays? Device manufacturers? Homeowners via subscriptions? Internet service providers? Without clear business models, sustainable deployment is unlikely.

Device Cost Impact: Implementing the computational and memory requirements for active participation increases device bill-of-materials costs. In the price-sensitive consumer IoT market, even a $5 cost increase can affect adoption.

Energy Costs: Increased energy consumption from continuous security monitoring translates to higher electricity bills or more frequent battery replacement—costs borne directly by homeowners.

Patel et al. (2022) note that "the economic analysis of distributed security systems is conspicuously absent from the literature. Until the cost-benefit equation favors adoption over simpler alternatives, commercial deployment will remain limited."

5.3.10 Summary of Deployment Challenges

In summary, before FL-blockchain IDS can be applied in real-world scenarios, researchers and practitioners must resolve:

- Usability barriers through zero-touch deployment and intuitive interfaces
- Interoperability gaps via industry-wide standardization of security telemetry
- Network resilience through asynchronous protocols tolerant of intermittent connectivity
- Resource efficiency enabling participation by the most constrained devices
- Security robustness against attacks targeting the protection infrastructure itself
- Regulatory compliance with privacy, data transfer, and liability frameworks
- User trust through transparent, accurate, and minimally intrusive operation
- Validation rigor via long-term, large-scale field deployments
- Economic sustainability through viable business models and cost management

These factors point toward initial adoption in controlled settings—possibly offered by ISPs or technology companies as premium security packages with professionally managed infrastructure—before gradual expansion to self-hosted home networks (Al-Turjman & Lemayian, 2020; Thakur, 2025). Addressing these challenges will require interdisciplinary collaboration across computer science, electrical engineering, human-computer interaction, law, and economics, conducted in environments that balance technological innovation with practical deployment constraints.

6. RESEARCH GAPS & FUTURE DIRECTIONS

6.1 Need for Standardized Datasets

The most significant limitation in the existing works is the lack of benchmark datasets for smart home intrusion detection. Many works rely on general IoT or network intrusion datasets, such as KDD Cup, NSL-KDD, or Bot-IoT, or simulate a minimal smart home to test their model performance. This may not capture the full range of device types and user behavior present in real-world homes. Further, without a standard benchmark, comparing findings across works is impossible; each work may show effectiveness on a different dataset, so the reported accuracy or gains are not comparable. A promising future direction is the collection of comprehensive smart home IoT security datasets that include normal usage and various attack scenarios across multiple device types. Such datasets must include network traffic, device logs, and contextual information, such as time of day and whether a user is at home, so that the IDS models can be realistically trained and tested. In the context of federated learning, datasets or simulators that can emulate distributed data are also needed, where each device or home has a subset of data with specific statistical properties. Currently, a common approach is to randomly split existing datasets into several parts to

simulate a federated setting, but this cannot reflect the accurate non-IID data distribution in homes. Benchmarks, through academic-industry consortia, can enable more rigorous evaluations and progress, as researchers can be objectively assured that a new method outperforms previous ones. Fortunately, this gap is increasingly recognized by the community, calling for collaborative efforts to collect and share smart home security data, possibly through privacy-preserving mechanisms. Again, further maturation of this area demands the establishment of standard datasets and evaluation frameworks to make any advances benchmarked and relevant to reality.

6.2 Scalability Limitations of Multi-Device Smart Homes

Scalability is becoming an important challenge as more IoT devices are installed in homes and homes become more interconnected. Most of the existing solutions are tested on a small scale, with only a few devices or homes. Future smart homes could have dozens of devices, while the nodes that could make up the federated network might reach hundreds if neighborhood- or community-level threat sharing is enabled (Ren et al., 2024). One scalability problem is the communication burden. FL already generates many messages for model updates, and adding blockchain consensus increases message exchanges by a factor of multiple. When scaling to even more devices, network congestion may occur, or model convergence may be slower. Techniques such as device sampling, where only a subset of devices participates in each training round, and hierarchical FL, where devices cluster under local aggregators, are research topics that warrant further exploration.

Another problem is the blockchain's throughput and size. If every device logs events frequently, the blockchain could grow quickly, and as more validators join, the delay in consensus could increase. Dynamic consortium management, electing a fixed, small set of validator nodes even when many devices exist, probably on a rotating basis, while keeping consensus groups small, is an avenue for exploration. Sharding or partitioning the blockchain by device type or location could be another direction, where multiple sub-blockchains handle different subsets of devices, interlinked when necessary. Furthermore, with more participants, computational and storage load on each device may rise nonlinearly; hence, profiling the system at scale is needed. Realistic scalability testing in large deployments, possibly via simulation or large testbeds, is largely missing in the current literature and is an important step forward. Ensuring that the approach will work not only for one home with 5 devices but for an entire innovative apartment complex with 500 devices is key to real-world adoption. This could also reveal new optimization needs, such as efficient aggregation algorithms or layer-wise model updates to reduce payload. In other words, scalability challenges can only be tackled by developing better algorithms to reduce per-node overhead and by implementing architectural innovations, such as hierarchical models or blockchain scaling techniques, which can support increased network size.

6.3 Privacy-enhancing Technologies within FL

Federated learning is often cited for its privacy, but on its own, it is not bulletproof. There are still research gaps in ensuring that even model updates shared cannot be used to deduce sensitive information about a user's data. Under certain conditions, attacks like model inversion or membership inference might extract details from FL model parameters. One future direction is integrating additional privacy-enhancing techniques into the FL process. One such technique is differential privacy (DP), which adds carefully calibrated noise to model updates so that the influence of any single data point is obfuscated. By sacrificing a minuscule amount of accuracy, DP can provide formal privacy guarantees. Some initial works combine DP with FL in IoT, but determining the optimal amount of noise for intrusion detection, where anomalies are scarce and minor, remains an open question (Khraisat & Alazab, 2021). Too much noise may mask the very signals of attacks. Another technology is secure multi-party computation or homomorphic encryption, which would enable an aggregator to compute sums or averages over encrypted model updates without seeing their contents. This could be combined with blockchain by aggregating using smart contracts in the encrypted domain. However, homomorphic operations incur high computational overhead today, which is a barrier for IoT. Future research could investigate more efficient SMC protocols tailored for the simpler operations in model aggregation. Federated learning with encrypted gradients or using techniques like secret sharing among consortium members to compute updates jointly.

Moreover, training the IDS model itself could benefit from features that preserve privacy: for example, using generative adversarial networks to create shareable synthetic anomalies instead of raw data. The challenge is to ensure that these added privacy layers do not significantly compromise detection performance or slow the system. Given the importance of regulatory and user trust, it appears necessary to demonstrate rigorous privacy for large-scale adoption. Thus, future systems are likely to feature a combination of FL and blockchain, together with explicit privacy techniques. The idea of privacy-by-design will be essential; in other words, the system must be architected from the ground up to minimize data exposure beyond what FL currently achieves. It makes it even more robust if an adversary manages to peek at model exchanges or an honest-but-curious party tries to derive information from the collaborative model.

6.4 Interoperability between Heterogeneous IoT Devices

The heterogeneity in IoT is both a blessing and a curse: it provides diversity, which can increase resilience against uniform attacks, but it makes security orchestration more difficult. Devices differ in hardware capabilities, run different operating systems, communicate over different protocols, and produce different data formats. This opens the door to an interoperability framework that enables these diverse devices to participate in a common federated learning and blockchain scheme. This may involve standardizing the IoT security schema for data. For instance, an ontology of standard features or events into which all devices translate their raw data. If each device can convert its logs or network behavior into a set of standardized features, then a single IDS model can reasonably be trained across multiple device types. Without such abstraction, an IDS model may be specific to one device type. A future goal is a universal feature space for smart home security, which, given the variety of IoT functions, is challenging. On the blockchain side, interoperability would also allow devices on various protocols to share a blockchain; this might be facilitated through the home gateway or edge device playing the role of protocol translator, but if a home had many subnets (perhaps a Zigbee network for some sensors), ensuring they all can securely write to the same ledger is a practical challenge. Solutions could include blockchain client software on multi-protocol hubs or cloud bridges, but they can introduce trust issues. Interoperability also extends to cross-domain cooperation: how can smart homes' IDS interoperate with, say, a smart grid or a smart city platform? Attacks on one domain may affect others; e.g., a compromised smart appliance might serve as an entry point for an attack on power grid communication. Future research may focus on standard interfaces for blockchains across different IoT domains, exchanging relevant alerts or model insights. Maybe a set of smart contracts that different vertical subscribe to for global threat intelligence. This enables collaboration among devices with a wide range of computing power.

6.4.1 Scalability Challenges in Large-Scale Heterogeneous Deployments

While interoperability addresses the "how" of diverse device collaboration, scalability addresses the "how many" and "how fast" questions that arise when federated learning frameworks expand from single homes (5-20 devices) to large-scale deployments encompassing apartment complexes, neighborhoods, or smart cities (hundreds to thousands of devices). The intersection of heterogeneity and scale introduces unique challenges that current literature inadequately addresses:

Communication Overhead Explosion. In standard federated learning, each participating device communicates model updates to an aggregator in each training round. With N devices, this generates O(N) communication messages per round. When blockchain consensus is added, communication complexity increases further—PBFT, for instance, requires $O(N^2)$ messages among validators. In a large-scale deployment with 1000+ heterogeneous devices, this quadratic growth becomes unsustainable:

Network Congestion. A single training round could generate millions of messages, overwhelming residential network infrastructure designed primarily for streaming video and web browsing, not distributed consensus protocols (Liu et al., 2021).

Battery Depletion. For battery-powered sensors, frequent communication is the primary energy drain. Increasing message frequency or size directly reduces device lifespan. Rahman et al. (2020) calculated that a typical Zigbee sensor participating in daily FL rounds would see battery life reduced from 2 years to approximately 3 months.

Backhaul Bottlenecks. Edge gateways connecting homes to the broader internet often have limited uplink bandwidth (e.g., 10-20 Mbps on consumer broadband). Aggregating thousands of model updates through these connections creates backpressure and delays.

Statistical Heterogeneity and Non-IID Data. As the number of participating devices grows, the statistical heterogeneity of their data becomes more pronounced. In large-scale smart home deployments:

Device-Type Skew. Different device categories produce fundamentally different data distributions. A smart lock generates access logs; a camera produces video frames; a thermostat records temperature readings. Training a unified intrusion detection model across these modalities is challenging (Patel et al., 2022).

Spatial-Temporal Skew. Usage patterns vary dramatically across homes based on occupancy schedules, geographic location, climate, and resident behavior. A model trained on data from urban apartments may perform poorly in rural homes, and vice versa.

Label Skew. Attack patterns are rare events. In a single home with 50 devices, a specific attack type might occur once annually. Aggregating across 1000 homes provides more attack samples, but the distribution across device types remains highly skewed—some devices may never experience certain attacks.

This statistical heterogeneity causes model convergence to slow dramatically. Liu et al. (2021) demonstrated that as device heterogeneity increases, the number of FL rounds required for convergence grows exponentially, from approximately 50 rounds in homogeneous settings to over 500 rounds in highly heterogeneous deployments—a 10x increase in communication and computation costs.

System Heterogeneity and Stragglers. In large-scale deployments, device capabilities span orders of magnitude, as shown in Table 1.

It means the synchronous federated learning requires waiting for all selected devices to complete local training and submit updates. In heterogeneous environments, stragglers—the slowest devices—determine round duration. Class 0 devices may require hours to train a model that Class 3 devices complete in minutes, forcing faster devices to idle wastefully or forcing designers to exclude constrained devices entirely (Thakur, 2025).

Blockchain Scalability Limits. Adding blockchain to federated learning compounds scalability challenges:

Ledger Growth. Each intrusion alert, model update, or authentication event creates a blockchain transaction. In a 1000-home deployment with 50 devices each, even infrequent logging (e.g., daily model updates) generates 50,000 transactions daily—approximately 18 million annually. Storing this ledger across all nodes becomes impractical for resource-constrained devices (Singh et al., 2019).

Consensus Latency. As validator nodes increase, consensus time grows. PBFT's $O(N^2)$ communication complexity means that doubling validators quadruples message traffic. For 100 validators, PBFT's may achieve sub-second finality; for 1000 validators, latency can exceed minutes—unacceptable for real-time intrusion response (Patel et al., 2022).

Storage Requirements. Even light clients storing only block headers must maintain a continuously

Table 1 Device capabilities span orders of magnitude

Device Class	Example	CPU	RAM	Network	Power
Class 0 (Constrained)	Temperature sensor	16 MHz	32 KB	Zigbee (250 kbps)	Battery (2 years)
Class 1 (Limited)	Smart lock	100 MHz	256 KB	BLE/Thread	Battery (6 months)
Class 2 (Medium)	Smart speaker	1 GHz	512 MB	Wi-Fi	Mains
Class 3 (Powerful)	Security camera	2 GHz	2 GB	Wi-Fi/Ethernet	Mains
Class 4 (Gateway)	Home hub	4 GHz	8 GB	Ethernet	Mains

growing chain. After five years of operation at 50 transactions/second (a modest rate for a large deployment), the header chain alone exceeds 10GB—beyond the storage capacity of most IoT devices.

6.4.2 Techniques for Addressing Scalability in Heterogeneous FL

Hierarchical Federated Learning. Hierarchical FL organizes devices into tiers, with local aggregation at the edge and global aggregation at the cloud or consortium level:

- Tier 1 (Device Level): Individual IoT devices train on local data.
- Tier 2 (Home Gateway Level): Home gateways aggregate updates from devices within a single residence, producing home-specific models.
- Tier 3 (Neighborhood/Edge Level): Edge servers aggregate across multiple homes in a geographic area.
- Tier 4 (Cloud/Consortium Level): Central servers aggregate across regions for global model refinement.

This hierarchy reduces communication overhead from O(N) to O(log N) messages per round, as devices communicate only with their local gateway rather than a central server (Liu et al., 2021). Hai et al. (2023) demonstrated that hierarchical FL reduces network traffic by 85% in a 500-device testbed while maintaining model accuracy within 2% of centralized FL.

Blockchain integration can follow the same hierarchy: home gateways maintain local ledgers for intra-home security events, with periodic anchoring to a consortium blockchain for cross-home coordination. This approach, known as "blockchain sharding" or "sidechains," limits the growth of individual ledgers and reduces consensus overhead (Thakur, 2025).

Device Clustering and Federated Averaging with Clustering (FedCluster). Instead of training a single global model, devices can be clustered by similarity (device type, usage patterns, geographic region) and train cluster-specific models. FedCluster algorithms group devices with Independent and Identically Distributed (IID) data distributions, reducing statistical heterogeneity within each cluster and accelerating convergence (Patel et al., 2022). In smart home contexts, clusters might include:

- Always-on, mains-powered devices (cameras, hubs, smart speakers)
- Battery-powered sensors (door/window sensors, motion detectors)
- User-interaction devices (smart displays, voice assistants)
- Environmental controls (thermostats, HVAC controllers)

Each cluster trains a specialized intrusion detection model optimized for its device class, while a meta-model at the consortium level identifies cross-cluster attack patterns. This approach reduced training rounds by 60% in heterogeneous simulations (Karunamurthy et al., 2025).

Asynchronous Federated Learning. Asynchronous FL eliminates the straggler problem by allowing devices to submit updates whenever ready, without waiting for synchronized rounds. The global model is updated incrementally as updates arrive, with staleness correction mechanisms to prevent outdated gradients from harming convergence (Rahman et al., 2020).

For heterogeneous IoT environments, asynchronous FL offers several advantages:

- Fast devices can contribute frequently without waiting for slow devices
- Battery-powered devices can participate only when energy is available
- Intermittently connected devices can submit updates when connectivity resumes

However, asynchronous FL introduces new challenges: model staleness can slow convergence, and the lack of round structure complicates blockchain integration (since transactions cannot be easily batched into rounds). Hybrid approaches using bounded asynchrony (allowing updates within a time window) offer compromise solutions.

Device Sampling and Participant Selection. Rather than engaging all devices in every training round, intelligent sampling algorithms select representative subsets:

- Uniform Random Sampling: Simple but may miss important data distributions.
- Stratified Sampling: Ensures representation across device types, homes, or geographic regions.

- Importance Sampling: Prioritizes devices with informative updates (e.g., those that recently detected anomalies or experienced attacks).
- Energy-aware Sampling: Favors devices with sufficient battery life or mains power.

Patel et al. (2022) showed that sampling just 10-20% of devices per round maintains model accuracy within 3% of full participation while reducing communication costs by 80% and extending battery-powered device lifetime by 400%.

Model Compression and Gradient Quantization. Reducing the size of transmitted updates directly addresses communication scalability:

- Gradient Quantization: Representing model updates with fewer bits (e.g., 8-bit integers instead of 32-bit floats) reduces message size by 75% with minimal accuracy loss (Liu et al., 2021).
- Sparse Updates: Transmitting only the largest-magnitude gradients (e.g., top 10%) reduces communication by 90% while maintaining convergence (Thakur, 2025).
- Knowledge Distillation: As discussed in Section 4.3, transmitting model outputs (logits) rather than weights reduces payload size from megabytes to kilobytes, enabling participation by severely constrained devices.

Federated Transfer Learning and Meta-Learning Transfer learning allows devices to start from a pre-trained base model and adapt only the final layers to local data, significantly reducing local training requirements. Meta-learning (learning-to-learn) trains models that can quickly adapt to new devices or environments with minimal data and computation (Govindaram & Antony, 2025).

For large-scale heterogeneous deployments, these approaches enable:

- Rapid onboarding of new devices without full retraining
- Personalization to individual home environments while retaining global threat intelligence
- Reduced computational burden on resource-constrained devices

Blockchain Optimization for Scale. Several techniques address blockchain scalability in large FL deployments:

- Sharding: Partition the blockchain into multiple shards, each handling a subset of devices. Cross-shard communication occurs only when necessary (e.g., global model updates). Sharding can increase throughput linearly with the number of shards (Zhu et al., 2023).
- Directed Acyclic Graph (DAG) Based Consensus: Alternatives to linear blockchains (e.g., IOTA Tangle) allow parallel transaction processing and eliminate miners, potentially scaling to IoT-scale workloads. However, DAG-based systems are less mature and have different security properties (Habibullah et al., 2024).
- Hierarchical Consensus: Different consensus mechanisms at different network tiers. For example, home networks might use lightweight PoA for local logging, while consortium networks use PBFT for cross-home coordination. This matches consensus overhead to trust requirements (Thakur, 2025).
- Checkpointing and Pruning: Periodically archive old blocks to cold storage and prune them from IoT devices, maintaining only recent history and cryptographic proofs of archived data integrity. This bounds storage growth (Singh et al., 2019).

6.4.3 Open Research Directions for Scalable Heterogeneous FL-Blockchain IDS

Adaptive Algorithm Selection. No single FL algorithm works optimally across all device classes and network conditions. Future systems should dynamically select algorithms based on context:

- For battery-powered sensors: Use asynchronous FL with extreme compression (1-bit quantization) and infrequent participation
- For mains-powered hubs: Use synchronous FL with full-precision updates and frequent participation
- For gateways: Participate in blockchain consensus with PBFT or PoA
- For cloud servers: Coordinate global model aggregation and consortium blockchain

Developing context-aware middleware that automatically selects and configures appropriate algorithms remains an open challenge.

Incentive Mechanisms for Participation. In large-scale deployments, why should homeowners contribute their devices' energy and bandwidth to collaborative security? Token-based incentive systems, where devices earn rewards for contributing useful updates, have been proposed but not validated in real deployments (Begum et al., 2024). Key questions include:

- How to measure contribution quality (not just quantity)?
- How to prevent gaming of incentive systems?
- How to value contributions from different device types?

Who funds the reward pool?

Privacy-Preserving Scalability. Many scalability techniques (e.g., sampling, clustering) potentially leak information about participants. For example, if a device is consistently sampled because it reports many anomalies, an observer might infer that home is under attack. Developing scalability techniques that preserve differential privacy guarantees is an important research direction.

Benchmarking and Testbeds. The research community lacks standardized benchmarks for evaluating scalability in heterogeneous IoT environments. Needed resources include:

- Large-scale emulation/simulation frameworks supporting thousands of heterogeneous devices
- Public datasets with realistic non-IID distributions across device types and homes
- Standardized metrics for communication efficiency, energy consumption, and convergence speed
- Long-term deployment testbeds (real homes instrumented over years)

Theoretical Foundations. Current understanding of FL convergence in heterogeneous environments relies on assumptions (e.g., convex loss functions, bounded gradient variance) that rarely hold in practice. Stronger theoretical guarantees for non-convex, heterogeneous, asynchronous settings would guide practical algorithm design.

6.4.4 Summary of Interoperability between Heterogeneous IoT Devices

Interoperability and scalability are intertwined challenges in large-scale heterogeneous smart home deployments. While interoperability ensures diverse devices can collaborate meaningfully, scalability ensures the collaboration remains efficient as the network grows. The techniques discussed—hierarchical architectures, clustering, and asynchronous updates, sampling, compression, and blockchain optimizations—offer promising directions, but significant research remains to validate these approaches in real-world deployments. The ultimate goal is a federated learning framework that gracefully scales from a single home with five devices to a smart city with millions, while maintaining security, privacy, and usability across orders-of-magnitude differences in device capabilities. This requires moving beyond one-size-fits-all solutions toward adaptive, context-aware systems that match algorithms to device classes and network conditions dynamically.

7. CONCLUSION

7.1 Summary of Findings

In this review, we analyzed how to combine federated learning and blockchain technology for intrusion detection in a smart home IoT network. The smart home network, consisting of a variety of heterogeneous, resource-constrained IoT devices, enhances ease of use but also poses potential security risks (Singh et al., 2019). The centralized nature of conventional intrusion detection systems is inadequate for such a scenario, as it creates single points of failure and allows potential breaches of confidentiality. However, federated learning is a decentralized method for model training that neither requires raw data transfer nor violates confidentiality (Rahman et al., 2020). While federated learning enables decentralized model training, it still lacks integrity/participation security, for which blockchain technology is a potential solution. With blockchain technology, a decentralized, immutable ledger, a secure pathway for federated learning is established, providing secure intrusion detection, authentication of model updates, and eliminating intermediate trusting for aggregations (Begum et al., 2024). We explained how different blockchain platforms (toward private/consortium blockchains) and consensus mechanisms (PBFT, PoA, among others) are leveraged to address the latency and power

constraints of IoT-driven IDS systems (Liu et al., 2021). While the combination of FL and blockchain is so potent, it also entails overhead costs; as such, a variety of improvement methods have been proposed. These range from edge computing to delegating complex operations to remote systems, to distillation to reduce the size of FL frameworks to device-executable levels, and to the optimization of smart contract logic for effective automated systems and responses. The literature indicates that the combined system, with the aforementioned optimizations, offers excellent accuracy and security, with tolerable latency and resource overhead (Thakur, 2025; Patel et al., 2022).

7.2 Implications for Future Research

The combination of federated learning and blockchain in smart home security is an area of research at a very early stage, and this review highlights several avenues for further exploration. One of these is a need for a holistic approach whereby the entire system, from the micro-code level in devices through networking and user interface, is taken into consideration by researchers working in this field. Future studies may explore co-designed approaches to IoT devices that are naturally compatible with collaborative security systems (Xu et al., 2020). Furthermore, this review suggests that one avenue for further study is open standards and open access to datasets, enabling this field to advance along these lines. However, this scalability problem implies that models or approaches at different levels (intra-home versus inter-home networking) are given particular emphasis (Liu et al., 2021).

There is, further, clear room for advancement in "lightweight" cryptographic approaches suited explicitly to IoT applications, potentially drawing on insights from sensor networking (Rahman et al., 2020). On a more practical note, success in this area would necessarily carry forward to successes in others such as competent healthcare and Industrial Internet, so it would be a valuable area of exploration for these researchers to see how a model developed here might be applied in these others, addressing their own unique needs and regulatory requirements. A further and important implication is a need for a cross-disciplined approach so that experts in computer security and developers in IoT technology (as well as likely sociologists and lawyers) are brought in so that studies can be appropriately designed for their application at home by private users, who would likely be deploying these aids in a semiautonomous form when attempting to detect threats (which would be a consideration, specifically regarding a federated model of blockchain associated with such a "smarter" home IDS design as this would necessarily imply, regarding a user consent and liability framework associated with such user-deployment at home by semiautonomous forms of these systems designed for threat detection in a semiautonomous form when attempting threat detection through sensors when operating at a "smarter" level of sophistication than is current today).

7.3 Final Remarks

Securing smart homes is becoming an increasingly pressing concern as IoT technology becomes an indispensable part of our daily lives. This paper emphasizes that the symbiotic use of federated learning and blockchain technology offers potential remedies for overcoming the inherent difficulties in smart home security by making it decentralized, reliable, and privacy-friendly. Technologies employed can leverage the strengths of social intelligence (fed. learning) and social trust (blockchain technology) to build IDS systems that are impenetrable against both hostile and internal threats and failures. Though there is indeed feasibility and benefit in adopting such a technology paradigm, translating research into practical use at this stage will involve overcoming specific difficulties and challenges. Encouragingly, however, its innovation track is showing that all obstacles and difficulties are being overcome, and that what is needed is understood and being met by all involved stakeholders.

In closing, it is safe to say that smart home security via federated learning and blockchain technology seems an important and exciting new wave of smart home technology. This technology, despite its progress and improvements, is expected to bring smart homes that are more secure, reliable, and trustworthy, protecting user privacy and freedom, and more effective at repelling attackers and potential thieves. This symbiotic technology of AI and blockchain in smart home security is an important example of using tools from multiple disciplines to develop solutions to complex, pressing cybersecurity issues.

ACKNOWLEDGMENTS

The authors' wish to express gratitude to AlMaarefa University Riyadh, Saudi Arabia, for their unwavering support and encouragement.

Author Contributions

Both the authors are equally contributed to this work.

Funding Statement

This research doesn't receive any fund.

Data availability Statement

The data used for conducting this research work will be available from the corresponding author.

Ethics in Publishing

"Not applicable"

Declaration of Competing Interest

"The authors declare no competing interest exists."

REFERENCES

Al-Turjman, F., & Lemayian, J. P. (2020). Intelligence, security, and vehicular sensor networks in internet of things (IoT)-enabled smart-cities: An overview. *Computers & Electrical Engineering, 87,* 106776. https://doi.org/10.1016/j.compeleceng.2020.106776

Alruwaili, F. F. (2024). Blockchain-powered deep learning for internet of things with cloud-assisted secure smart home networks. *IEEE Access, 12,* 119927-119936.

Begum, K., Mozumder, M. A. I., Joo, M.-I., & Kim, H.-C. (2024). BFLIDS: Blockchain-Driven Federated Learning for Intrusion Detection in IoMT Networks. *Sensors, 24*(14), 4591. https://doi.org/10.3390/s24144591

Berger, C., Toumia, S. B., & Reiser, H. P. (2024). Exploring Scalability of BFT Blockchain Protocols through Network Simulations. *Formal Aspects of Computing, 36*(4), 1-29. https://dl.acm.org/doi/10.1145/3689343

Ghasempour, A. (2019). Internet of Things in Smart Grid: Architecture, Applications, Services, Key Technologies, and Challenges. *Inventions, 4*(1), 22. https://doi.org/10.3390/inventions4010022

Govindaram, A., & Jegatheesan, A. (2025). FLBC-IDS: A federated learning and blockchain-based intrusion detection system for secure IoT environments. *Multimedia Tools and Applications, 84*(17), 17229-17251. https://doi.org/10.1007/s11042-024-19777-6

Habibullah, S. M., Alam, S., Ghosh, S., Dey, A., & De, A. (2024). Blockchain-based energy consumption approaches in IoT. *Scientific Reports, 14*(1), 28088. https://doi.org/10.1038/s41598-024-77792-x

Hai, T., Wang, D., Seetharaman, T., Amelesh, M., Sreejith, P. M., Sharma, V., ... & Liu, H. (2023, February). A novel & innovative blockchain-empowered federated learning approach for secure data sharing in smart city applications. In *International conference on advances in communication technology and computer engineering* (pp. 105-118). Cham: Springer Nature Switzerland.

Idrissi, M. J., Alami, H., El Mahdaouy, A., El Mekki, A., Oualil, S., Yartaoui, Z., & Berrada, I. (2023). Fed-ANIDS: Federated learning for anomaly-based network intrusion detection systems. *Expert Systems with Applications, 234,* 121000. https://doi.org/10.1016/j.eswa.2023.121000

Karunamurthy, A., Vijayan, K., Kshirsagar, P. R., & Tan, K. T. (2025). An optimal federated learning-based intrusion detection for IoT environment. *Scientific Reports, 15*(1), 8696. https://doi.org/10.1038/s41598-025-93501-8

Khraisat, A., & Alazab, A. (2021). A critical review of intrusion detection systems in the Internet of Things: Techniques, deployment strategy, validation strategy, attacks, public datasets and challenges. *Cybersecurity, 4*(1), 18. https://doi.org/10.1186/s42400-021-00077-7

Kumar, M., Ali, T., Anguera, J., & Tripathi, S. L. (Eds.). (2026). *Emerging technologies in AI, computation, communication, and cybersecurity.* CRC Press/Taylor & Francis Group.

Liu, H., Zhang, S., Zhang, P., Zhou, X., Shao, X., Pu, G., & Zhang, Y. (2021). Blockchain and federated learning for collaborative intrusion detection in vehicular edge computing. *IEEE Transactions on Vehicular Technology, 70*(6), 6073-6084. https://ieeexplore.ieee.org/abstract/document/9420262

Patel, C., Bhatt, D., Sharma, U., Patel, R., Pandya, S., Modi, K., ... & Ghayvat, H. (2022). DBGC: Dimension-based generic convolution block for object recognition. *Sensors, 22*(5), 1780.

Rahman, M. A., Zaman, N., Asyhari, A. T., Al-Turjman, F., Bhuiyan, M. Z. A., & Zolkipli, M. F. (2020). Data-driven dynamic clustering framework for mitigating the adverse economic impact of Covid-19 lockdown practices. *Sustainable cities and society, 62*, 102372.

Ren, C., Jiang, B., & Lu, N. (2024). Federated learning with potential partnership identification for accurate prediction in flexible manufacturing system. *IEEE Transactions on Reliability, 74*(2), 2549-2560. https://doi.org/10.1109/TR.2024.3427813

Saleh, F. (2021). Blockchain without waste: Proof-of-stake. *The Review of financial studies, 34*(3), 1156-1190.

Shalan, M., Hasan, M. R., Bai, Y., & Li, J. (2025). Enhancing smart home security: Blockchain-enabled federated learning with knowledge distillation for intrusion detection. *Smart Cities, 8*(1), 35. https://doi.org/10.3390/smartcities8010035

Singh, P. K., Singh, R., Nandi, S. K., & Nandi, S. (2019, June). Managing smart home appliances with proof of authority and blockchain. In *International conference on innovations for community services* (pp. 221-232). Cham: Springer International Publishing.

Swathi, K., Durga, P., Prasad, K. V., Chaitanya, A. K., Santhi, K., Vidyullatha, P., & Rao, S. V. A. (2025). Secure blockchain integrated deep learning framework for federated risk-adaptive and privacy-preserving IoT edge intelligence sets. *Scientific Reports, 15*(1), 41133. https://doi.org/10.1038/s41598-025-24895-8

Thakur, S. S. (2025, March). Federated learning over LEO satellite networks for scalable and secure global IoT connectivity. In *ECCSUBMIT Conferences* (Vol. 3, No. 1, pp. 113-118).

Xu, L., Gao, Z., Fan, X., Chen, L., Kim, H., Suh, T., & Shi, W. (2020, December). Blockchain based end-to-end tracking system for distributed IoT intelligence application security enhancement. In *2020 IEEE 19th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)* (pp. 1028-1035). IEEE.

Zhu, J., Cao, J., Saxena, D., Jiang, S., & Ferradi, H. (2023). Blockchain-empowered federated learning: Challenges, solutions, and future directions. *ACM Computing Surveys, 55*(11), 1-31. https://doi.org/10.1145/3570953