課題研究報告書

IoT ネットワークにおける連合学習を用いた DoS/DDoS 防御に関する調査

Liang Zhenghao

主指導教員 BEURAN Razvan

北陸先端科学技術大学院大学
先端科学技術研究科
(情報科学)

令和 8 年 3 月

Abstract

In recent years, the Internet of Things (IoT) has expanded into a wide range of domains, including smart homes, healthcare and nursing care, agriculture, transportation, and industrial control (IIoT/OT), increasing its importance as a component of social infrastructure. Security issues in IoT networks are not limited to the compromise of information assets; they can also directly affect physical safety and business operations. In this context, DoS/DDoS remains a major threat, and recent industry reports repeatedly highlight both the growing scale and increasing frequency of attacks. Cloudflare's quarterly reports indicate not only an increase in observed large-scale DDoS incidents but also a shortening of peak durations, suggesting stronger requirements for rapid detection and response. Moreover, during Q2–Q3 of 2025, 8.3 million incidents were reported, including extremely large attacks reaching up to 29.7 Tbps and 14.1 Bpps. In addition, in OT/ICS domains, cyber incidents have been noted to lead to operational disruption such as business shutdowns and service outages, and cases have also been reported in which DoS attacks escalated to physical damage.

However, IoT networks operate in distributed settings where edge devices have limited CPU, memory, and power, and communications are often unstable (wireless links, bandwidth limits, variable delays, and loss). Multi-stakeholder operations further induce Non-iid data, making a single centralized model hard to apply uniformly. Under these constraints, DoS/DDoS can reduce availability via bandwidth exhaustion and latency spikes, disrupting monitoring/control. Thus, IoT/OT requires operational feasibility beyond accuracy—sustained detection/mitigation across devices, time bounds, and network conditions—because DDoS-driven congestion can break log collection, inference, and even model updates.

Against this backdrop, DoS/DDoS defense using federated learning (Federated Learning; FL) has attracted attention as a means to protect privacy, preserve data sovereignty, and mitigate data silos. FL enables collaborative learning without sharing raw data and thus helps protect information privacy; however, it depends on communication rounds and is sensitive to the status of participating devices and network conditions. In other words, when applying FL to IoT, the ability for training to "keep running" becomes a fundamental feasibility requirement, and communication degradation caused by DoS/DDoS can undermine FL's basic assumptions. This tension cannot be fully addressed by conventional IDS/classifier-centered evaluation frameworks alone, and a structured analysis is needed that includes behavior under degraded communication, the ability to sustain updates, time requirements such as inference latency and end-to-end (E2E) response, and the feasibility of implementation on low-performance devices.

This survey aims to systematically organize recent research trends and practical

operational challenges in studies that apply FL to DoS/DDoS detection and mitigation in IoT environments. We searched the literature using Google Scholar and ScholarGPT and selected 50 papers according to predefined criteria. The selected studies are concentrated in 2024–2025, confirming rapid growth of the field in recent years. We organized the literature using seven frameworks, including categories based on detector architectures (CNN/DNN/RNN/GNN/AE/Transformer) as well as system/network-mechanism-driven studies that emphasize operational feasibility—such as communication, participation, and trust—rather than detectors alone. We then extracted FL design aspects (e.g., aggregation, participation control, reliability and security) and operationally relevant metrics (inference latency, E2E response time, communication volume, evaluation under congestion/loss conditions, hardware-based measurement, and low-performance device evaluation), and analyzed reporting trends by publication year.

As a result, studies evaluated in computation-only settings (i.e., without confirmed implementations of update transmission/reception involving communication) were the most common(about 70%), indicating that many works do not experimentally examine how DoS/DDoS affects communication and latency. Dataset usage still heavily relies on public datasets from 2015–2019, with 68% of uses tied to datasets from 2019 or earlier. Among the 44 papers that explicitly mentioned IoT deployment, only eight reported real-device measurements (six on low-performance devices), while most relied on PC/server experiments. low-performance devices require model lightweighting, control of update frequency, and reductions in communication rounds due to limited computation and power, DoS/DDoS can impair training continuity through bandwidth exhaustion and increased latency; thus, beyond accuracy, design and evaluation must consider whether "training can continue running." We also found that some papers provide insufficient descriptions of measurement units and measurement points (device/gateway/server) for inference latency, E2E response, and communication volume, raising concerns about reproducibility. In addition, performance variability due to Non-iid data and environmental differences, zero-day attacks, the need to clarify re-training procedures under congestion, and the lack of cross-dataset evaluation were identified as cross-cutting challenges.

Based on these findings, this survey compiles reporting items into a checklist and organizes evaluation conditions and experimental environments in tabular form, with the expectation of providing practical guidance for research design aimed at real-world deployment in IoT environments.

概要

近年，IoT（Internet of Things）はスマートホーム，医療・介護，農業，交通，ならびに産業制御（IIoT/OT）など多様な領域へ広がり，社会インフラとしての重要性を高めている。IoT ネットワークのセキュリティ問題は，情報資産の侵害に留まらず，物理世界の安全や事業へ直接影響し得る点が特徴である。この状況下で DoS/DDoS は依然として主要な脅威であり，近年の業界レポートでは攻撃規模の巨大化と頻発が繰り返し報告されている。Cloudflare の四半期レポートでは，大規模 DDoS の観測件数増加に加え，ピークが短時間化していることが示されており,「早く検知して対応する」要件が一層強いことが示唆される。さらに 2025 年 Q2–Q3 には 830 万件の観測が報告され，最大 29.7Tbps ／ 14.1Bpps といった超大規模攻撃も観測されている。加えて，OT/ICS 領域では，サイバー事案が事業停止や業務中断といった運用上の中断へ接続し得る点が指摘され，DoS 攻撃が物理的破壊へ波及した事例も報告されている。

一方，IoT ネットワークは分散的な環境で継続的にデータを生成・共有する反面，端末側は CPU・メモリ・電力といった計算資源に強い制約を持ち，通信も無線，帯域幅制限，可変遅延，通信損失など不安定性を伴うことが多い。さらに IoT は広域かつ多主体に跨るため，運用主体の違い（管理ポリシー，ログ構成，ネットワーク構成）に起因してデータ分布が不均一（Non-iid）になりやすく，単一の集中学習で一律に対処することが難しい。このような環境において DoS/DDoS は帯域幅枯渇や遅延増大を通じてサービス可用性を低下させ，監視・制御系の応答遅延や誤動作を誘発する可能性がある。したがっIoT/OT では，単に検知精度が高いだけでは不十分であり,「どの設備で」「どの程度の時間で」「どのような通信条件下でも」検知・緩和を継続できるかという運用成立性が重要となる。特に DDoS は通信を攻撃対象とするため，検知器が正常に動作していても，通信混雑や損失によってログ収集，推論，さらにはモデル更新の継続が阻害され，防御が機能不全に陥る可能性がある。

こうした背景の下，プライバシー保護，データ主権，およびデータサイロの解消を目的として，連合学習（Federated Learning; FL）を用いた DoS/DDoS 防御が注目されている。連合学習は生データを共有せずに協調学習を実現でき、情報のプライバシーを保護する一方，通信ラウンドに依存し，参加する端末の状態やネットワーク状態に敏感である。すなわち IoT に連合学習を適用する場合,「学習が回り続けるか」が成立要件となり，DoS/DDoS による通信劣化は連合学習の前提そのものを崩す可能性がある。この矛盾は従来の IDS ／分類器中心の評価枠組みだけでは十分に扱い切れず，通信悪化時の挙動，更新継続能力，推論遅延・E2E 応答など時間要件，および低性能デバイス上での実装可能性を含めた整理が必要である。

本調査は，IoT 環境における DoS/DDoS の検知・緩和へ連合学習 を適用した研究を対象に，近年の研究動向と現実運用上の課題を体系的に整理することを目的とする。Google Scholar および ScholarGPT を用いて文献を探索し，選定基準に基

づき 50 本を選定した。対象文献は 2024〜2025 年に集中しており，当該分野が直近数年で急速に拡大していることが確認された。本調査では，検知器アーキテクチャ（CNN ／ DNN ／ RNN ／ GNN ／ AE ／ Transformer）に基づく分類に加え，検知器そのものよりも通信・参加・信頼などの運用成立性を重視するシステム／ネットワーク機構主導の研究を含め，7 つの枠組みで整理した。その上で，連合学習設計（集約，参加制御，信頼性・安全性等）と，運用に直結する指標（推論遅延，E2E 応答時間，通信量，混雑・損失条件下の評価，ハードウェア実測，低性能デバイス評価）を抽出し，年次別の報告状況を分析した。

その結果，実験環境は計算のみ（通信を伴う更新送受信の実装が確認できない）研究がもっとも多く、約 7 割を占め，DoS/DDoS が通信と遅延へ与える影響を実験的に検証していない事例が多いことが分かった。データセットについては，公開されている 2015〜2019 年代のデータセットへの依存が強く，使用回数ベースで約 68%が 2019 年以前に属していた。さらに，IoT 環境適用を言及した 44 本のうち，実機デバイスでの実測は 8 件（低性能デバイス 6 件＋それ以外 2 件）に留まり，多くの研究は PC ／サーバ上での検証に依存していた。低性能デバイスでは計算資源と電力の制約から，モデル軽量化，更新頻度制御，通信回数削減が不可欠である一方，DoS/DDoS 下では帯域幅枯渇や遅延増大により学習継続性が損なわれ得るため，精度のみならず「学習が回り続けるか」という観点での設計・評価が求められる。また，推論遅延・E2E 応答・通信量の単位や測定点（端末／ゲートウェイ／サーバ）の記載が不十分な文献もあり，再現性の確保が課題であった。加えて，Non-iid や環境差に起因する性能ばらつき，ゼロデイ攻撃，通信混雑時の再学習手順の明確化，およびクロスデータセット評価の不足も横断的課題として確認された。

これらを踏まえ，本調査は報告項目をチェックリストでまとめ，評価条件や実験環境を表整理することで，IoT 環境での現実運用への実装を見据えた研究設計の参考となることを期待する。

目 次

図 目 次

表 目 次

第1章　はじめに

1.1　調査背景

IoT（Internet of Things）および OT（Operational Technology）を含むサイバー・フィジカル領域は、スマートホームから製造業・重要インフラに至るまで急速に普及し、攻撃面が拡大している。とりわけ IoT/OT のセキュリティ問題は、情報資産の侵害に留まらず、物理世界の安全や操業へ直接影響し得る点が特徴である。

この状況下で DoS（Denial of Service）/DDoS（Distributed DoS）は依然として主要な脅威であり、近年の業界レポートでは攻撃規模の巨大化と頻発が繰り返し報告されている。例えば、Cloudflare の四半期レポートでは DDoS の大規模化・短時間化がが進んでいると報告されている。 [1]。また Nokia の脅威レポートでは、DDoS トラフィックが他のトラフィック種別を上回る勢いで増加していること、およびボットネット (botnet) が主要因であることが報告されている [3]。OT 領域では、サイバー事案が操業停止や業務中断といった「運用上の中断」へ接続し得る点がレポートで指摘されている他、DoS 攻撃による IoT 設備の加熱で物理的破壊されるケースも確認されている [5]。

加えて、IoT 機器は攻撃「対象」であると同時に、侵害された機器群がボットネット化して攻撃「基盤」になり得る。Nokia はボットネットが DDoS の大きな駆動要因であることを述べており [3]、家庭内 IoT を対象とした大規模観測でも短期間に膨大なセキュリティイベントが起こり得ることが示されている [4]。

その上、IoT 機器の多くは計算資源やセキュリティ機能が限定的であるため、マルウェア感染によるボットネット化になりやすい。このデバイスの侵害は単なる個別の端末問題に留まらず、大規模な DDoS トラフィックを生成する攻撃の起点へと変貌することを意味する。したがって、IoT における DDoS 防御は、ネットワーク境界での遮断だけでなく、分散した端末側でボットネット特有の異常な振る舞いを早期に検知する枠組みが不可欠である。

一方、IoT/OT の現場では、端末計算資源の制約、低遅延要求、通信帯域・品質の不確実性、プライバシーの制約、データ分布の偏りなどが同時に存在し、精度中心の議論だけでは実運用へ繋がりにくい。この点で、ネットワークのエッジでの処理は遅延・帯域制約の緩和の枠組みとして位置づけられ [7]、GDPR みたいな個人情報保護政策や JC-STAR のような IoT セキュリティラベリング政策により、IoT ではプライバシーとセキュリティの要請が強いことも整理されている [8]。

こうした背景から、端末側データを外部へ集約せずに学習を進める連合学習（Fed-

表 1.1: 業界レポートにみる IoT/OT 領域の脅威規模

出典	年次	注目すべき観測値・報告内容
Cloudflare（DDoS レポート）	2025 年	2025 年 Q2–Q3 に 830 万件を観測、最大 29.7 Tbps/14.1 Bpps。ピークの短時間化も報告→“早く検知して対応する ” 要件が強い [1]。
Nokia（携帯、IoT 脅威レポート）	2024 年	DDoS トラフィックの増加やボットネット関与が報告される [3]。
Bitdefender（家庭内 IoT 観測）	2024 年	家庭内 IoT を含む環境で、多数のセキュリティイベントが発生し得ることが示される [4]。
Dragos（OT 年次レポート）	2025 年	OT/ICS における脅威動向と、運用中断リスクの顕在化と DoS 攻撃による IoT 設備の物理的破壊が指摘される [5]。
VicOne（自動車セキュリティ）	2025 年	車両・サプライチェーン等を含む自動車領域の脅威とリスクが整理され、EV の充電設備による DoS 攻撃観測される。 [6]。

erated Learning; FL）は、IoT/OT のような分散環境で検知モデルを更新するための有力候補として検討されている。しかし現状の文献は、モデル精度を中心に議論するが、通信を伴う学習更新の実装、推論時間・応答時間、通信量、輻輳・損失条件、ハードウェア・実験環境といった実運用に直結する観点の報告が不十分な場合が多い。精度が高くても、更新が通信制約下で回らない、あるいは対処までの時間が攻撃の時間スケールに間に合わなければ、実運用上の有効性は担保できない。そこで本調査は、IoT 環境下の DoS/DDoS 検知に関する FL 研究を、検知器とシステム観点（通信・遅延・環境条件）から整理し、運用へ繋がるための論点を明確化する調査として位置づける。

1.2 調査目的と調査方針

本調査の目的は、新規アルゴリズムの提案そのものではなく、IoT 環境下の DoS/DDoS 検知における連合学習適用研究を対象とした調査である。近年の業界観測では DDoS の大規模化・短時間化が示され、またボットネットが主要因であることも報告されている [2]。この状況では、高精度だけでなく、即応性（速く検知し対処が間に合うか）、攻撃・帯域制約下でも運用として成立するかが重要となる。

しかしながら、既存文献はモデル精度（Acc/AUC/F1 等）の報告が中心となりやすく、実運用に直結する前提条件や評価指標（遅延・通信・ネットワーク悪化条件・環境条件）が十分に揃っていない可能性がある。そこで本調査は、DoS/DDoS 下の実運用に直結する調査軸として、以下の五つの指標を用いて文献を整理する：

1. 推論遅延
2. E2E 応答時間（検知から対処まで）
3. 通信量（学習更新に伴う送受信量）
4. ネットワーク条件（帯域制限・遅延付与・損失等の劣化条件）
5. 実験環境・実装（実機性、およびモデル更新の送受信が実験として確認できるか）

（1）（2）は「速く検知できるか／対処が間に合うか」を左右し、（3）（4）は分散環境における学習・運用の成立性を規定する。また（5）は、評価結果の解釈可能性を担保する前提条件であり、同じ連合学習と言っても、実験としてモデル更新の送受信が行われていない場合、実運用へ繋ぐ議論が成立しにくい。なお、上記指標の操作的定義および「報告あり／なし」の判定・記録規則は第 3 章 3.6 節に集約する。

本調査の目的は、「IoT における DoS/DDoS 検知連合学習研究が、どの程度現実ネットワークへ展開可能な形で整理・評価されているか」を明確化し、さらに「主張を裏付けるために不足している観点は何か」を体系的に示すことである。具体的には、次の問いに答えることを目標とする：

(Q1) 研究動向：IoT 環境の DoS/DDoS 検知に連合学習を用いる研究は、どのような検知器系統・設計方針で発展してきたか。

(Q2) 貢献の整理：各研究は DoS/DDoS 検知に対し、精度向上・データ分布の偏りへの対応・ロバスト性・軽量化など、どのような貢献を主張しているか。

(Q3) 現実展開可能性：現実ネットワークでの運用を想定したとき、評価はどこまで行われているか（例：実機性、通信の実装、遅延、通信量、輻輳・損失条件）。

(Q4) 不足点：Q2 の主張に対して、運用評価として何が欠けているか、又は何が測られていないか。

本調査では、上記の問いに答えるため、対象文献を検知器の系統、FL 設計、運用評価（遅延・通信・条件・環境）の観点から整理し、文献間で比較可能な形に再構成する。

1.3 対象範囲と用語に関する補足

本調査の対象は、IoT ともっと特化した IIoT/OT/IoV 等の分野も含む環境の論文を調査する。DoS/DDoS を検知対象とする連合学習関連研究である。用語や技術背景（IoT/OT の特性、DoS/DDoS、IDS、連合学習の基本、攻撃耐性等）については、第 2 章で体系的に整理するため、本章では定義の列挙は最小限に留める。

1.4　本調査で分かったこと

本調査は 50 本の論文を読み、まとめ、以下の発見があった：

1. 2023 年と比べて、2024 年や 2025 年の連合学習の論文は比較的に見つけやすく、この分野での研究が増加しつつある。そして、論文で重視する課題も多様化になり、発展が進んでいる。

2. 連合学習の論文は、まだ正確度を中心に討論しているケースが多い。連合学習を現実で実用化するためには、それなりの指標を計測する必要があるが、その割合はまだまだ少ない。

3. 連合学習の実験は、まだ計算のみのものが多く、ネットワークをシミュレーションして実験するケースが少ない。つまり、現在の研究は、通信状況を考慮せずに行っているものが多数存在する。

4. データセットが比較的に古い問題が存在する。サイバー攻撃が多様化、膨大化、複雑化になっている今では、2018 年以前のデータセットは現実運用まで支えきれない可能性が高い。

1.5　本調査の構成

本調査の構成は以下の通りである：

- 第 2 章では IoT/OT、DoS/DDoS、IDS、および連合学習の前提知識を整理する。

- 第 3 章では文献検索および選定基準、記録項目、分類方法および判定基準を定義する。

- 第 4 章では選出した文献の説明を行う。使用するデータセットのまとめや論文の概要を述べる。

- 第 5 章では現実運用を展開する可能性の観点から不足点・課題を整理する。

- 第 6 章で結論をまとめる。

第2章　関連技術の整理

2.1　IoT ネットワーク

IoT（Internet of Things）とは、センサやアクチュエータ等の物理デバイスがネットワークを介して接続され、データ収集・監視・制御を行う分散システムである。そのネットワークアーキテクチャは、大きく分けて、ゲートウェイが集約を行うゲートウェイ型と、デバイスがセルラー回線等を用いて直接インターネットへ接続する直接接続型（Cellular IoT/NB-IoT）の 2 つに分かれる。

2.1.1　IoT ネットワークアーキテクチャ

ゲートウェイ型

図 2.1 の示す通り、3 層構造としてモデル化される。

1. **感知層（Perception Layer）**：センサー、RFID やアクチュエータなどの物理デバイスで構成され、物理世界からのデータ収集と識別を担う。これらのデバイスは計算資源が極めて限定的であるため、ボットネット化の標的となりやすく、感知層における最初のリソース枯渇型 DDoS 攻撃の起点となりやすい。また、特化された DoS 攻撃により、物理的破壊されるリスクも存在する。

2. **ネットワーク層（Network Layer）**：認識層で取得したデータをゲートウェイやルーターを介してインターネットへ伝送する役割を持つ。DDoS 攻撃が起こるときに、先に帯域枯渇されやすい。また、感知層より性能が高い傾向を持つことにより、IDS を設置されるケースもある。

3. **アプリケーション層（Application Layer）**：スマートシティやヘルスケアなどの具体的なサービスを提供し、クラウドと連携してグローバルな管理を行う。この層ではデータの高度な処理が行われる一方で、アプリケーション固有の脆弱性を突く DDoS 攻撃が直接的な打撃を与える可能性が存在する。

こうした階層構造において、特にエッジに近い認識層やネットワーク層のデバイスは、CPU、メモリ、電力供給において厳しい制約を抱えている。したがって、

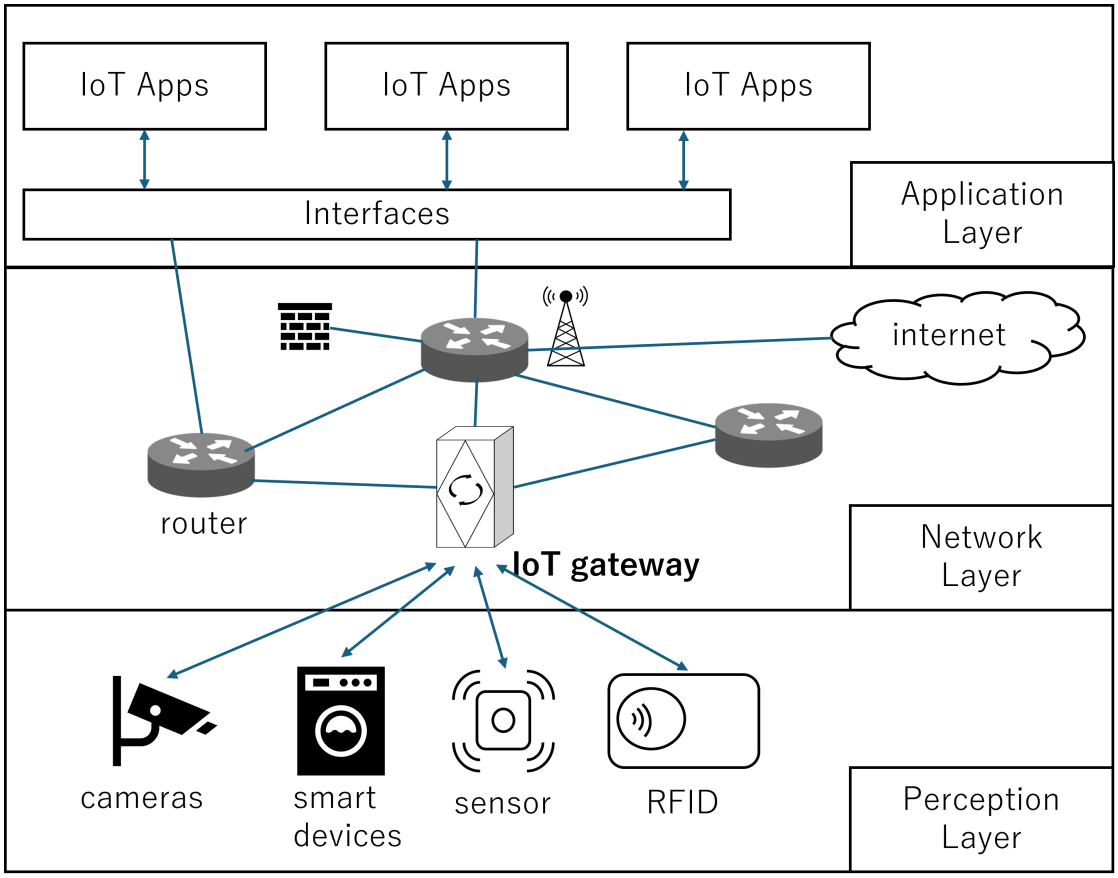

図 2.1: ゲートウェイ型のアーキテクチャ [9]

本調査で扱う連合学習の適用を検討する際には、単なる検知精度のみならず、低性能デバイスも許容できる推論遅延や通信オーバーヘッドの評価が不可欠な評価軸となる。

SD-IoT: SDN（Software Defined Networking）とはルーターやスイッチといったネットワークを構成する機器をソフトウェアで仮想化、一括して制御するシステムである。SDN を IoT ネットワークへ組み込みすることにより、Bluetooth や ZigBee といった異種デバイスが混在する IoT 環境においても、インテリジェントな経路選択やキャッシュ技術を活用した柔軟なトラフィック管理が可能となる。

IoT で SDN を使用する時、ネットワーク層とアプリケーション層の間に、コントロールプレン（control plane）層が設置される。この場合、インターネット層の設備は自主的にトラフィックを管理することでなく、SDN サーバーからの指示を従うこととなる。つまり、SDN サーバーは DoS/DDoS 攻撃の対象となりうる。ネットワーク層における管理の柔軟性を向上させる SDN の重要性を踏まえ、SDN アーキテクチャに関連するクラウドやサーバへ設置する IDS に関する研究も、本調査の対象範囲の対象とする。

直接接続型

直接接続型アーキテクチャとは、IoT 端末がローカルのゲートウェイを介さず、携帯電話ネットワークのベースステーションを通じて直接コアネットワークへ接続する構成を指す。図に示すように、各 NB-IoT 端末はベース（NB-IoT base station）を経由して NB-IoT コアネットワークへ接続され、その上位にクラウドプラットフォーム、さらに業種別の Vertical Industry Center が位置する階層構造を形成している。

この構成は、短距離通信をゲートウェイで集約する従来型 IoT アーキテクチャとは異なり、端末自身が広域通信機能を有し、ネットワーク上の独立したエンドポイントとして振る舞う点に特徴がある。代表的な実装技術としては、3rd Generation Partnership Project により標準化された NB-IoT が挙げられる。NB-IoT は、既存

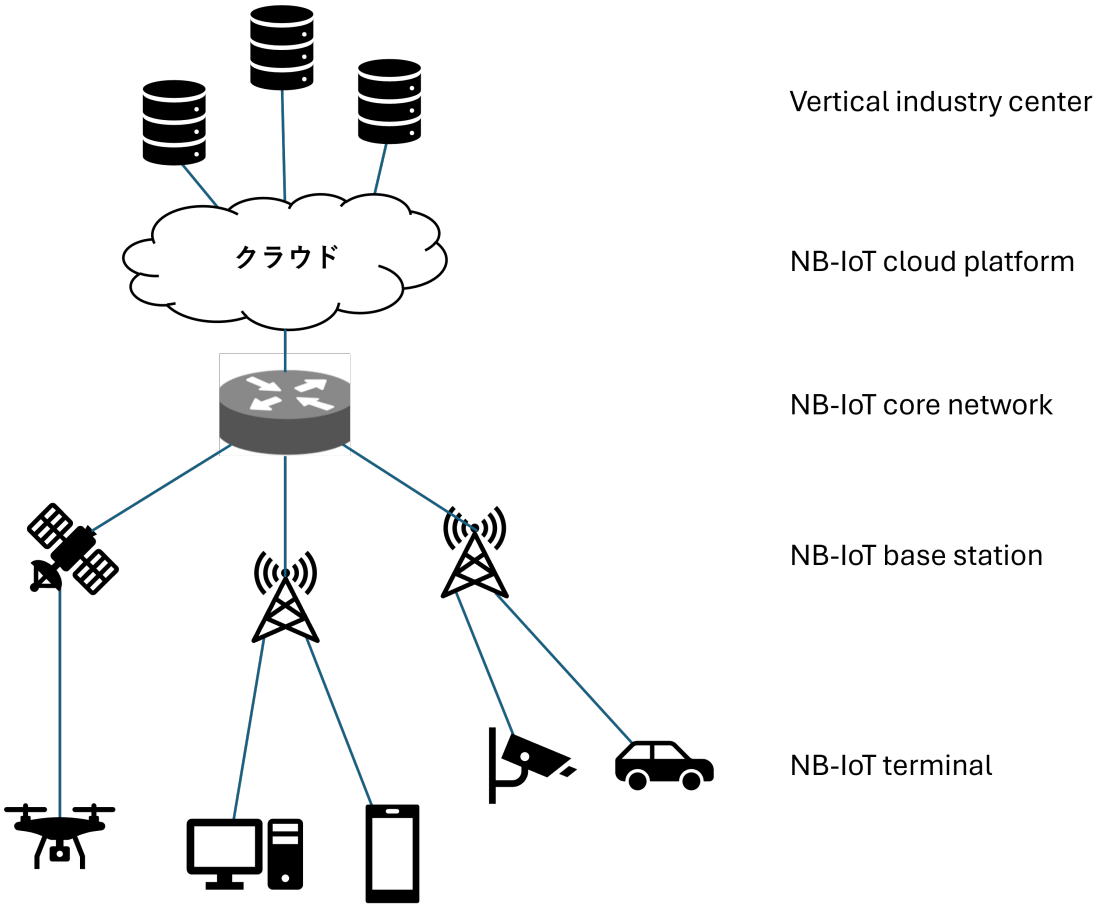

図 2.2: 直接接続型のアーキテクチャ [10]

の Long Term Evolution（LTE）インフラを活用する技術であり、広域カバレッジと高い屋内浸透性を両立している。特に重要なのは、小容量データの送信に際して、通常必要となる専用のデータ通信セッションを確立せずに通信を完結できる

設計である。すなわち、端末はネットワークとの制御手続きの過程で、追加の通信経路を生成することなくデータを伝送できる。この方式により、接続確立やセッション維持に伴う信号交換を大幅に削減できるため、通信オーバーヘッドが抑制される。これは、IoT 端末が周期的に少量のデータのみを送信するというトラフィック特性に適合しており、無駄な接続維持処理を回避することで消費電力の低減にも寄与する。

すなわち、この構成は常時接続状態を維持することを前提とする従来のモバイル通信モデルとは異なり、必要時のみ最小限の手続きで通信を完結させるという思想に基づいて設計されている点を重視する。

図 2.2 の示した通り、各端末はベースステーションと直接無線接続され、コアネットワークに個別に収容される。すなわち、端末は PC やスマートフォンと同様にネットワーク上の独立ノードとして存在し、論理的には個別の IP コンテキストを持つ構造となる。その結果、ゲートウェイで通信を一括制御する境界型セキュリティモデルは適用しにくい。

このような直接接続構造は、セキュリティ設計の前提条件そのものを変化させる。従来のゲートウェイ型アーキテクチャでは、ローカルネットワークの入口に防御機構を集中配置することで、通信の監視や制御を一元的に実施することが可能であった。しかし、直接接続型では各端末が広域ネットワークに個別に接続されるため、単一の境界点に防御機構を集約することが困難となる。結果として、防御境界はネットワーク外周ではなく、各端末へと分散する構造となる。

この構造下では、認証、通信の保護、鍵管理、ソフトウェアの完全性確保といった機能を端末側で行う必要がある。すなわち、セキュリティ責任はネットワーク装置からエンドデバイスへと移行し、各端末が自律的に信頼性を維持することが前提となる。

このように、直接接続型アーキテクチャは通信効率と広域展開性を実現する一方で、防御機能の集中管理を困難にし、セキュリティ運用を構造的に複雑化させるという課題が存在する。

JC-STAR　近年，IoT 機器の急速な普及に伴い，機器レベルでのセキュリティ確保の重要性が高まっている。このような背景の下，日本では IoT 機器のセキュリティ対策を評価する制度として JC-STAR（Japan Cyber-Security Technical Assessment Requirements） が整備されている。JC-STAR は，IoT 機器に対するセキュリティ要件を定義し，製品の安全性を評価及び可視化することを目的とした制度であり，主に機器の設計段階や出荷段階におけるセキュリティ対策を対象としている。具体的には，認証，ソフトウェア更新機能，脆弱性管理などの基本的なセキュリティ機能の実装状況に基づいて複数のレベルが定義されており，IoT 機器の利用者や導入者がセキュリティレベルを判断するための指標として機能する。

一方で，JC-STAR が対象とするのは主として機器自体のセキュリティ要件であり，運用段階におけるネットワークトラフィックの監視や攻撃検知の仕組みまでは

直接的に扱わない。特に，IoT ネットワークでは DoS/DDoS のような大規模なネットワーク攻撃が発生する可能性があり，これらに対する防御にはネットワークレベルでの異常検知技術が重要となる。近年では，プライバシー保護や分散環境への適応性の観点から連合学習を利用したトラフィック分析や侵入検知の研究が注目されている。

表 2.1: JC-STAR の階級区分

星級	評価方式	定義・位置付け
★1	自己適合宣言	製品類型共通の最低限のセキュリティ技術要件を満たすことを事業者自らが宣言するレベル。基本的な認証機能、通信の保護、ソフトウェア更新機構、データ保護等の基礎的対策を対象とする。
★2	自己適合宣言	★1 を基礎とし、より高度な要件への移行を想定した中間的区分として位置付けられる。将来的な高水準評価への橋渡しを目的とする。
★3	第三者評価	独立した評価機関による客観的な適合確認を伴うレベル。設計・実装の妥当性が外部検証され、より高い信頼性が求められる。
★4	第三者高度評価	高度なセキュリティ保証を目的とする最上位レベル。高リスク用途や重要インフラ分野への適用を想定し、厳格な評価手続が実施される。

2.1.2 OT/ICS ネットワーク

OT（Operational Technology）/ICS（Industrial Control Systems）は、IoT ネットワークの一種であるが、工場・発電・輸送などの物理プロセスを制御する領域であり、IT システム以上に可用性（止めないこと）と安全性（人身・設備被害を防ぐこと）が重視される。このため、パッチ適用や機器更新は計画停止・検証・認証を要し、「更新して解消する」というの発想がそのまま適用できない場合が多い。通常の IoT ネットワークと比べて、脆弱性管理が難しく、攻撃耐性のばらつきが存在する。

加えて OT/ICS では、可視性（visibility）のギャップが課題となりやすい。OT/ICS では資産把握・トラフィック監視が不十分な状態、侵入元の見落としなどの状況は珍しくない。さらに、制御プロトコルや独自仕様が多いことにより、通常の監視・防御手段の適用が難しい場面もある。その上、便宜を図るために、IT ネットワークに設備をそのまま暴露しているケースも少なくはない。そのため、DoS/DDoS 攻撃による影響は、通常の IoT ネットワークよりも被害が大きい可能性が存在する。

2.1.3 Edge Computing

Edge Computing [7] は、データ発生地点に近いエッジ（端末・ゲートウェイ・サーバ）で計算処理を行い、クラウド往復通信を最小限に抑える考え方である。これにより通信遅延を抑制し、帯域消費を削減しつつ、プライバシ保護を図ることが可能である。IoT/OT ではリアルタイム性・通信不安定性が顕著であるため、エッジ側での人工知能による推論・検知が求められている。

しかしながら、エッジで AI を扱うことは「計算・電力・メモリ制約の中でモデルを動かす」ことを意味する。したがって、モデル軽量化や特徴量設計だけでなく、学習形態（集中学習か分散学習か）を含めたアーキテクチャ選択が必要となる。連合学習は、分散データを活用しつつ生データ共有を避ける学習方式として、Edge Computing で注目を集めている。

2.2 DoS/DDoS 攻撃と検知

2.2.1 DoS/DDoS

DoS（Denial of Service）/DDoS（Distributed DoS）は、標的サービスの可用性を低下させる攻撃であり、帯域・処理能力・状態管理資源を枯渇させるのが目的とする。

図 2.3 の示した通り、DoS 攻撃は単一の設備から被害者へ大量なトラフィックを送ることで、相手の設備の性能を低下させることができる。このような攻撃は、ip アドレスを制限すれば、防ぐことが可能となる。しかしながら、最近では、IoT 端末まで DoS 攻撃を行い、発熱による破壊やメモリの破壊するケースもあるため、そのリスクも過小評価するべきではない。

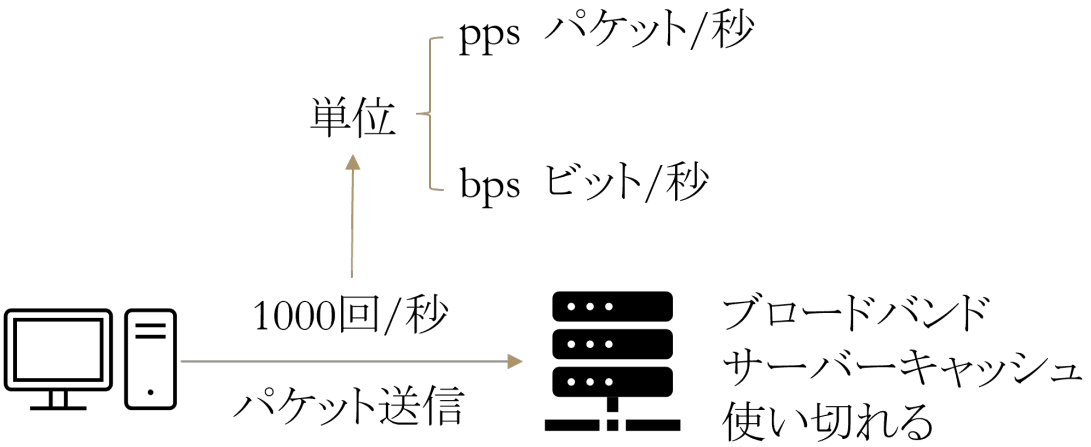

図 2.3: DoS 攻撃

図 2.4 の示した通り、DDoS は複数の攻撃源から同時にトラフィックを送出する点で DoS より強力であり、検知・緩和の難易度が高い。

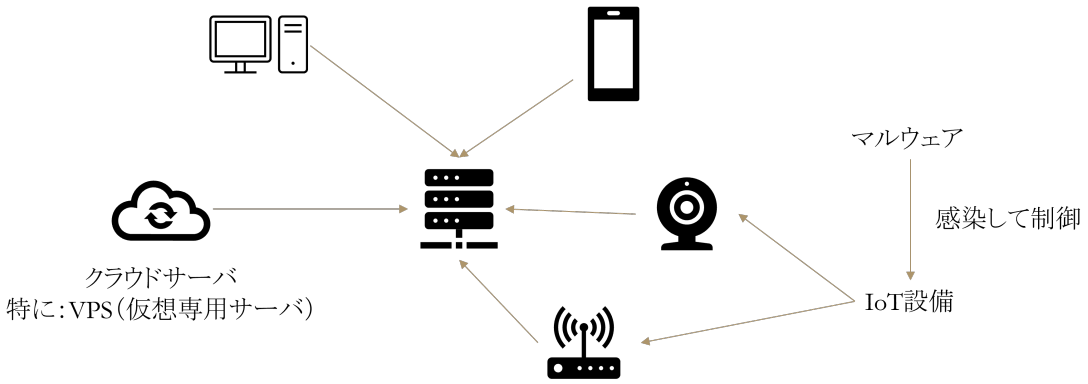

図 2.4: DDoS 攻撃

2.2.2 IDS

IDS（Intrusion Detection System） [12] は、ネットワークトラフィックやシステム挙動を監視し、通常から逸脱する行為、または既知の攻撃パターンを検知する仕組みである。検知方式は代表的に特徴検出 (Signature-based) 型と異常検出 (Anomaly-based) 型に整理できる：

- **特徴検出型**：既知攻撃に強く高速である一方、未知攻撃には弱く、データベースの更新が必要である。
- **異常検出型**：未知攻撃にも対応し得る一方、誤検知や性能コストの課題を抱えやすい。

また、配置の観点からは以下の通り分類される：

- **ホストベース (HIDS)**: 個別の端末内で動作し、暗号化された通信内容や内部攻撃の検知に優れるが、各ホストの資源を消費し、管理負荷が高い 。
- **ネットワークベース (NIDS)**: 境界ゲートウェイ等に配置してネットワーク全体を監視し、脅威の拡散を早期に防ぐ 。ただし、高速ネットワーク下での全パケット解析や、暗号化されたトラフィックの精査には限界がある 。

性能制約の厳しい IoT/OT 環境（ICS）では、端末側での高度な HIDS が困難な場合がよくある。そこで、ゲートウェイやエッジ側に検知機能を統合し、NIDS やファイアウォールを組み合わせた多層防御を構築する設計をするケースが多い。

2.2.3 「検知が速い」ことと「対応が速い」ことの違い

DoS/DDoS の実運用では、検知器の推論が高速であっても、緩和策（遮断・レート制限・経路制御等）が有効化されるまでに遅延があれば被害は継続する。すな

わち「推論遅延」と「エンドツーエンド（E2E）応答時間」は同一ではない。E2E応答には、観測・集計・推論・意思決定・制御適用までの処理と通信が含まれ、配置（端末/ゲートウェイ/エッジ/クラウド）やネットワーク条件に強く依存する。

この観点は、IoT/OT のように通信が不安定で更新が困難な環境において、実装可能性を左右する要因となる。従って、精度のみの比較では不十分となり、遅延や通信条件を含む評価の有無が重要となる。

2.3 連合学習 (Federated Learning)

2.3.1 基本フローと通信ラウンド

連合学習は、複数デバイスまたは組織が生データを直接共有せず、各拠点で学習したモデル更新（モデルパラメータや重み）を集約することで協調的にモデルを学習する方式である。このような方法を用いて、データがネットワークへ広がらないことにより、プライバシーの保護を図る。また、生データの送信をモデルデータへ切り替えることで、通信負担を減らす目的をもつ。その流れは図 2.5 の示す通り：

1. 中央サーバがグローバルモデルを配布する。
2. エッジがローカルデータで学習する。
3. ローカルモデルの更新データまたはモデルデータを中央サーバへアップロードする。
4. 中央サーバが各エッジサーバからのデータを統合して、グローバルモデルを更新かつ配布をする。

こののように繰り返して送信する流れが連合学習の思想である。また、一回の流れは通信ラウンドとして扱われ、ラウンド数とラウンド当たりの通信量が運用コストに直結する。また、モデルを統合する方法としては、FedAvg が一番典型的で、多くの論文で使用または改造をされている。

FedAvg： 中央サーバが各エッジサーバから集めたローカルモデルの更新値を全部まとめた上で、平均値でとる、その平均値をグローバルモデルの更新に使用する。また、更新したグローバルモデルを各エッジサーバへ再度配布するアルゴリズムである。

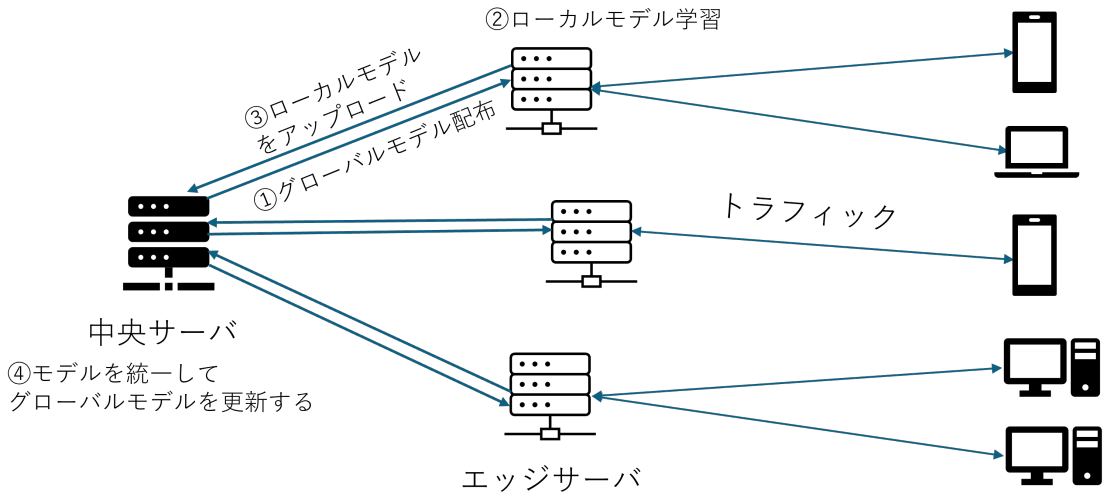

図 2.5: 連合学習

2.3.2 Non-iid

連合学習における代表的難点は、エッジの間のでデータ分布が異なる非独立同分布（Non-iid）問題である。IoT/OT では設置環境・機器種別・利用状況が多様であり、分布の偏りが生じやすい。Non-iid 下では収束不安定や精度低下が起こり得るため、分布偏りを緩和する集約設計や、クライアント毎に適応させる個人化の発想が導入される。

また連合学習は学習過程への介入により、ポイズニング・バックドア等という独自の脅威も伴う。信頼できないエッジ設備の混入が想定される環境では、異常更新の検出やロバスト集約、信頼管理の問題は現実運用へ直結する。

2.3.3 IoT/OT での現実運用上の論点

IoT/OT で連合学習を現実運用へ近づけるには、以下の制約を同時に満たす必要がある：

1. 通信量（アップロード/ダウンロードの総データ量）
2. 通信遅延と学習完了時間
3. エッジ設備の参加率や切断、及び信頼性
4. 端末側の性能、電力の制約

特に OT/ICS では導入・更新に制約があるため、「モデル精度が高い」だけでは採用判断に至りにくい。従って、推論遅延、通信量、通信混雑・損失条件下の挙動、ハードウェア条件の明示など、現実で運用することに直結する観点が重要となる。

2.4　本章のまとめ

本章では、IoT/OT/Edge、DoS/DDoS と IDS、連合学習の基礎概念を、本調査の調査目的に必要な範囲で整理した。次章では、文献収集・選定、抽出項目、および整理方法をまとめる。

第3章　調査方法

3.1　本調査の狙いと方法論

本調査は、連合学習を用いたIoT/OT/Edge環境におけるDoS/DDoS検知研究を対象とする調査研究である。調査の焦点は、提案手法の精度比較に留まらず、実ネットワークで運用を含めて整理・評価する点にある。具体的には、通信の成立性（モデル更新の送受信が実験として実行されたか）、遅延、通信量、混雑・損失条件下での挙動、およびハードウェア条件の明示といった観点を重視する。

この目的のため、本章では以下の項目を中心にまとめる：

1. 文献の収集方針
2. 文献ごとに記録する項目
3. 分類方法
4. 実験環境の判定

3.2　文献検索

3.2.1　検索対象と検索ツール

文献検索はGoogle Scholarを一次の探索基盤として用い、補助的にScholarGPTを併用して行う。Google Scholarではキーワード検索に加え、被引用文献・関連文献の追跡を行い、ScholarGPTではキーワード検索・同義語の展開を補助し、検索漏れの低減を図る。

本調査では、検索対象を正式に出版された学術文献に限定する。具体的には、学術誌・ジャーナル論文、および査読付き国際会議の論文集に正式収録された論文を対象とし、プレプリントは積極的に採用しない方針である。ただし、プレプリント由来の記述であっても、最終的に正式出版版が確認できる場合は、正式版を参照対象とする。

検索対象期間は近年の連合学習適用研究の増加を踏まえ、主に2019年以降を中心に設定した。

3.2.2 検索キーワードとクエリ設計

検索キーワードは、攻撃・タスク、学習枠組み、適用領域の 3 ブロックで構成する。攻撃・タスクは「DoS/DDoS」「intrusion」「detection」をメインで、学習のほうは「federated learning」を必須要素とする。適用領域は IoT を中核に、IoV、Smarthome、IIoT、CPS 等の表記を含める。主に使用するキーワードは表 3.1 に示す。

表 3.1: 検索キーワード

ブロック	キーワード例
攻撃・タスク	DDoS / DoS / denial of service、intrusion、detection / IDS / intrusion detection
学習枠組み	federated learning / federated / FL
適用領域	IoT / IIoT / industrial IoT、IoV / V2X、Smarthome、CPS / cyber-physical systems

3.2.3 論文の絞り込み

絞り込みは段階的に実施する。まずタイトル・アブストラクトで連合学習と対象領域（IoT/OT/Edge 関連）および DoS/DDoS 関連性の有無を確認し、次に本文を精読して評価の有無と必要情報（検知器、学習設定、データセット、指標）が読み取れるかを確認する。この段階で、ジャーナル掲載または会議論文集への正式収録であることも併せて確認し、正式収録が確認できない文献は一次研究として除外する。その後、論文の引用文献・関連文献を追跡し、同様の基準で追加収集を行う。サーベイ論文・業界レポートは一次研究の比較対象としては含めず、背景・用語整理・動向補助として利用する。

3.3 選定基準

3.3.1 基準

以下を満たす文献を対象に含める：

1. 出版形態：学術誌掲載論文、または査読付き国際会議の論文集に正式収録された論文である。

2. 連合学習である：連合学習に基づく学習・集約プロセスが記述されている。

3. IoT である：IoT/IIoT/IoV/Smart Home/CPS 等の文脈で検知を扱う、または想定環境がそれに該当する。

4. DoS/DDoS と関わる:DoS/DDoS 検知が主題、または評価対象として DoS/DDoS（関連攻撃）が含まれる。

5. 実験評価：データセットと評価指標に基づく実験結果が提示されている。

3.3.2 除外基準

以下に該当する文献は除外する：

1. 正式出版版が確認できない文献（プレプリントのみ等）。

2. 集中学習のみで連合学習を用いない文献。

3. DoS/DDoS との関連が確認できない文献。

4. 概念の提案のみで実験がない、または評価条件・指標が提示されない文献。

5. サーベイ論文（一次研究比較対象からは除外し、背景として利用する）。

3.4 記録項目

本節では、各文献から記録する項目を定義する。調査研究として第 4 章・第 5 章で傾向分析や不足指標の議論を行うためには、「何を記録したか」を先に固定しておく必要がある。本調査では、検知器の特徴と実際に運用に関わる情報を中心に、以下に示す表 3.2 の内容を記録する。

表 3.2: 文献から記録する主要項目

カテゴリ	記録内容（例）
基本情報	年、論文名（または著者年表記）、掲載形態（ジャーナル／会議）
対象・適用先	対象シーン（ネットワーク IDS、IoT エッジ等）、検知の配置（端末／ゲートウェイ／エッジ等）
データセット	使用データセット名、DoS/DDoS が評価対象に含まれるか、IoT 特化データか否か
検知器（detector）	モデル系統（例:CNN/RNN/GNN/Autoencoder/Transformer 等）、特徴抽出の方法など
連合学習設定	集約方法（例：FedAvg 等）、ラウンドや参加数に関する記述、Non-iid への言及
セキュリティ・プライバシ関連	プライバシ保護の仕組み（例：暗号・秘匿化等）や学習過程の攻撃耐性（ポイズニング・バックドア等）への言及
通信最適化	圧縮・量子化・スパース化等、通信量削減に関する仕組みの有無
評価結果	精度系指標（Accuracy/F1/AUC 等）と主要結果、評価条件の要点
運用に関わる指標	推論遅延、E2E 応答、通信量、混雑・損失条件下の評価、ハードウェア条件
環境判定の根拠	通信実行（モデル更新の送受信）の有無、実機利用の有無、判定に用いた証拠の要約

3.5 分類

3.5.1 基本方針：検知器中心の分類

本調査の分類は、原則として検知器のモデル系統を主軸とする。これは、DoS/DDoS 検知において、特徴抽出と判定方法の違いが性能・計算量・適用先に直結しやすいためである。

3.5.2 例外方針：連合学習のアルゴリズム・運用機構が主貢献の文献の扱い

文献の貢献が検知器自体ではなく、連合学習のアルゴリズム・運用機構（参加制御、信頼制御、帯域制御、同期/非同期、堅牢集約など）に主として置かれる場合

がある。この類の文献は、検知器の種類よりもシステム全体的に対しての議論を中心とするため、独立カテゴリ（システム／ネットワーク機構主導）として扱う。

ただし、MLP 系検知器を基盤とする文献については、連合学習のアルゴリズムの改善をメインとする論文が多いため、一つの特徴として扱うと考える。そのため、検知器中心の整理の連続性を優先し、原則として MLP 系に残す。したがって、分類規則は次の優先順位で適用する：

1. 主貢献が連合学習のアルゴリズム・構造であり、かつ検知器が MLP 以外である場合：システム／ネットワーク機構主導に分類する。

2. それ以外の場合：検知器のモデル系統に基づき分類する。

3.6 運用に関わる指標

現実での運用に直結する指標について、本節では文献間で比較可能な形に整理するため、操作的定義と記録基準を定める。

以降の集計は本節の規則に従い、数値としての報告有無、条件（計測範囲・ネットワークの悪化条件等）の明示有無を区別して記録する。なお、通信の実行有無に基づく環境判定（simulation/computation 等）は 3.7 節の規則に従う：

- **推論遅延：**検知器が入力を処理して出力を得るまでの時間である。測定対象（端末／ゲートウェイ／サーバ）、入力条件、単位（ms/s）を可能な限り併記する。

- **E2E 応答（End-to-End response）：**観測（データ取得）から検知・判断・制御反映（遮断/緩和）までを含む応答時間である。推論遅延と混同しないため、論文が「応答」に含める範囲（通信・集計・制御適用の有無）を読み取る。

- **通信量**：学習過程で送受信されるモデル更新に関するデータ量、またはその評価である。1 ラウンド当たりの送受信量、総量、圧縮の有無などを読み取る。

- **混雑・損失条件下の評価**：帯域制限、遅延付与、パケット損失、ジッタ等の条件下で評価したかを指す。条件値が提示される場合はそれを記録する。

- **ハードウェア条件**：推論・学習がどの計算基盤で実行されたかを示す。実機端末、エッジサーバ、GPU サーバ、仮想環境などを可能な限り具体に記録する。また、低性能デバイスでの実験の有無についても併せて記録する。

3.7　実験環境の判定基準

3.7.1　判定基準：連合学習構造＋通信動作

本調査の環境判定は、モデル更新の送受信が実験として実行されたかを核とする。判定は以下の条件を満たすかで行う：

- 連合学習構造：クライアント/サーバ（または参加側/集約側）、学習ラウンドまたは集約設備が存在し、データが複数のクライアントに分配される（少なくとも実験設定上そうである）。
- 通信動作：「モデル更新伝送」の証拠が存在する（アルゴリズム説明のみでは不可）、そして以下の条件を満たす：
  - 連合学習のフレームワークを用い、実験実行に使用した記述があり、または通信/ミドルウェア（gRPC、HTTP、MQTT、RabbitMQ、socket、REST API 等）を用いた記述がある。
  - 実装・実験の記述として、クライアントがモデル更新を送信し、サーバが集約後に配布することが明示されている（概念図やコードのみは不可）。

ここではネットワーク遅延・損失・混雑の評価や、多プロセス/多デバイス構成は必須条件としない。アップロード/ダウンロード/集約が実際に起きたことを示せれば十分である。

3.7.2　環境カテゴリの定義

環境カテゴリは以下の通り定義する：

- simulation：3.7.1 の二つの条件を同時に満たす。
- simulation+real：simulation の条件を満たし、さらに実験評価に実物理デバイスを用いた測定が明示される。
- computation：通信実行の証拠がなく、単機の順次計算として解釈できる、または通信が概念的記述に留まる。
- computation+real：computation に該当し、追加で実物理デバイス上で性能測定があるが、連合学習通信過程の証拠がない。
- unknown：記述不足により通信動作の有無を確証できない。

3.8　本章のまとめ

本章では、Google Scholar および ScholarGPT を用いた文献収集方針、文献の選定基準、記録項目、分類規則、運用指標の操作的定義、ならびに実験環境判定規則を提示した。次章以降では、本章の手順に基づき整理した文献集合を集計し、研究動向を説明する。

第4章　連合学習ベースの DDoS 検知に関する研究現状

4.1　研究焦点の推移

本節では、第 3 章の選定基準に基づいて整理した対象文献について、年次分布と研究焦点の推移を概観する。特に、2023 年前後の「実現可能性の確認（探索期）」と、2024 年以降の「最適化・運用志向（発展期）」という流れを、本研究の整理で付与したタグ（Non-iid 対応、通信最適化、個性化、ポイズニング攻撃耐性、セキュリティ／プライバシー）に基づいて示す。併せて、実運用を議論する上で重要な指標（推論遅延、通信量、通信混雑、低性能デバイスでの評価、E2E 応答）の報告状況を年次別に整理し、第 5 章の現実運用に近づけるための論点整理へ繋ぐ。

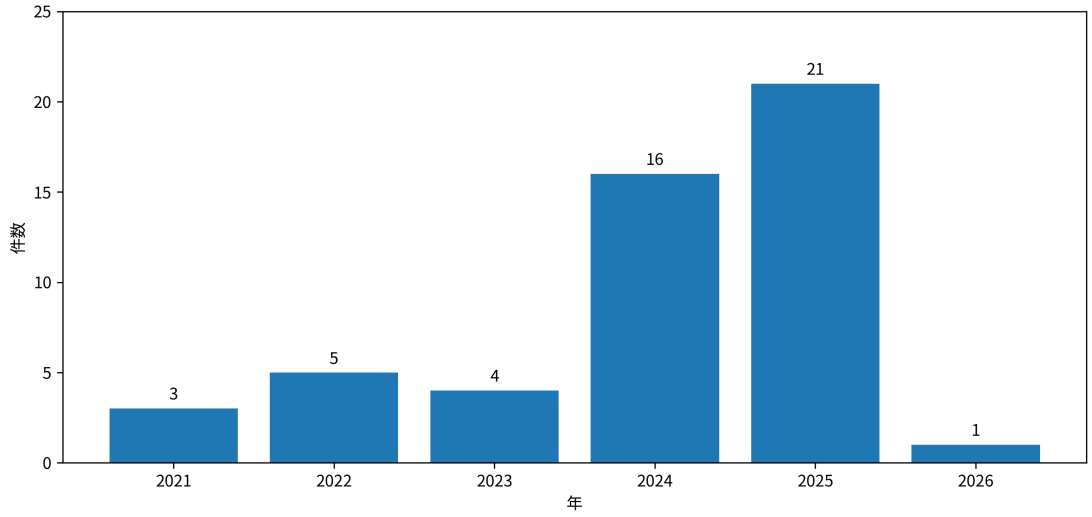

図 4.1: 対象文献の年次分布 (論文数=50)

年次推移（探索期 → 発展期）　図 4.1 より、対象文献は 2024–2025 年に大きく集中しており、当該分野が近年急速に拡大していることが分かる。2023 年前後は、連合学習 を DoS/DDoS 検知へ適用すること自体の有効性（精度、検知の成立）を検証する探索的研究が中心になりやすい。一方、2024 年以降は論文数が急増し、

表 4.1: 連合学習の課題改善傾向（数（割合%））

年次区分	Non-iid 対応	通信最適化	ポイズニング攻撃耐性	セキュリティ／プライバシー
∼2023（n=12）	5（41.7）	2（16.7）	2（16.7）	7（58.3）
2024（n=16）	6（37.5）	4（25.0）	3（18.8）	9（56.2）
2025∼（n=22）	9（40.9）	8（36.4）	1（4.5）	8（36.4）

Non-iid（分布の偏り）や通信コストなど、FL 特有のボトルネックに対して最適化を試みる研究が増える傾向が見られる。

表 4.1 より、2024 年以降は「通信最適化」など、性能最適化や運用適合性を意識した設計要素が顕在化する。この傾向は、研究の関心が「連合学習を用いた検知の成立性確認（探索期）」から、「通信制約を含む運用条件下での効率化（発展期）」へ移行していることを示唆する。また、Non-iid 対応は全期間で一定割合存在し、分布の偏りが継続的な主要課題であることが分かる。一方で、ポイズニング攻撃耐性は年次により変動があり、体系的に取り上げられる段階には至っていない可能性がある。

評価設定と「現実性」指標の年次差　実運用を議論する上で重要な観点として、本研究では以下の五つの指標に着目する：

- 推論遅延（Inference latency）
- E2E 応答（検知から遮断・緩和動作まで）
- 通信量／通信オーバーヘッド
- 通信混雑・損失・攻撃されている時の条件
- 低性能デバイスでの評価（例：Raspberry Pi、Jetson Nano、ESP32 等）

表 4.2: 年次別に見た「現実性」指標の報告状況（数（割合%））

年次区分	推論遅延	通信量	通信混雑	低性能デバイス評価	E2E 応答
∼2023（n=12）	1（8.3）	1（8.3）	0（0.0）	0（0.0）	0（0.0）
2024（n=16）	4（25.5）	3（18.8）	0（0.0）	2（12.5）	0（0.0）
2026（n=22）	8（36.4）	7（31.9）	2（9.1）	4（18.2）	0（0.0）

表 4.2 に、年次別の報告状況を示す。推論遅延や通信量は 2025 年にかけて報告数が増える傾向が見られる一方、通信混雑の考慮は依然として少ない。また、E2E

応答は全期間でほぼ未報告であり、研究評価と運用要件の間にギャップが残る。特に、低性能デバイス（Raspberry Pi 等）を用いた実測評価は少数であり、2025 年に入ってから増加するものの、全体としては例外的である。したがって、端末性能制約を前提とした議論（処理能力、消費電力、メモリ、モデル更新の実行可能性）は、多くの研究で十分に裏付けられていない可能性がある。

表 4.2 より、2023 年前後から 2024 年にかけては、推論遅延や通信量といった指標の報告割合は高いとは言えず、特に 2024 年は論文数が増えたにもかかわらず報告割合が低い。2025 年には推論遅延および通信量の報告割合が上昇するが、それでも多数派ではない。また、E2E 応答は全期間でほぼ未報告であり、研究評価と運用要件の間にギャップが残る。

4.2 データセット

本節では、本調査で対象とした論文群が用いるデータセットを整理し、三つの観点から特徴をまとめる：

1. 想定する展開シナリオ（IoT/IIoT・一般ネットワーク・SDN・車載 CAN など）
2. データ形態（PCAP フロー特徴量／ NetFlow ／ログ・マルチモーダル）
3. 攻撃カバレッジ（DoS/DDoS を中心とした攻撃種別やラベル設計）

連合学習を用いた DoS/DDoS 防御では、**通信・端末側推論・特徴抽出コスト**が現実での運用に直結するため、**どの形態のデータを前提に評価しているか**を明示することが重要である。

4.2.1 データ形態と評価上の意味

本調査で確認できたデータ形態は大きく以下に分類できる：

- **PCAP（パケットキャプチャ）**：生データに近く再現性は高い一方、前処理（フロー化・DPI など）や端末側負荷が大きい。
- **フロー特徴量（CICFlowMeter など）**：多くの研究で採用される代表形態である。ただし、特徴抽出器の実装・設定差が結果に影響し得るため、学習器の比較だけでなく **前処理パイプライン** の明記が望ましい。
- **NetFlow ／標準化特徴セット（NF-*／ CIC-ToN-IoT など）**：特徴が統一され軽量で、端末側・エッジ側での推論／学習に適する。一方で、生パケット由来の情報（ペイロードなど）は失われ、攻撃識別の上限に影響し得る。

- **ログ／テレメトリ（マルチモーダル）**：ネットワークトラフィックに加えて OS ログや IoT テレメトリを含む。展開現実性は高いが、クライアント間でモダリティが揃わない（Non-iid）状況が生じやすい。

4.2.2 公開データセットの整理

表 4.3 に、本調査で参照されたデータセットのうち、公開かつ入手可能なものを中心に整理する。整理する時の分類基準はデータセットの展開シナリオや特性を基に分類する。主に以下の 4 分類となる：

A. **一般ネットワーク IDS ／ DDoS**： 観測点は組織ネットワーク境界（ルータ／ FW ／ IDS）を想定し、推論や学習はサーバ側になりやすい。比較や再現性に優れる一方、IoT 端末側の特徴（低性能、低帯域幅、不安定リンク、特殊プロトコル）を直接反映しにくい。

B. **IoT/IIoT 指向**：観測点は IoT ゲートウェイ／エッジを想定し、展開デバイス主にゲートウェイ機器やエッジサーバである。端末と環境の異質性が強く、Non-iid や時変性を含む「現場寄り」の評価に向く。

C. **標準化・派生（NetFlow ／特徴標準化）**：フロー化・標準化により前処理と更新量を抑えやすく、低負荷デバイス（ゲートウェイ／エッジ）での運用仮定と相性が良い。一方、変換に伴う情報欠落により、軽量化と検知性能のトレードオフを意識した解釈が必要となる。

D. **特定ネットワーク（SDN ／車載 CAN 等）**：観測点と展開デバイスが領域固有（例：SDN コントローラ／スイッチ、車載 ECU ／ゲートウェイ）で、遅延制約と観測可能特徴が明確なため、展開条件を具体化した議論がしやすい。ただし適用範囲は当該領域に限定されるため、一般 IoT への使用は慎重に扱う。

表 4.3: 本調査で用いられた主要な公開データセット（特性サマリ）

データセット	年	想定領域	データ形態	主要特性（DoS/DDoS との関係・備考）
(A) 一般ネットワーク IDS ／ DDoS				
CIC-IDS2017 [14]	2017	一般 IDS	PCAP Flow	多種攻撃＋ benign を含む代表的 IDS 基準である。フロー特徴量系研究で頻出する。
CSE-CIC-IDS2018 [15]	2018	一般 IDS	PCAP Flow	IDS2017 の拡張（規模・シナリオ）である。前処理差の影響に注意が必要である。
CIC-DDoS2019 [17]	2019	DDoS	PCAP Flow	反射／増幅を含む DDoS 評価に特化する。DDoS 分類・検知のベンチマークとして頻用される。
UNSW-NB15 [19]	2015	一般 IDS	PCAP Flow	混合トラフィック基準である。旧来ベンチマークだが比較目的で依然として用いられる。
NSL-KDD [29]	2009	一般 IDS	Flow	古典的基準（KDD99 系）である。比較には便利だが現実性は低く、運用結論の根拠としては限定的である。
CAIDA DDoS Attack 2007 [13]	2007	DDoS	PCAP (Trace)	実ネットワーク由来の DDoS イベントである。取得に規約同意等が必要な場合がある。
(B) IoT/IIoT 指向（DoS/DDoS ／ボットネットを含む）				

続く...

データセット	年	想定領域	データ形態	主要特性 備考
ToN-IoT [21]	2020	IoT/IIoT	トラフィック＋ ログ／テレメトリ	マルチモーダルで展開現実性が高い。クライアント間の Non-iid が顕在化しやすい。
BoT-IoT [20]	2018	IoT	PCAP Flow	IoT ボット ネット・DoS/DDoS 等を含む。IoT-IDS の定番基準として参照される。
IoT-23 [22]	2020	IoT	PCAP Zeek ログ	シナリオ単位でラベル化される。マルウェア／ボットネット解析や IDS に利用される。
IoTID20 [25]	2020	IoT	トラフィック Flow	IoT ネットワーク侵入検知向けである。攻撃種別や前処理条件は原典に依存する。
N-BaIoT [27]	2018	IoT	時系列 Flow	実 IoT 機器を対象としたボットネット検知で知られる。AE などの異常検知とも相性が良く、ウィンドウによる時系列データも持つ。
Edge-IIoTset [23]	2022	Edge/IIoT	トラフィック Flow	エッジ／ IIoT を意識した包括的データセットである。連合学習／集中学習の比較で用いられる。
CICIoT2023 [18]	2023	IoT/IIoT	PCAP Flow	大規模 IoT 攻防ベンチマークとして比較的新しい選択肢である。
(C) 標準化・派生（軽量化／比較容易性を重視）				

続く...

データセット	年	想定領域	データ形態	主要特性 備考
NF-ToN-IoT-v2 [31]	2023	IoT/IIoT	NetFlow 特徴量	ToN-IoT を NetFlow／標準特徴へ変換した派生版である。端末側負荷を下げた評価に向く。
NF-BoT-IoT-v2 [32]	2023	IoT	NetFlow 特徴量	BoT-IoT の標準特徴版である。軽量推論やクロスデータ比較に利用しやすい。
NF-CICIDS2018-v2 [33]	2023	一般 IDS	NetFlow 特徴量	CICIDS2018 の標準特徴版である。前処理の統一により比較可能性を高める。
CIC-ToN-IoT [30]	2023	IoT/IIoT	標準化特徴量	ToN-IoT 派生（標準化特徴）として扱われる。定義は配布元の仕様に従う。
(D) 特定ネットワーク（SDN／車載 CAN）				
InSDN [24]	2020	SDN	トラフィック Flow	SDN 環境の侵入検知向けである。制御プレーン等を意識した評価で用いられる。
Car-Hacking [35]	2016	車載 CAN	CAN フレーム	CAN 注入（DoS／Fuzzy など）を含む代表的ベンチマークである。
OTIDS [36]	2018	車載 CAN	CAN フレーム	車載ネットワーク向け IDS データセットとして参照される。

不確定要素（データセット名・公開が不明確なもの） データセットには、論文によって独自収集（例：実車 3 台の自収集データ、特定 5G コア監視データ）や、名称が一般名詞に留まる記述（例：`Automotive`、`Modbus`）も含まれる。これらは、**データセットとしての同定（名称・配布元 URL・版・ライセンス）**が困難であるため、統計・比較では「自作／不明」として区別して扱う。

データセット傾向　対象文献では公開データセットの利用が多く、図 4.2 に公開データセットの使用頻度上位を示す。CIC-DDoS2019 [17] と CIC-IDS2017 [14] の使用頻度が高く、N-BaIoT [27]、UNSW-NB15 [19]、CICIoT2023 [18]、ToN-IoT [21] がこれに続く。加えて、NF-BoT-IoT-v2 [32]、InSDN [24]、CSE-CIC-IDS2018 [15]、NSL-KDD [29] といったデータセットも用いられている。第 5 章では、データセット選択と評価設計の関係について整理し、検討する。

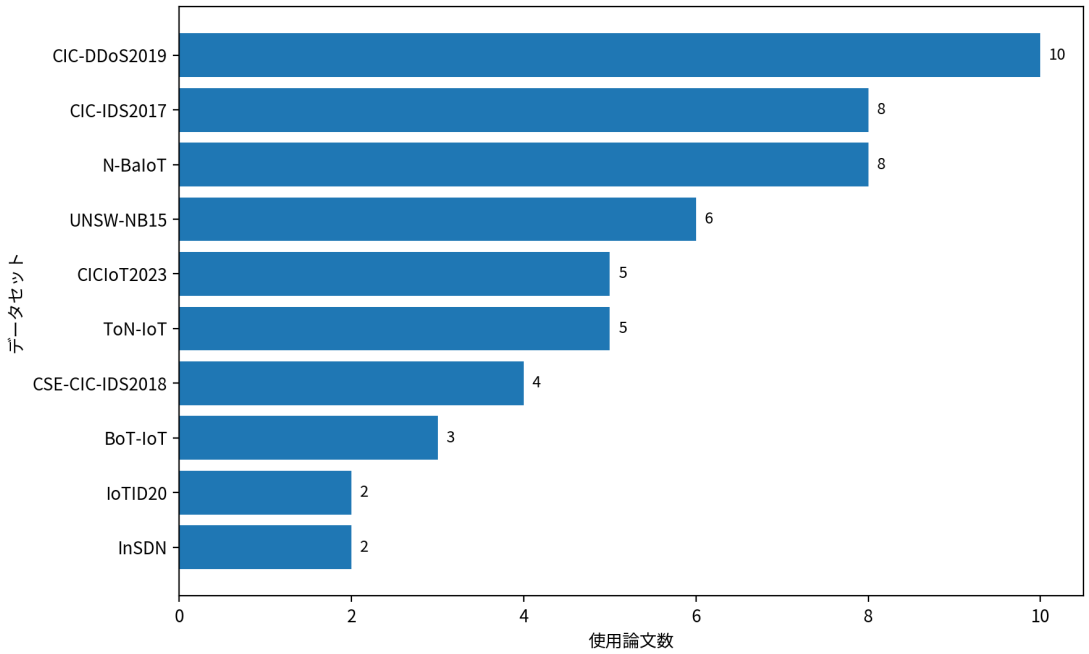

図 4.2: データセット使用頻度 Top10

4.3　検知モデル・アーキテクチャ別の整理

本節では、対象文献を検知モデルの系統に基づいで分類し、各系統の研究目的と設計上の着眼点の違いを整理する。本章の最後に、表 4.4 に、各文献の基本属性（対象、データセット、検知方法、連合学習 の枠組み、目的、評価指標）を一覧として示す。本表は対象（IoT 端末／エッジ／ネットワーク／制御層）、検知器（CNN/DNN/RNN/GNN/Autoencoder/Transformer）、連合学習手法、目的、評価指標を比較できるように構成している。

続く各節では、各グループの違いを、入力・特徴表現（フロー統計量、時系列、グラフ構造、潜在表現など）と、目標（Non-iid の影響緩和、通信・計算資源の制約、ポイズニング攻撃への耐性、プライバシー確保等）で要約する。

4.3.1 G1：CNN / ResNet / 1D-CNN 系（局所パターン抽出による高精度化）

G1 は、フロー特徴量等から局所的な判別パターンを抽出し、分類性能（Acc/F1 等）を極めて高く引き上げる手法である。近年では、FedAvg を基盤としつつ、動的重み付けやクラスタリングによる Non-iid への強力な適応、さらには 8-bit 量子化や通信圧縮を用いたエッジデバイス向けの大幅な軽量化・通信量削減に成功している研究が多い。

一方、こうした精度の高さや計算効率の向上が示される反面、実際の DoS/DDoS 攻撃に伴う通信混雑やパケット損失下での更新成立性、あるいは E2E（End-to-End）の応答遅延まで含めた運用評価は依然として限定的である。

各論文概要

1. **AFL-SecNet (2026)** [39]：個性化 CNN を用いた連合学習に、同態暗号ベースの安全集約（更新秘匿）を組み合わせ、複数ドメイン間でも生データを共有せずに DDoS 検知モデルを共同学習する設計である。プライバシー要件が強いネットワーク環境でも協調検知を成立させる点が狙いである。

   UNSW-NB15 において 97.58%の精度と 1.15 秒の推論遅延、CICDDoS2019 で 96.60%の精度と 0.63 秒の低推論遅延を達成し、強力なプライバシー保護と高い特徴抽出能力を両立させた。

   一方で、鍵管理や暗号計算に伴う追加負荷、協調復号における信頼前提の置き方（運用上の依存関係）が、実装・拡張のボトルネックになり得る。

2. **FLAME (2025)** [40]：Non-iid で更新が「外れ値化」しやすい点に着目し、クライアント更新の分布推定（KDE など）と分布お違い（JSD など）で逸脱を検知しつつ集約を制御する。ゼロデイ攻撃に対しても過度な精度崩れを避け、収束を安定させることを狙う。CICDDoS2019 では特定クライアントで最大 99.81%の高精度を記録するなど、優れたロバスト性を示した。

   ただし、距離推定や制御ロジックはスケール時の計算コストやパラメータ感度が課題になりやすく、実環境のドリフトや多様な回避攻撃を含む条件での体系的検証が今後の論点となる。

3. **FedDWC (2025)** [41]：データ分布の類似性に基づくクライスタリングと、ローカルモデルの性能（検知精度）に応じた動的重み付け（Dynamic Weighting）を革新的に統合し、Non-iid 環境下での学習効率を飛躍的に向上させた。200 台の IoT デバイスを想定した Non-iid 環境において、IoTID20 データセットで 99.21%、CICIoT2023 で 99.10%という極めて高い検知精度を達成し、大規模ネットワークにおいても 99.41%の高精度を維持した。

一方で、モデルポイズニングやバックドア攻撃といった敵対的攻撃に対する防御機構はまだ実装されておらず、悪意のあるノードが混入した非理想的な実運用環境下において、この動的重み付けやクラスタリングがどの程度堅牢に機能するかの検証が、実用化に向けた次の焦点となる。

4. **Albanbay et al. (2025)** [42]：Albanbay et al. (2025) は、計算資源が限られた IoT デバイス上でのローカルモデル展開の実現可能性と、データ規模やクライアント数がモデルに与える影響を検証した。同研究は、最大 150 台の IoT デバイスを想定した大規模な連合学習環境を構築し、さらに Raspberry Pi 5 実機を用いた展開テストを実施するという極めて実践的な貢献を果たした。

   特に、CNN モデルが約 94%〜95%の高い検知精度を維持しつつ、1 サンプルあたり約 104ms の低い推論遅延と安定した熱特性（72 ℃以下）を達成し、エッジハードウェアにおける精度と計算効率の最適なトレードバランスを実証した。

   一方で、本実験は静的なデータセットを用いたオフラインの制御環境下で実施されており、実際の IoT 環境で発生し得る通信混雑、パケット損失、端末の突発的な離脱、あるいは悪意のあるノードの混入（ポイズニング攻撃）といった非理想的な運用条件下での包括的な検証は行われていない。動的な実ネットワーク環境における学習更新の成立性や E2E のロバスト性評価が、今後の実用化に向けた重要な課題として残されている。

5. **FL-DAD (2024)** [43]：CNN を連合学習で学習し、IoT ネットワークで分散 DDoS 検知を成立させることに特化した論文である。分類性能に加えて通信量や遅延といったシステム指標も併記し、「分散運用として何が律速になるか」を意識して構成した。CICIDS2017 データセットを用いた評価において 98%以上の高い検知精度を達成し、学習の進行に伴い 1 エポックあたりの通信オーバーヘッドを 7.9MB まで削減し、IoT ノード数が増加しても高い精度を維持し続けるという優れたスケーラビリティと通信効率を実証した。

   一方、アルゴリズム自体に新規性が見当たらなく、Non-iid 問題が対策を取り組んでない。連合学習プロセス自体は実装されているものの，更新受送信の実測/通信シミュレーションや，異種端末の性能制約，ノード増加に伴う集約負荷，更新頻度増大時の通信負担といった運用上の制約を含めた E2E 評価は限定的であり，非理想環境下での包括的検証が今後の課題となる。

6. **FEDSA-ResnetV2 (2024)** [44]：FEDSA-ResnetV2 (2024) は、車路協調（IoV）のエッジ環境に向けて、自己注意機構（Self-Attention）を組み込んだ軽量な SA-ResnetV2 と 8-bit 後量子化技術を統合した連合学習フレームワークを提案した。

研究では、検知性能をベースライン比で最大34%向上させただけでなく、8-bit 量子化により推論時間、モデルサイズ、パラメータ数、mFLOPs といった実行効率指標を最大72.6倍も劇的に最適化することに成功し、計算資源が限られたエッジデバイスへの実用的な展開基盤を証明した。

一方、モデルや通信の軽量化には大きく貢献しているものの、IoV 特有の高速移動に伴う動的なネットワーク変動や、実際のサイバー攻撃により深刻な通信混雑が発生した非理想的な条件下で、この最適化された更新プロセスがどの程度安定して成立するかについての E2E 検証は、今後の運用課題として残されている。

7. **NIDS-FGPA (2024)** [45]：NIDS-FGPA (2024) は、トラフィックを画像化して 2D-CNN と BiGRU で時空特徴を抽出するとともに、Paillier 同態暗号と勾配相似度（GSA）に基づく動的集約を組み合わせた。

   研究では、Edge-IIoTset と CIC IoT 2023 データセットでそれぞれ 94.5%と99.2%の高精度を達成した。さらに、GSA に基づく集約を導入することで、従来手法と比較して通信オーバーヘッドを 22%〜51%削減することに成功し、暗号化がもたらす負荷の相殺に大きく貢献した。

   一方、同態暗号の導入による絶対的な通信量は依然として極めて大きく、論文内でもわずか 256-bit 鍵の構成でさえ最大 363.2MB の通信オーバーヘッドに達することが報告されている。本論文の第 5 章で論じるように、帯域幅が制限された IoT 環境や DDoS 攻撃に伴う通信混雑時において、このような大規模な暗号化パラメータの送受信を安定して完遂できるかどうかが、実用化に向けた最大の課題となる。

4.3.2 G2：DNN / MLP 系

G2 は、表形式特徴量に対して計算負荷の小さい DNN（深層ニューラルネットワーク）や MLP（多層パーセプトロン）を適用し、推論を軽量に保ちつつ、連合学習の運用プロセス自体（非同期化、適応的なクライアント選択、動的重み付け、早期停止など）を高度化する手法である。

このアプローチによりエッジ環境での学習効率と安全性が劇的に向上している。CPU 負荷や収束時間を大幅に削減しつつ、99%以上の高精度を維持する研究も存在している。

一方、これらの成果の多くは理想的な環境に留まっている。実際の DDoS 攻撃に伴う深刻な通信帯域の枯渇やパケット損失下において、こうした高度な制御ロジックがタイムアウトせずに正常に機能するかどうかといった、非理想環境下での更新の成立性や E2E 評価は依然として限定的である。

各論文概要

1. **ARAFL-BAD (Lohani & Sangal、 2025)** [46]：IoT エッジ環境において、ポイズニングやバックドアといった対抗的攻撃（Adversarial attacks）への強力な耐性とゼロデイ攻撃の検知を両立する、非同期型の連合学習フレームワークを提案した。クライアント側では MLP を学習しつつ、FGSM/PGD 等で生成した対抗例も混ぜた学習を行い、サーバ側は信頼度に基づく重み付け集約でポイズニング・バックドア等の影響を抑える。

   Raspberry Pi 4B を含む実機エッジデバイスを用い、WebSocket ベースの非同期通信による実践的な実装評価を実施した。その結果、平均 98.29%の高精度と 98.30%の F1 スコアを達成しつつ、多様な対抗的攻撃の成功率（ASR）を平均 5.59%まで抑え込むことに成功した。さらに、推論時間を約 0.3ms、1 ラウンドあたりの通信オーバーヘッドを約 0.78MB という極めて軽量な水準に最適化し、リソース制約の厳しいエッジハードウェアにおける高い実用性とセキュリティの両立を実証した点である。

   一方で、計算資源の消費や非同期更新の有効性については実機ネットワーク環境で詳細に評価されているものの、実際のボットネット攻撃発生時や DDoS 攻撃に伴う深刻な通信帯域の枯渇、パケット損失が発生する非理想的なネットワーク環境下での評価は行われていない。

2. **DTFL-CD (Salim et al.、 2024)** [47]：医療 IoT 環境を対象に、エッジ層にデジタルツイン（DT）を構築し、適応的な閾値判定と早期停止（ATES）を用いた連合学習（FLDS）によってゼロデイサイバー攻撃を検知するシステムを提案した。

   VMware 上に構築された仮想エッジサーバと仮想 IoT デバイスを用いたシミュレーション環境において、最大 0.98 の F1 スコアを達成した。さらに、ATES を用いることで従来の手法（FedAvg）が 120 ラウンドで 83%の CPU を消費するのに対し、50 ラウンドで 75%の CPU 消費に抑えるという高い計算効率を実証した点である。また、1Mbps から 1000Mbps までの帯域幅を変化させた通信テストを行い、エッジベースの DT がクラウドベースの IDS と比較して応答遅延を最大 33%削減し、DT の同期遅延も平均 0.02ms という極めて低い水準に維持できることを証明した。

   一方で、これらの評価は帯域幅を制御した仮想マシン上のシミュレーション環境で行われており、リソースが枯渇する物理エッジデバイスでの実証には至っていない。さらに、DDoS 攻撃に伴うパケット損失や深刻な通信混雑が発生した非理想的なネットワーク環境下での評価は行われていない。

3. **FLAD (Doriguzzi-Corin & Siracusa、 2024)** [48]：サーバ側にテストデータが存在しないというセキュリティ上の厳しい前提の下で、クライアン

トの局所的な検証スコアに基づいて適応的に計算量（エポック数や勾配降下ステップ数）を割り当てる連合学習メカニズムを提案した。

研究は、CIC-DDoS2019 データセットを用いた評価において、50 クライアントの大規模環境で 0.9899 の F1 スコアを達成し、全体の学習時間をわずか 363 秒に短縮した。また、13 クライアントの極端な Non-iid 環境（各クライアントが 1 種類の攻撃のみを保持）においても、FedAvg や既存手法を凌駕する 0.9667 の F1 スコアを達成しつつ、1 ラウンドあたりの学習時間を約 9 秒に劇的に最適化した。このように、通信や計算リソースを最適配分することで、高い精度と収束の高速化を見事に両立している。

一方で、これらの優れた学習効率は、通信遅延の影響を排除するためにサーバとクライアントを同一マシン上の単一 Python プロセス（ローカル呼び出し）として実装した環境下で得られたものである。実際の DDoS 攻撃に伴う深刻な帯域枯渇やパケット損失が発生する非理想的なネットワーク環境下の実験は課題となる。

4. **Weighted FL for LR-DDoS in SDN (Ali et al.、2023)** [49]：IoT ネットワークを管理する SDN のコントロールプレーンを標的とした低レート DDoS（LR-DDoS）攻撃を検知するため、重み付き連合学習（WFL）モデルを提案した。

   研究は、CAIDA データセットと MATLAB を用いたシミュレーション環境において、Levenberg-Marquardt（LM）、Bayesian Regularization（BR）、Scaled Conjugate Gradient（SCG）という 3 つの異なる学習アルゴリズムで局所モデルを訓練した。フェデレーテッドサーバにおいて、これら局所モデルの予測性能に基づいて独自の優先度（重み）を割り当てて集約することで、98.85%の高い分類精度と 94.21%の F1 スコアを達成した。さらに、1 レコードあたりの推論時間をわずか 0.019ms に抑え、局所重みのみを送信することでメモリや処理要件を直感的に削減し、SDN ベースの IoT ネットワークにおける侵入検知の高い有効性を実証した。

   一方、局所モデルに対する優先度の割り当ては試行錯誤（hit and trial）に基づく静的な設定であり、動的な環境変化への適応性に課題が残る。また、実験は MATLAB 上の単一 computation 環境で実施されており、モデルのパラメータ送信による通信オーバーヘッドの定量的な計測や、実際の LR-DDoS 攻撃発生時に生じるネットワーク遅延、パケット損失、帯域幅枯渇が連合学習の更新プロセスに与える影響については評価されていない。

5. **FedDDoS (Pinto Neto et al.、 2022)** [50]：分散型マルチテナント IoT 環境において、異なるテナントが経験した DDoS 攻撃の知識を、生データを共有することなく協調的に学習するフレームワーク（CIDD）を提案した。連合学習の DDoS 攻撃検出では、かなり早期で模索を始めた代表の一つである。

同研究は、CICDDoS2019 データセットを 6 つのテナントおよび複数のエッジ地域に論理的に分割したシミュレーション環境において、ローカルスケーリングと多層パーセプトロン（MLP）を用いた連合学習を評価した。その結果、128 ノードの隠れ層を持つモデルが 84.8%、64 ノードのモデルが 84.2%という分類精度を達成し、個別学習よりも協調学習がテナントの保護能力を向上させることを実証した。さらに、ローカルデータを用いた標準化（Local scaling）を導入することで、データやスケーラ（標準化パラメータ）自体を交換することなく、プライバシーを維持したまま高い分類精度を実現できることを示した。

一方で、本実験は Flower フレームワークを用いた単一のワークステーション上で論理的なデータ分割により実施されており、実際のマルチテナントエッジ環境における仮想マシン間のネットワーク通信や、実データの送受信を伴うエミュレーションは行われていない。また、モデルパラメータの送信に伴う通信オーバーヘッドの定量的な計測も未報告である。

4.3.3 G3：RNN / LSTM / GRU 系

G3 は、セッション継続や低レート攻撃など「時間的依存」を取り込むことで、静的特徴だけでは拾いにくい挙動を捉えることである。RNN、LSTM、GRU などの再帰型ニューラルネットワークを採用した文献である。

近年、このアプローチにより、単発のパケットや静的な特徴量だけでは検知が困難な低レート DDoS 攻撃や、複雑な分散型攻撃に対する検知精度が飛躍的に向上している。例えば、CNN と Bi-LSTM を組み合わせたハイブリッドモデルや、Self-Attention 機構と最適化アルゴリズムを統合した GRU モデルにより、99%以上の極めて高い検知精度が複数報告されている。

しかし、時系列データの処理は系列長や時間窓の設計に強く依存し、モデルの計算量やメモリ消費が増大しやすいという弱点を持つ。特に、実際の DDoS 攻撃に伴う通信帯域の枯渇やパケット損失が発生した場合、入力となる時系列データ自体に欠損が生じたり、モデル更新の遅延によって「平常時の遷移パターン」の学習が追従できなくなるリスクが高い。

各論文概要

1. **FedXAI（Kalakoti et al.、 2025）** [51]：IoT 環境におけるボットネットおよび DDoS 攻撃の検知に向けて、LSTM モデルを用いた連合学習フレームワークを提案した。さらに、深層学習特有のブラックボックス性を解消するため、クライアント側で生成した SHAP や LIME などの XAI（説明可能な AI）の解釈結果を安全性を確保した上で集約し、生データを共有せずにサーバモデルの予測根拠を可視化する画期的なアプローチを導入した。

研究では、N-BaIoT データセットを用いた単一高性能サーバ上の評価において、マルウェアの二値分類で精度 99.90%、F1 スコア 99.90%、ボットネットタイプの多クラス分類で精度 99.28%、F1 スコア 99.54%という極めて高い検知性能を達成した。加えて、XAI を用いた特徴量重要度の可視化により、誤検知の原因特定や SOC（セキュリティオペレーションセンター）での信頼性構築に大きく貢献している。

一方、時系列データを処理する LSTM モデルそのものの計算負荷に加え、SHAP 等の解釈生成や暗号化された安全集約プロセスは、IoT エッジデバイスに対してさらなる計算およびメモリのオーバーヘッドを強いる。また、現在の評価は各エッジデバイスのデータが独立同分布（IID）であることを前提としており、現実の IoT 環境で生じる Non-iid データに対する有効性が未検証である。その上、SHAP による解釈生成はクライアント数が増加するにつれて計算コストが著しく増大するため、大規模な IoT ネットワークへの拡張性に欠ける点が指摘されている。

2. **SAMFL-SCDCOA（Alohali et al.、 2025）** [52]：GRU のゲート機構（入力/更新ゲート）による「必要な履歴だけを保持して過去依存を制御する」性質に Self-Attention を重ね、系列中の重要区間に重みを置いて識別を強化する検知器を連合学習で学習するフレームワークを提案した。

   さらに特徴選択やハイパラ最適化を多段で組み合わせ、系列モデルの性能を最大化する方針をとり、CICIDS-2017 データセットおよび UNSW-NB15 データセットを用いた評価において、それぞれ 99.14%および 99.04%の極めて高い検知精度と、97.76%および 91.69%の F1 スコアを達成した。さらに、複雑なモデル構造を持ちながらも、処理時間（Computational Time）を約 5.08 秒から 7.88 秒に抑え、従来手法（RNN や LSTM 単体）と比較して計算効率と分類精度の優れたバランスを実証している。

   一方、特徴選択から注意機構、ハイパーパラメータ最適化に至る多段的な処理パイプラインは、各エッジデバイスに対して少なからぬ計算負荷を要求する。公開データセットへの依存により現実のサイバー攻撃の複雑性を完全に網羅できていない点が著者らにより指摘されている。さらに、リアルタイム環境やエンタープライズ規模のネットワーク環境での評価が行われていないため、ネットワークトラフィックが高負荷な状況下でのスケーラビリティや応答遅延（レイテンシ）、および既存のセキュリティインフラへの統合やハードウェア要件といった現実のデプロイメントに関する課題が未解決のまま残されている

3. **SCNN-Bi-LSTM（Bukhari et al.、 2024）** [53]：無線センサネットワーク（WSN）における侵入検知に向けて、スタック型 CNN（SCNN）と双方向 LSTM（Bi-LSTM）を組み合わせたハイブリッド深層学習モデルを連合学

習フレームワークに統合した手法を提案した。

研究は、SCNN を用いてトラフィックデータから空間的な局所特徴を抽出し、Bi-LSTM を用いて時間的な依存関係や文脈を学習するアプローチを採用している。WSN-DS データセットおよび CIC-IDS-2017 データセットを用いた評価において、前者で 99.70%の精度と 99.60%の F1 スコア、後者で 99.93%の精度と 99.93%の F1 スコアという極めて高い分類性能を達成した。生データをクライアント側に保持したままモデルパラメータのみを同期させることで、プライバシーを保護しつつ、集中学習手法や従来の機械学習モデル（SVM、LightGBM 等）を凌駕する検知能力を実証している。

一方、大規模かつ複雑なネットワーク環境へ展開した場合のスケーラビリティに懸念があり、計算コストを法外に増加させることなく増大するデータ負荷を処理できるかどうかが課題とされている。さらに、現在の評価は特定のネットワークアーキテクチャに偏っており、多様なネットワーク環境への適応性が未検証である。また、リアルタイム検知が求められる環境における計算時間（推論・更新にかかる時間）の制約や、トラフィックが動的に変動する実際のネットワーク条件下での性能評価が不足しており、広範な実世界でのテストとデプロイメントが今後の重要な研究方向として挙げられている。

4. **FL-AD（Mothukuri et al.、 2022）** [54]：IoT ネットワークにおけるセキュリティ攻撃の検知において、データプライバシーを保護するため、GRU を用いた異常検知モデルを連合学習フレームワークに統合した手法を提案した。

   研究は、PySyft を用いた仮想環境上で IoT デバイスの分散学習をシミュレートし、Modbus プロトコル等のネットワークデータに対する評価を行った。評価の結果、異なるウィンドウサイズや層構造を持つ GRU モデルにおいて、最高で約 95.65%の検知精度（Accuracy）を達成した。さらに、ランダムフォレストを用いたアンサンブル学習により複数モデルの予測を統合し、従来の集中型機械学習モデルと比較して、ユーザデータのプライバシーを保護しつつ、同等以上の最適な検知性能と学習効率（計算リソースの分散化）を実現できることを実証した。

   一方、現在の評価が既存の静的なデータセットを用いたシミュレーション環境に留まっている点を率直に指摘している。実運用の IoT 環境への完全な適用に向けては、実際の IoT デバイス群を用いた物理的なテストベッドを構築し、ライブデータ（リアルタイムなトラフィック）を用いた運用評価を行うことが必要不可欠であると述べている。さらに、デバイス固有のデータから、既知の攻撃だけでなく未知の脆弱性（ゼロデイ攻撃など）を実環境で正確に分類・検知できるかどうかの実証が、今後の重要な研究課題として残されている。

5. **FIDS（Li et al.、 2021）**[55]：RNN 系の非独立同分布（Non-iid）なデータセット学習では、各クライアントの系列分布（平常時の遷移パターン）が異なると隠れ状態の学習方向がずれ、FedAvg 的な平均が効きにくい。そこで本研究は、クライアントのデータ空間をプロトタイプ（代表ベクトル）で要約し、グローバル空間との距離・相関に基づいて寄与を調整することで、系列学習の「ぶれ」を抑え、収束を安定化させる立場を取る。

   一方、課題として、極端にサンプル数が少ない攻撃タイプに対する分類精度の低下を率直に報告している。例えば、論文内の実験において、データセット内にわずか 1 サンプルしか存在しなかった「DrDoS UDP」攻撃については、特徴が類似する「UDP-Lag」と混同しやすく、十分な分類ができなかったと述べており、このような極端なデータ不均衡や少数サンプルに対するモデル性能の改善が今後の課題とされている。さらに、本実験は単一の computation 環境で行われており、実際の DDoS 攻撃に伴う深刻な通信帯域の枯渇やパケット損失が発生した場合に、このプロトタイプベクトルや勾配の送受信がタイムアウトせずに成立するかどうかの E2E 評価は行われておらず、実ネットワークへの展開にはさらなる検証が必要である。

6. **FLDDoS（Zhang et al.、 2021）**[56]：データ分布の不均衡が連合学習のモデル性能に与える悪影響を解決するため、AE ＋ RNN 系で系列特徴を自動抽出しつつ、研究は、ローカルデータの不均衡を解消するために SMOTEENN に基づくデータ再サンプリング手法を採用し、少数派クラスのサンプルを生成して分類器の性能を向上させた。さらに、グローバルデータの不均衡（Global imbalance）に対処するため、中央サーバ側でモデルパラメータを K-Means アルゴリズムによってクラスタリングし、階層的に集約する手法を設計した。

   CICDDoS2019、CICIDS2017、および NSL-KDD データセットを用いた単一計算環境（Computation）での評価において、それぞれ 94.82%、93.26%、99.13%の極めて高い分類精度を達成した。また、K-Means に基づく階層的集約を導入することで、従来の FedAvg アルゴリズムと比較して通信ラウンド数を 40.62%削減することに成功し、モデルの収束速度と通信効率の大幅な改善を実証した。

   一方、未知の攻撃タイプに対する検知能力の不足を率直に指摘しており、既存のモデルをより現実のネットワーク環境に適合させるための拡張が今後の課題として挙げられている。さらに、本実験は既知のデータセットを用いたシミュレーションに留まっており、実際の DDoS 攻撃に伴う深刻な通信帯域の枯渇やパケット損失が発生する物理的な通信環境下において、この K-Means に基づく階層的なモデルパラメータの送受信がタイムアウトせずに機能するかどうかといった、実運用における検証が求められている。

4.3.4 G4：GNN 系

G4 は、ネットワークトラフィックをグラフ構造としてモデル化し、GNN（グラフニューラルネットワーク）を用いてエンティティ間の複雑な相互作用や空間的依存関係を学習する手法を使用する研究の類である。

近年、このアプローチにより、従来の深層学習モデルでは見逃されがちな分散型ボットネットや協調的な DDoS 攻撃の検知能力が飛躍的に向上している。例えば、GraphSAGE や GAT（グラフ注意機構）、さらには時間的推移と空間的相関を同時に捉える STGNN（時空間グラフニューラルネットワーク）を連合学習に統合することで、97%から 99%以上の極めて高い検知精度が複数報告されている。また、一部の最新研究では、コミュニティ抽出やサブグラフの最適化を通じて、グラフ構造の表現力を維持しつつ通信オーバーヘッドを大幅に（例えば 85%）削減することに成功し、IoT エッジ環境での分散協調検知の基盤を確立しつつある。

一方、グラフの動的構築（エッジの生成や近傍探索）とメッセージパッシング（ノード間の特徴集約）は、本質的に計算リソースとメモリを大量に消費する。さらに連合学習環境では、各クライアントが観測する局所的なサブグラフの特徴をグローバルモデルに統合する際、通信遅延やパケット損失が発生すると、グラフのトポロジ情報や時間的パターンが欠落・崩壊しやすいという致命的な弱点を持つ。特に、実際の DDoS 攻撃によって通信帯域が物理的に枯渇した非理想的なネットワーク環境下において、この重いグラフ構造の更新やパラメータ同期がタイムアウトせずにリアルタイムで完遂できるかはまだ課題とされている。

各論文概要

1. **FedSTGCN（Wang et al.、 2025）** [57]：エッジコンピューティング環境におけるデータプライバシーの懸念に対処しつつネットワーク侵入検知の精度を向上させるため、時空間グラフニューラルネットワークと連合学習を統合した新しい手法を提案した。ネットワーク攻撃が持つ空間的なトポロジ特徴だけでなく、同一時間帯のデータフローが似た性質を持つという時間的特性にも着目している。プライバシー漏洩を防ぎつつノード情報を集約するため、グラフ分離手法を用いて内部グラフと外部グラフを切り分け、さらに LSTM 手法を用いて時間的な相関を処理するアーキテクチャを設計した。NF-BoT-IoT-v2 および NF-ToN-IoT-v2 データセットを用いた環境での評価において、20 クライアントでの二値分類精度はそれぞれ 96.75%および 97.52%に達した。

   一方、現在の実験や評価がすべて独立同分布なデータセットに基づいている点を率直に指摘している。現実のネットワークシナリオにおいて、異なるクライアントから収集されるデータは通常、独立同分布ではなく極端な偏りを持つことが多い。このような不均一なデータ分布は連合グラフ学習にとって重大な課題を引き起こすため、データ分布が偏った非理想的な環境下におけ

る分散型ネットワーク侵入検知の有効性検証が今後の重要な研究方向として残されている。

2. **FedGATSage（Al Tfaily et al.、 2025）** [58]：oT ネットワークにおける侵入検知において、構造的パターンと時間的パターンの両方を保持しつつプライバシーを保護するため、クライアント側で GAT を用いた処理を行い、サーバ側で GraphSAGE を用いるハイブリッド連合学習アーキテクチャを提案した。生のネットワークデータやモデルパラメータを直接共有する代わりに、各クライアントがネットワークをコミュニティとして抽象化し、その埋め込みのみをサーバへ送信する手法を採用している。NF-ToN-IoT および CIC-ToN-IoT データセットを用いた環境での評価において、前者でバランス精度 78.58%およびマクロ F1 スコア 61.93%、後者でバランス精度 82.78%およびマクロ F1 スコア 80.03%を達成した。さらにコミュニティ抽象化により、ノードレベルの共有と比較して通信オーバーヘッドを 85%削減することに成功し、帯域幅に制約のある環境での実用性を高めている。

   一方、スキャニングのような一部の攻撃タイプにおいて F1 スコアが 0.20 を下回るなど、検知性能に課題が残されている。また、DoS や XSS といった特定の攻撃では誤検出率が高くなる傾向があり、極めて重要なネットワークでは事後処理メカニズムが必要となる可能性がある。さらに、現在の評価は静的なデータセットで行われているが、現実のトラフィックは動的であるため、進化する脅威状況に継続的に適応するためのスライディングウィンドウアプローチの導入などが今後の課題として挙げられている。

3. **Fed-GNN（Li & Qi、 2025）** [59]：連合学習で GNN を回す際の悪性クライアント更新やデータポイズニング、更新漏えいといった攻撃面を強く意識したセキュリティアウェアなフレームワークを提案した。

   同研究は、サーバ側で各ノードのローカルモデル更新とグローバルモデルとの相関度をフロベニウスノルム（Frobenius norm）を用いて算出し、softmax 関数で重み付けを行う注意機構ベースの集約アルゴリズムを導入している。これにより、信頼性の高い更新を強調し、モデルの堅牢性を高めている。さらに、モデルパラメータの更新伝送時の秘匿性を確保するため、関数隠蔽型マルチ入力関数暗号（FH-MIFE: Function-Hiding Multi-Input Functional Encryption）を組み合わせている。この二重の暗号化技術により、GNN の表現力を維持したまま「更新の安全性」を強固に固め、ネットワーク侵入検知タスクやオープンドメインデータセットにおいて、従来手法を凌駕する高い F1 スコアと検知精度を達成した。

   一方で、注意機構による集約と FH-MIFE 暗号の組み合わせがもたらす計算および通信コストの増大を実運用上の課題として率直に認めている。複雑で大規模なネットワーク環境にそのまま適用するには、Fed-GNN の計算効率

とスケーラビリティをさらに向上させる必要があり、通信オーバーヘッドを削減するための「圧縮を考慮した暗号化スキーム」の設計が今後の課題として挙げられている。加えて、本手法の防御メカニズムはノイズ干渉に対しては有効であるものの、悪意のあるクライアントが通常の学習後に直接値の反転を行うパラメータ反転攻撃（Parameter inversion attacks）に対しては、依然として防御効果が不十分であるという限界も報告されている。

4. **FLARE（Ma et al.、2025）**[60]：6G-IoT および IoMT（Internet of Medical Things）主導の医療システムにおいて、動的なネットワークトラフィックとエッジデバイスの計算資源制約に対処するため、資源適応型のマルチレベルグラフモデリングを用いた連合学習フレームワークを提案した。ネットワークトラフィックをスライディングウィンドウ（最適値として 15 秒を採用）によって動的な有向重み付きグラフに変換し、エッジ特徴を考慮した集約メカニズムと適応型マルチホップ埋め込みを用いて特徴抽出を行う。CICIDS2017 および UNSW-NB15 データセットを用いた論理的な分散シミュレーション評価において、それぞれ F1 スコア 0.935 および 0.927、AUC-ROC 0.947 および 0.938 という極めて高い検知性能を達成した。さらに、グラフ構造の複雑化に伴う計算コストを抑えるため、デバイスの計算能力に応じた層の適応的選択を取り入れ、従来の GNN や他の Federated Graph Learning システムと比較して、検知精度と学習時間（約 186 秒）の優れたバランスを実証している。

   一方、現在のモデルがトラフィックデータのみに依存しており、多様な IoMT センサーからのマルチモーダルデータの融合に至っていない点を指摘している。さらに、極めて複雑かつ異質な実際の医療環境へ適用するためには、クライアントごとのパーソナライズされたモデル適応技術や、バッテリー駆動の医療機器にとって死活問題となるエネルギーを考慮した最適化戦略の導入が必要であると率直に認めている。また、本実験は単一の Computation 環境上でデータを論理的に分割して実施されたものであり、実際の 6G-IoT 医療ネットワークにおける深刻なパケット損失や帯域幅の枯渇、重厚なグラフ埋め込みの同期プロセスに与える影響が、今後の重要な検証課題として残されている。

5. **GraphFedAI（Anjum et al.、 2025）**[61]：論文は動的な IoT 環境に特有のプライバシー保護やスケーラビリティといった課題を解決するため、グラフニューラルネットワークと連合学習を統合した DDoS 攻撃検知フレームワークを提案した。IoT ネットワーク上のデバイス間通信パターンを時間の経過とともに進化する動的グラフとしてモデル化している。欠損データを補間技術で復元しピアソン相関に基づく特徴選択で冗長な情報を削減した上で、グラフニューラルネットワークを用いて空間的および時間的な特徴を同時に抽出するアプローチを採用した。CICIoT-2023 データセットを用いた環

境および Raspberry Pi 等の実機を用いたベンチマーク評価において、最高で 98.7%の検知精度と 98.4%の F1 スコアを達成した。さらに 8 ビット固定小数点量子化や勾配の疎化といった圧縮技術を導入することで、各通信ラウンドにおけるアップロードデータ量を約 9.3MB に抑止し、帯域幅が制限されたネットワーク条件下でも低い誤警報率と遅延を維持しながらスケーラブルな分散学習が成立することを実証した。

一方、連合学習の同期プロセス中にグラフ上で発生する計算オーバーヘッドの大きさを率直に指摘している。またグラフのトポロジ自体に対する敵対的な操作やポイズニング攻撃に対する脆弱性も懸念事項として挙げられている。したがって実用的な分散学習フレームワークにおける通信コストのさらなる最適化や、敵対的なグラフ改ざんに対する堅牢な保護メカニズムの構築が、動的かつ大規模な IoT シナリオへ円滑に適用するための今後の重要な研究課題として残されている。

6. **FeCoGraph（Mao et al.、 2025）** [62]：ラベル付きデータが極めて少ない（Few-shot）シナリオにおいて、高度に不均衡かつ Non-iid なサイバー攻撃を検知するため、ラベル認識型のグラフ対比学習を組み込んだ連合学習フレームワークを提案した。

   研究は、ネットワークフローをライングラフ構造に変換してフロー間の高次なトポロジ関係を抽出し、さらに教師あり対比学習を導入することで、クラス内の特徴を凝縮しつつクラス間の差異を最大化するアプローチを採用した。NF-BoT-IoT-v2、NF-ToN-IoT-v2、および NF-CSE-CIC-IDS2018-v2 の 3 つのデータセットを用いた評価において、二値分類で平均 98.27%の精度、多クラス分類で平均 96.92%の精度を達成した。特に、パーソナライズされた連合学習を採用することで、クライアント間のデータ分布の偏りを効果的に吸収し、少数のラベル情報だけでも高い検知性能と収束速度を実証している。

   一方で、グラフニューラルネットワークと対比学習の組み合わせがもたらす「極めて重い計算負荷」と「CUDA のメモリ不足」の問題を率直に報告している。この計算コストを回避するため、本実験ではフルデータセットを使用できず、全体のわずか 2%にまでデータをダウンサンプリング評価せざるを得なかったと述べている。本実験は高性能な GPU（RTX 3090 × 2）を搭載した環境で実施されており、これほど計算・メモリ要件の重いグラフベースのモデルが、リソース制約の厳しい実際の IoT エッジデバイス上で動作するかどうかは極めて疑わしい。

7. **AGAT-FL（Sanjalawe et al.、 2025）** [63]：連合学習ベースの IoT 侵入検知システムにおいて、モデルポイズニング攻撃の脅威を軽減するため、適応型グラフ注意ネットワーク（GAT）を用いたフレームワークを提案した。N-BaIoT および CIC-ToN-IoT データセットを用いた評価において、クライ

アントの更新履歴と振る舞いに基づいて動的に信頼スコアを割り当てる手法を採用した。さらに、ハイブリッド CNN-GRU モデルによる時空間的な異常検知と、マハラノビス距離を用いたフィルタリングを組み合わせることで、悪意のある更新を効果的に除外しつつ、N-BaIoT で精度 94.01%、F1 スコア 93.75%、CIC-ToN-IoT で精度 91.02%、F1 スコア 91.15%という優れた検知性能を達成した。また、悪意のある更新を平均 22〜28%拒否することで通信オーバーヘッドを削減し、SHAP や LIME などの XAI 技術を統合して異常分類の根拠を可視化することに成功している。

一方、GAT の学習やグラフ構築に伴う高い計算複雑性が、リソースが極めて限られた IoT エッジノードへの展開を困難にする懸念が挙げられている。また、非同期なクライアント参加の処理において、ネットワークに断続的にしか接続できないデバイスが、その更新頻度の低さから信頼スコアにおいて不当にペナルティを受けるという評価上の課題が残されている。さらに、現在の評価は 2 つの既存ベンチマークデータセットに留まっており、未知の攻撃タイプやハードウェアの差異、マルチテナントネットワークなどが混在する、より多様で予測不可能な現実の IoT 環境への汎化能力の実証が今後の重要な研究課題とされている。

8. **FedDGNN（Ghazal et al.、 2025）** [64]：医療分野などで見られる異種混在型の IoT システムにおいてリアルタイムな侵入検知を実現するため連合学習と深層グラフニューラルネットワークを統合した新しいアルゴリズムを提案した。ネットワーク内のデバイス群が織りなす複雑な関係性や通信トラフィックの異常を深層グラフニューラルネットワークを用いて学習しつつ連合学習によって生データをデバイス内に保持したままモデル更新のみをクラウドへ送信する仕組みを採用している。システム全体の検証にはネットワークシミュレータやクラウド基盤を組み合わせた simulation 環境が構築され実際の医療センサーから得られたデータストリームを用いて異常検知性能が評価された。実験の結果提案モデルは従来のサポートベクターマシンやランダムフォレストといった機械学習手法を凌駕し最高で 95.3%の分類精度および 93.3%の F1 スコアを達成した。これにより分散型ネットワークにおける機微なプライバシーデータを保護しつつ高度な脅威をリアルタイムで特定できる能力が実証されている。

   一方、特に深層グラフニューラルネットワークと連合学習の組み合わせは極めて計算負荷が高くリソースの限られた末端のデバイスにそのまま導入することは非常に困難であると認めている。この課題を克服するためにはモデルの刈り込みや知識蒸留あるいは量子化といった手法を用いて大幅な軽量化を図る必要があると述べている。さらに大規模なシステムにおける通信遅延を最小限に抑え真のリアルタイムデータ処理を実現するためにはエッジコンピューティングとの深いレベルでの統合が不可欠であり運用要件を満たす持

続可能なアーキテクチャの構築が今後の重要な研究課題として残されている。

9. **FedGAT（Wu et al.、 2024）** [65]：ログを時系列に並べ、ログ密度に基づいてグラフを構成し、GAT の注意でノード間インタラクションの重要度を学習することで、部門横断・階層横断の攻撃検知を狙う。システムが記録したネットワークトラフィック情報を時系列順に整理し隣接するタイムスタンプ内のログ数をシステムの振る舞い密度として定義することで時系列のログシーケンスからグラフ構造を構築している。このアプローチにより各ノード間の情報の流れや相互作用を注意機構を用いて評価し内部ネットワーク通信の精度を向上させることに成功した。

   NSL-KDD データセットを用いた実験において提案手法は従来のランダムフォレストやディープコンボリューショナルニューラルネットワークなどの検知手法と比較して極めて高い分類性能を示した。具体的には DoS や Probe あるいは U2R といった多様な攻撃シナリオに対して最高で 99.99%に達する検知精度および F1 スコアを達成しデータセキュリティを確保しつつ分散デバイス上で協調的にモデルを訓練できることを実証している。また異なる層における計算負荷とデータ応答の変動を浮き彫りにし大規模データセットにおける更新頻度の最適化に関する重要な洞察も提供した。

   一方で、部門やドメインをまたぐ複雑なネットワーク環境では攻撃者が大規模言語モデルを利用して強力な検知回避能力を持つ攻撃コードを短期間で生成する可能性があり現在の枠組みでは適時性やスケーラビリティに重大な課題が生じると指摘している。さらに提案されたアプローチがより大規模なネットワークインフラストラクチャに直面した際のスケーラビリティは完全には検証されておらず計算効率を維持したままシステム性能を最適化する手法の確立が今後の重要な研究方向として残されている。

10. **CAN-FedGNN（Zhang et al.、 2023）** [66]：車載ネットワークにおける暗号化や認証の欠如に起因するメッセージ注入や改ざんなどの多様なサイバー攻撃を同時にかつ高速に検知するためグラフニューラルネットワークと連合学習を組み合わせた新しい侵入検知システムを提案した。連続する CAN メッセージのストリームを有向属性グラフとして構築しノードの属性にはメッセージのデータ内容を含めエッジの属性には特定のメッセージ識別子ペアの出現頻度を割り当てるという手法を採用している。攻撃データの収集が困難で正常データが極端に多いという不均衡なデータ分布に対処するため提案手法は二段階の分類器カスケード構造を導入した。第一段階では一クラス分類を用いて異常検知を行い異常と判定されたデータのみを第二段階の多クラス分類器へ送る。この第二段階には未知のクラスからの新規な異常に対処するためのオープンマックス層が組み込まれている。さらに車両の状態や運転シナリオの違いによるデータの非独立同分布性に対応しつつユーザー

のプライバシーを保護するために連合学習を用いて汎用的なモデルを構築した。データを用いた評価では毎分回転数や速度情報の注入攻撃に対して最高で 100%の精度と F1 スコアを達成し一つのメッセージグラフあたりの平均予測時間を 3 ミリ秒に抑えるなど極めて高いリアルタイム性と検知性能を実証した。

一方、特にリプレイ攻撃とスプーフィング攻撃についてはランダムに生成されたメッセージが過去に観測された正常なメッセージと酷似する場合があり互いに誤分類されやすいという問題が指摘されている。さらにファジー攻撃が統計的なメッセージシーケンスに明確な変化をもたらさない場合も正確な識別が困難となる。加えて分類精度を上げるためにメッセージの収集区間を長く設定するとデータ収集にかかる時間が 100 ミリ秒を超えてしまいリアルタイム性に重大な悪影響を及ぼすというトレードオフの存在も今後の実運用に向けた未解決の課題として残されている。

4.3.5 G5：Autoencoder / 表現学習系

G5 はオートエンコーダや生成敵対的ネットワークなどの表現学習を活用しネットワークトラフィックの正常な振る舞いを学習することで侵入検知を行うものである。正常なデータのみを用いてモデルを訓練し入力データの再構成誤差や潜在空間における距離を異常スコアとして算出するメカニズムがメインとなる。近年この手法はラベル付けされた攻撃データが不足しがちな IoT 環境において未知の攻撃やゼロデイ脅威を検知する能力が非常に高い点から大きな注目を集めている。連合学習と組み合わせることで各デバイスが自身の正常な通信パターンをローカルで学習し生データを共有することなくグローバルな異常検知モデルを構築できるためプライバシー保護と未知の脅威への適応を両立する強力な枠組みとなっている。

一方、これらの教師なし学習または半教師あり学習アプローチは実運用に向けて極めて深刻な課題を抱えている。最大の弱点は異常と正常を分けるしきい値の設定が非常に困難でありかつハイパーパラメータの微調整に極度に依存する点である。連合学習環境においては各クライアントが観測する正常なトラフィックの分布が全く異なる非独立同分布の性質を持つためグローバルモデルで平均化された潜在空間や単一のしきい値を強引に適用すると誤警報や見逃しが急増する。

各論文概要

1. **FedGAN-ID（Chen et al.、 2025）** [67]：高度に接続された知能型車両の車載ネットワークにおけるサイバー攻撃の脅威に対処するため敵対的生成ネットワークと連合学習を組み合わせた新しい侵入検知モデルを提案した。車載ネットワークの主流であるコントローラエリアネットワークが暗号化や

認証メカニズムを欠いており攻撃に対して非常に脆弱であるという問題に着目している。従来の集中型データ処理や限られた実際の攻撃データに依存するモデルのプライバシー懸念と検知能力の限界を克服するため提案手法は二段階のカスケード検知器を構築した。第一段階では正常なデータ分布を模倣するように訓練された敵対的生成ネットワークを用いて異常を検知し第二段階ではエネルギーベースの分布外検知手法を導入して未知のゼロデイ攻撃を含む多様な脅威を分類する。さらに差分プライバシーを組み込んだ連合学習スキームを採用し各車両がプライバシーを保護された合成データを生成してクラウドサーバにアップロードすることで実際の生データを共有することなくモデルの攻撃分類能力を向上させている。三台の実際の知能型車両を用いた環境および実機ベンチマークにおける評価においてサービス拒否やなりすまし等の五種類の攻撃に対してすべての評価指標で 94.66%以上の性能を達成し車両の処理ユニット上でわずか 1.96 ミリ秒以内で正確に攻撃を検知できることを実証した。

一方、差分プライバシーの導入によるプライバシー保護レベルとモデルの精度との間に明確なトレードオフが存在しプライバシー予算を厳しく設定すると検知精度が約 1%低下するなどデータの有用性が損なわれることが確認されている。連合学習のプロセスにおいて各通信ラウンドでのモデルパラメータ送信に加えて合成データやロジットベクトルの共有に合計で約 3109 メガバイトという極めて大きな通信オーバーヘッドが発生することが指摘されている。

2. **FL-DAE（2025）** [68]：IoT 環境における分散サービス妨害攻撃を緩和するため深層オートエンコーダと連合学習を統合した新たな侵入検知フレームワークを提案した。計算資源の限られた末端デバイスの負荷を軽減するため深層オートエンコーダを用いてデータの次元削減と特徴抽出を行い、再構成誤差に基づく異常検知モデルを各デバイス上で構築している。さらに Non-iid かつ不均一なデータがもたらす学習の不安定性を克服するため各通信ラウンドにおいて一部のクライアントのみをランダムに抽出する部分選択手法と集約前にローカルモデルを再訓練する手法を組み込んでいる。実際のボットネットトラフィックを含むデータセットを用いた環境での実験においてモメンタム付きの連合平均化アルゴリズムを採用した場合に従来の標準的な連合平均化アルゴリズムを上回る性能を示した。具体的には九つのデバイスデータを用いた評価で 93.95%の曲線下面積および 93.10%の F1 スコアを達成し誤警報率を 2.19%に抑えつつ一ラウンドあたりの平均学習時間を約 49 秒に短縮するなど高い検知精度と効率的な計算資源の活用を実証した。

一方、具体的には選択されたすべてのデバイスが指定された回数の確率的勾配降下法による反復を完全に実行できると仮定しているが実際の異種混在ネットワークでは処理能力の低い遅延デバイスが多数存在しそれらが学習全

体の収束速度を著しく阻害する問題がある。

3. **FL-IDS（Olanrewaju-George & Pranggono、 2025）** [69]：教師なし AE（再構成誤差に基づく異常スコア）と教師あり DNN を同一枠組みで比較し、連合学習で更新のみ共有することでプライバシーと分散運用を両立する。九つの異なるボットネット感染デバイスのトラフィックを記録したネットワークデータセットを用いた環境での実験において提案された教師あり深層ニューラルネットワークモデルは従来の集中型機械学習手法を大きく上回る性能を示した。具体的には精緻なハイパーパラメータの最適化を経た結果として 90.39%の分類精度と 99.99%の適合率および 93.12%の F1 スコアを達成し生データを一切外部に送信することなく高度なボットネット攻撃を正確に識別できることを実証した。また教師なし手法との比較を通じてネットワークの特性に応じた最適な学習アプローチの選択に関する重要な洞察を提供している。

   一方、データセットの特性やモデルの複雑さに合わせて最適な構成を見つけ出すためのハイパーパラメータ微調整に極めて大きく依存している点を率直に認めている。最適なパラメータを探索するためのランダムサーチ手法は分散した各エッジノードにおいて多大な計算リソースと時間を消費するためリソースが極めて限定的な末端のデバイス群へそのまま展開するには大きな障壁となる。さらに現在の評価は特定の静的なトラフィックデータに基づいており絶えず変化する実際のネットワーク環境において未知の脅威やトラフィックパターンの変動に即座に適応するための軽量化戦略や継続的な最適化メカニズムの確立が今後の重要な研究課題として残されている。

4. **FedMSE（2024）** [70]：この研究では、収縮オートエンコーダ（SAE）とセントロイド（CEN）アルゴリズムを組み合わせたハイブリッド異常検知モデルである SAE-CEN を各ゲートウェイに配置している。各ゲートウェイは、ローカルの正常なトラフィックデータのみを用いてモデルを訓練し、そのパラメータを中央サーバへと送信する。サーバ側では、平均二乗誤差（MSE）を利用した新たな集約アルゴリズムである MSEAvg を導入している。このアルゴリズムは、サーバ側で保持する検証用データセットに対して各ローカルモデルがどれだけ正確に正常データを再構成できるかを評価し、再構成誤差が小さいモデルに対してより大きな重みを割り当てる仕組みである。実際のボットネットトラフィックデータを用いた環境での実験において、データ分布が極端に偏る Non-iid の設定下でも、提案手法は 97.30%という極めて高い AUC スコアを達成した。さらに、全体のわずか 50%のゲートウェイのみが学習に参加するリソース制約の厳しいシナリオにおいても、一貫して高い検知性能と収束の安定性を維持できることが実証されている。

   一方、学習段階において異常データが一切存在しないという前提があるため、

モデルのハイパーパラメータを最適にチューニングすることが非常に困難であり、経験則に基づいた一般的な値に依存せざるを得ない点が挙げられている。また、優れたモデルを優先して集約するMSEAvgアルゴリズムは、サーバ側で受信したすべてのローカルモデルに対して開発用データセットを用いた推論と誤差計算を毎ラウンド実行する必要があるため、サーバ側の計算負荷と処理時間を必然的に増大させるという課題がある。

5. **FL-LSTM-AE（2024）** [71]：この研究は、5Gコアネットワークから収集される多変量時系列データに対して、LSTMオートエンコーダを用いて異常な振る舞いを学習し、再構成誤差に基づいて侵入を特定するアプローチを採用している。集中型学習のベースラインモデルでは適合率92.19%、再現率97.50%、F1スコア94.78%、精度93.12%という非常に高い性能を示した。これをフレームワークを用いたFL環境に拡張した場合でも、精度89.74%およびF1スコア92.19%を達成し、十分な検知能力を維持している。さらに本研究の大きな特徴として、悪意のあるクライアントが偽のデータを注入するデータポイズニング攻撃がモデルに与える脅威を具体的に検証し、KrumやBulyanといった複数のロバストな集約手法を比較評価することで、computation環境および実機データ収集環境においてノイズ注入やバックドア攻撃に対する防御メカニズムの有効性を実証している。

   一方、データポイズニング攻撃に対して高い耐性を示す高度な集約アルゴリズムは、各通信ラウンドにおいて悪意のある更新を特定し除外するために複雑な計算を要求するため、リソースに制約のあるエッジ環境でのスケーラビリティに懸念を残す。また、LSTMオートエンコーダの再構成誤差に基づいて異常を定義する手法は、ネットワークの正常なトラフィックパターン自体が極めて動的かつ急速に変化する現実の5G環境においては、誤警報を防ぐためのしきい値の適応的な調整が困難になる可能性がある。

6. **CAE-Agri（Khacha et al.、 2024）** [72]：農業IoT（AIoT）インフラストラクチャを保護するため、畳み込みオートエンコーダ（CAE）と連合学習を統合した新しい侵入検知システムを提案した。Conv AEで局所パターン（近傍特徴のまとまり）を畳み込みで圧縮し、復元誤差を異常スコアとして用いる。画像的な局所相関に強いCAEの性質を、農業IoTのセンサ/通信特徴に転用する狙いで、ラベル不足や未知異常に対して有利になり得る。最新の大規模データセットであるCICIoT-2023を用いたcomputation環境での実験において、独立同分布（IID）データの条件下では、提案モデルは中央集権型の学習手法を上回る99.39%の精度、99.99%の再現率、および99.40%のF1スコアを達成し、スマート農業環境における未知のサイバー攻撃の検知に極めて有効であることを示した。

   一方で、デバイスごとにデータ分布が異なる非独立同分布（Non-iid）の現

実的なシナリオにおいてはモデルの性能が著しく低下し、特に参加クライアント数が増加すると適合率が 72.77%まで急落するなど、クライアント間のデータ不均一性が異常検知の正確性に与える悪影響が顕著に現れている。また、現在の FL プロセスではクライアントとサーバ間でやり取りされる CAE の重みパラメータ自体に対する暗号化技術が導入されていない。そのため、通信経路上の傍受やパラメータのリバースエンジニアリングによるプライバシー侵害の脆弱性が残されており、より強固な暗号化の統合による安全性の確保が実用化に向けた不可欠な研究課題として指摘されている。

7. **AE-IoT（2024）**[73]：IoT デバイスが生成する膨大なデータに伴うプライバシーおよびセキュリティの懸念に対処するため、深層オートエンコーダ（AE）と連合学習を統合した分散型のボットネット攻撃検知手法を提案した。この研究は、ネットワークの末端にあるデバイスからデータを外部へ転送することなく、ローカルで異常検知モデルを学習させるアプローチを採用している。検知器のコアとして、三つの線形層と正規化層からなるエンコーダおよびデコーダを持つ深層 AE を構築し、正常なトラフィックデータのみを用いてモデルを訓練する。推論時には、入力データの MSE を算出し、事前に設定されたしきい値を超過したパケットストリームを異常として判定する。実際の Mirai および BASHLITE ボットネットに感染した 9 種類のデバイスデータを用いた環境での実験において、適応的最適化アルゴリズムである FedAdam を用いて重みを集約した結果、モデルは 99.831%の精度、100%の再現率、および 99.888%の F1 スコアを達成した。これは同じ構造を用いた従来の集中型機械学習モデルによる精度 90.34%と比較して、攻撃検知性能が大幅に向上していることを示しており、分散環境において生データを保護したまま高いセキュリティ能力を確保できることを実証した。

   一方で、著者らの評価手法やシミュレーションの枠組みから、実運用環境への適用に向けた明確な限界も読み取れる。本手法における異常判定は、事前にシミュレーションのエポックを消費し尽くした後に生成される固定的な再構成誤差のしきい値に完全に依存している。しかしながら、トラフィック量やデバイスの挙動が絶えず変動する実際の多様なネットワークにおいては、正常な通信パターン自体が時間とともに変化するため、シミュレーション環境で静的に決定されたしきい値の品質をそのまま維持することは極めて困難である。さらに、この固定されたしきい値を超過したものを単純に数え上げて偽陽性を評価する現在のアプローチでは、未知の環境変動が起きた際に誤警報が急増するリスクがあり、しきい値の動的かつ自律的な調整メカニズムの欠如が、実際の動的ネットワークへ展開する上での大きな課題として残されている。

4.3.6 G6：Transformer / ViT 系

G6 には自己注意機構を備えた transformer モデルや視覚 transformer を侵入検知に応用し、複雑なネットワークトラフィックの長期的な依存関係や広域的な特徴を捉える手法を使用する論文である。

このアプローチは、パケットのシーケンスを自然言語の単語の並びのように扱ったり、トラフィック全体を画像に変換して視覚的なパターンとして処理したりすることで、従来の軽量モデルでは見逃されがちな微細な悪性トラフィックや高度に分散化されたサイバー攻撃を極めて高い精度で分類する能力を持つ。

しかしながら、これらの巨大で複雑なアーキテクチャは実運用に向けて致命的な運用ギャップを抱えている。最大の問題点は、transformer が本質的に要求する莫大なパラメータ数と計算リソースであり、これが連合学習における各通信ラウンドの転送量とエッジデバイスの処理負荷を劇的に増大させる点である。多くの研究はパラメータ効率化手法や量子化を導入して軽量化を試みているものの、実際の分散サービス妨害攻撃によって引き起こされる深刻な帯域枯渇やパケット損失が発生するネットワーク条件下において、このような重厚なモデルの更新がタイムアウトせずに成立するかどうかの実証はまだ充分ではないことが存在する。

各論文概要

1. **FLoV2T（Zeng et al.、 2025）** [74]：AIoT 環境における悪性トラフィック分類を目的として、視覚 transformer（ViT）と低ランク適応（LoRA）技術を統合した。生のネットワークトラフィックを個々のパケットに基づいてパッチに分割し、それを二次元の画像フォーマットに変換して入力する手法を採用している。これにより、自然言語処理や画像認識で実績のある事前学習済み ViT の強力な自己注意機構を最大限に活用し、従来の統計的手法では捉えきれない複雑な攻撃パターンを抽出している。エッジデバイスの厳しい計算資源制約を克服するため、提案手法は LoRA を用いた局所的なパラメータの微調整メカニズムを導入した。

   CICIDS2017 および CICIDS2018 データセットを用いた実験の結果、提案手法は平均 97.26%の精度と 96.99%の F1 スコアを達成し、既存のベースライン手法を大きく上回った。また、LoRA の導入により、高い分類性能を維持したまま学習および通信の対象となるパラメータ数を約 64 分の 1 にまで大幅に削減することに成功し、軽量なエッジ展開の可能性を示した。

   一方、実際の AIoT ネットワーク環境に特有の高遅延問題や不安定な接続といった通信の不確実性が、連合学習の同期的または非同期的な更新プロセスに影響を与え、モデルの収束とシステム全体の安定性を著しく阻害する可能性があると述べている。さらに、パラメータ数を削減したとはいえ、ViT アーキテクチャ自体が本質的に持つ重い計算負荷は完全に解消されたわけで

はなく、バッテリ駆動の AIoT デバイスにおいてはエネルギー効率が依然として最大の懸念事項であると説明している。

2. **ResVGG-SwinNet（Alshdadi et al.、 2025）** [75]：この研究では、複雑なトラフィックの特性を網羅的に捉えるために、特徴抽出を担う ResNet、特徴を精緻化する VGGNet、そして文脈依存関係を学習する Swin-Transformer を直列に配置する極めて重厚な設計を採用している。さらに、クラス不均衡やデータ冗長性に対処するため、動的比例クラス調整（DPCA）や双対適応セレクタ（DAS）といった複数の高度な前処理および特徴最適化手法をパイプラインに組み込んでいる。これらの複雑な機構を FL 環境に適用し、単一の高性能な計算サーバー上で各クライアントの資源制約や通信遅延、パケット損失を論理的に模倣した環境において評価を行った。CIC-DDoS2019、UNSW-NB15、および IoT23 という三つのデータセットを用いた実験の結果、提案モデルは 99.0%の精度、99.3%の曲線下面積、および 98.5%の F1 スコアを達成し、誤警報率を 2.5%に抑えつつ、一ラウンドあたり 2.8 ミリ秒という推論時間で分散データからの高精度な脅威識別が可能であることを示した。

   一方、提案モデルは三つの重厚な深層学習ネットワークを重ね合わせ、さらに多段の前処理を組み込んでいるため、学習率やドロップアウト率などのハイパーパラメータの微小な変動に対してモデルの精度が極めて不安定に波打つことが感度分析の三次元可視化において明確に示されている。また、シミュレーション上の計算では推論時間をミリ秒単位に抑えられたと報告しているものの、著者らは実際の資源制約が厳しい IoT エッジデバイスへのリアルタイムな実装や展開検証が今後の課題であると率直に認めている。

3. **SecureFogDL（Zheng & Pu、 2025）** [76]：医療 IoT システムにおいて継続的な患者監視や個別化治療の安全性を確保するため、フォグコンピューティングと連合学習、および BERT を統合した新しい侵入検知フレームワークである SecureFogDL を提案した。

   この研究は、医療データが持つ極めて高い機密性を保護するため、生データを中央クラウドに送信せず、ネットワークのエッジに近いフォグノード上で分散的にモデルを学習させるアプローチを採用している。特徴量の抽出には、多次元かつ複雑なトラフィックデータを低次元の潜在ベクトルに圧縮するオートエンコーダを用い、リソースの限られたデバイスでの計算負荷を大幅に軽減している。その後、抽出された特徴シーケンスは BERT ベースの分類器に入力され、自己注意機構によって時間的および空間的な依存関係を捉えることで、正常なトラフィックと DDoS 攻撃を正確に識別する。TON-IoT データセットを用いた環境での評価において、提案フレームワークは 98.350%の精度と 98.349%の F1 スコア、および 0.998 の曲線下面積を達成した。さらに、ローカルでの学習を 3 エポック行うごとに通信を実施する設定により、1 ラ

ウンドあたりの通信コストを 8.2 メガバイトに抑えつつ、1 レコードあたりの平均推論遅延を 0.021 秒に短縮し、リアルタイム性が求められる医療環境での有効性を実証した。

一方、現在の評価は正常トラフィックと攻撃トラフィックを区別する二値分類のベースライン検証に留まっている点である。実際の医療環境では、データインジェクションやランサムウェア、さらには推論ベースの脅威など、より広範で多様な攻撃ベクトルが存在するため、多クラス侵入検知への拡張が今後の開発における極めて重要な方向性であると率直に認めている。オートエンコーダによるデータ圧縮を適用しているものの、transformer 構造自体が本質的に持つ重い計算要件は完全に解消されておらず、非常に制約の厳しいフォグ環境における効率をさらに高めるためには、DistilBERT のようなより軽量化されたモデルアーキテクチャの導入や、差分プライバシーとの組み合わせによる堅牢性の向上が不可欠な課題として残されている。

4. **Access-side Transformer + Personalized FL（Luo et al.、 2024）** [77]：Luo らによる 2024 年の研究は、第 5 世代および第 6 世代移動通信システム（5G/6G）の IoT 環境において、攻撃の標的となるサーバ側でのリソース枯渇問題を回避するため、ネットワークのアクセス側となる入口に分散配置された検知点において悪性トラフィックを識別する、transformer と個人化連合学習を組み合わせた新しいフレームワークを提案した。

   研究では、各検知点におけるトラフィックの統計的特徴を時系列データとして入力し、強力な特徴抽出能力を持つ transformer モデルを用いて悪性トラフィックの複雑なパターンを学習する。さらに、異なるデバイスが接続される検知点ごとにトラフィックの分布が大きく異なる Non-IID の課題に対処するため、共通層と個人化層を分離したアルゴリズムを導入した。共通層のパラメータは中央クラウドを介して全検知点間で共有および集約されることでグローバルな攻撃の知識を取り込み、個人化層のパラメータは各ローカルデータのみで独立して更新される。N-BaIoT データセットを用いた評価では、三つの検知点が協調するシナリオにおいて平均 99.2%の精度および F1 スコアを達成し、パラメータの共有のみで中央集権型モデルに匹敵する高い検知性能と、各検知点の局所的なトラフィック特性への優れた適応性を実証した。

   一方、著者ら自身の考察からも、このフレームワークを実際のネットワーク入口に展開する際の深刻なジレンマが読み取れる。transformer アーキテクチャはその優れた自己注意機構によって極めて高い分類精度を実現するものの、本質的に膨大な計算リソースとメモリを要求する極めて重いモデルである。著者らはネットワーク入口の検知点が持つ計算能力は非常に限定的であることを率直に認めており、このようなリソース集約型のモデルを制約の厳しいアクセス側エッジサーバにそのまま配備することは運用上極めて困難であると指摘している。さらに、学習の枠組みにおいて、巨大なモデルの更新

パラメータをネットワークの入口と中央クラウドの間で頻繁にやり取りすることは、帯域幅の枯渇に直結する。したがって、高い検知性能を維持したままモデルを大幅に軽量化し、計算コストをエッジの許容範囲内に収めるための最適化が、実運用に向けた極めて重要な未解決課題として残されている。

5. **securityBERT（Adjewa et al.、2024）**[78]：ネットワークフローを BERT 系 Transformer で扱い、文脈理解（トークン間依存）を利用して異常（未知）を早期にあぶり出す「Federated anomaly-based IDS」を構築する。エッジ適用の現実解として、モデルの軽量化・圧縮（線形量子化）を組み合わせ、IID/Non-iid の両条件で学習挙動を比較しながら、端末上で回る推論時間も提示する点が実装寄りである。ただし、BERT は入力表現（どの特徴をどう並べてトークン化するか）に性能が強く依存し、運用環境が変わった場合に“ 文脈 ”が崩れて精度が落ちる可能性がある。量子化はサイズを下げる一方で境界事例に効きやすく、誤警報と見逃しのバランスを含めた長期評価（ドリフト/概念変化）が今後の焦点となる。

   5G エコシステムにおける IoT デバイスのセキュリティ課題に対処するため、大規模言語モデル（LLM）のアーキテクチャをエッジデバイス向けに最適化した軽量な侵入検知システムである securityBERT と、連合学習を統合したアプローチを提案した。

   研究では、従来の機械学習手法が未知の攻撃に対して限界を持つことや、標準的なトランスフォーマーモデルが資源制約の厳しい端末には重すぎるという問題に着目している。そこで、双方向エンコーダ表現トランスフォーマーを基盤としつつ、エンコーダ層を四層に削減し、隠れ層の次元数や注意ヘッド数を最適化することで、高い文脈理解能力を維持したままパラメータ数を大幅に縮小した。さらに、生のネットワークトラフィックデータをバイトレベルのバイト対エンコーディングを用いてトークン化し、学習後の線形量子化技術を適用して計算負荷を低減させている。computation 環境およびラズベリーパイ等の実機ベンチマークを用いた評価において、集中型学習で 97.79% の精度を達成した。また、十台のクライアントを用いた分散シナリオにおいて、独立同分布（IID）データで 97.12%、データの偏りがある非独立同分布（Non-IID）データでも 96.66%の高い検知精度を維持しつつ、モデルサイズを 30.38 メガバイトまで圧縮できることを実証した。

   一方、エッジ端末上での単一パケットの推論時間は 0.45 秒と報告されているが、これはミリ秒単位の超低遅延が要求され、毎秒膨大なトラフィックが飛び交う実際の 5G ネットワークにおけるリアルタイムな監視としては処理遅延が大きすぎる懸念がある。分散環境におけるデータの偏りに関して、著者ら自身が指摘するように、各クライアントが保持するデータクラスが豊富で均等に近い理想的な条件下では高い精度が得られるものの、そのような設定は実際の運用環境では非現実的である。現実の高度に偏ったトラフィック分

布や厳しいリソース制限下において、推論速度と分類精度のトレードオフをいかに克服するかが、LLM ベースのセキュリティソリューションを社会実装する上での重要な未解決課題として残されている。

4.3.7 G7：システム／ネットワーク機構主導

G7 は侵入検知器の内部アーキテクチャの改良よりも、帯域幅の割り当て、参加ノードの信頼性管理、悪意のある更新の排除、および検知後のトラフィック緩和といった、システムおよびネットワーク機構の運用最適化に主眼を置いた手法がメインとなる。

このアプローチは、IoT 環境特有の通信制約や運用上のボトルネックに対処するために設計されている。例えば、DDoS 攻撃による通信帯域の枯渇を防ぐためのゲーム理論に基づくリソース配分や、ブロックチェーンとスマートコントラクトを活用したマルチドメイン間の信頼評価、あるいはポイズニング攻撃を行う悪性クライアントを特定して排除する動的なスコアリング機構などが含まれる。これらのネットワーク主導の最適化により、データ分布が不均一な環境や制約の厳しいエッジネットワークにおいても、連合学習プロセス自体の崩壊を防ぎ、安全で協調的なセキュリティ基盤を持続させることを目指している。

しかしながら、これらのシステム管理に基づくアプローチは、実運用に向けて極めてパラドックス的な課題を抱えている。最大の弱点は、信頼管理や帯域最適化あるいは暗号化処理といった複雑な運用機構を追加すること自体が、エッジデバイスおよびネットワーク全体に多大な計算オーバーヘッドと通信遅延をもたらす点である。

各論文概要

1. **EGDA（Xu et al.、 2025）** [79]：EGDA はモバイルエッジコンピューティング（MEC）における連合学習の帯域幅リソースを標的とした DDoS 攻撃を緩和するため、進化ゲームとダブルオークションを統合した新しい帯域幅割当フレームワークを提案し、限られた無線帯域を悪意のある攻撃者が大量に要求して学習プロセスを遅延あるいは停止させるという深刻なネットワーク課題に着目している。この問題に対処するため、提案フレームワークは二つの階層で構成されている。第一の階層では、サービスプロバイダとユーザ間のユーザ帯域幅割当に対して進化ゲーム理論を適用し、レプリケータ動力学を用いて各ユーザへの帯域割当を反復的に調整することで、限られた情報下での FL 完了遅延を最小化する。第二の階層では、ネットワークオペレータと SP 間のサーバ帯域幅割当に対してダブルオークションメカニズムを導入し、参加者のプライバシーを保護しつつ真実の要求申告を促し、市場全体の社会的余剰を最大化する。MATLAB を用いたシミュレーションにおいて、攻

撃者が通常の数倍の帯域を要求する極端な状況でも、提案手法は他のヒューリスティックな割当手法と比較して通信遅延を大幅に削減し、安定した学習の収束と高い社会的余剰を維持できることを実証した。

一方、ユーザ側とサーバ側の帯域割当を個別に最適化する現在の構造では、UBA フェーズで設定された遅延制約が SBA フェーズにおける割当の自由度を制限してしまい、システム全体を統合的に最適化する単一のアプローチに比べて、最終的な社会的余剰の最大化において準最適な解に陥る可能性があると認めている。また、フレームワークはあくまで DDoS 攻撃による資源枯渇の影響を緩和することに特化しており、ネットワーク内の悪意のあるトラフィックを積極的に特定する検知機能は備わっていないのが課題である。

2. **FedLAD（Fotse et al.、 2025）** [80]：SDN において制御プレーンを担う単一のコントローラが DDoS 攻撃の標的となりやすい問題に対処するため、マルチコントローラアーキテクチャと連合学習を統合した新たな侵入検知フレームワークを提案した。単一の集中型コントローラに依存する従来の防御手法が、大規模ネットワークにおけるスケーラビリティの要求を満たせず、計算資源の枯渇を引き起こすという限界に着目している。提案手法では、ネットワークを複数のドメインに分割し、それぞれのローカルコントローラが自身の管理領域でネットワークトラフィックの統計情報を収集する。各ローカルコントローラは、XGBoost アルゴリズムを用いて局所的な DDoS 検知モデルを学習させ、そのモデルパラメータのみをルートコントローラへ送信する。ルートコントローラはこれらを集約してグローバルモデルを構築し、再びローカルコントローラへ配信することで、機密性の高い生データを交換することなくネットワーク全体の防御能力を向上させる。CICDoS2017 や CICDDoS2019 などのデータセットを用いたオフライン評価では、集中型モデルに匹敵する平均 98.33%の精度を達成した。さらに、実際のネットワークトラフィックをエミュレートした環境でのリアルタイム評価において、15 台のローカルコントローラを協調させた場合に最高 97.64%の精度を達成し、集中型アプローチと比較してコントローラの CPU 使用率やメモリ消費量を大幅に分散および削減できることを実証した。

一方、異なるローカルコントローラ間でトラフィックデータの Non-iid 性が強い場合、標準的な連合平均化アルゴリズムでは精度が著しく低下することが実験により示されており、精度を維持するためにはランキングクライアントなどのより複雑な集約技術に依存せざるを得ない点が挙げられる。このシステムに組み込まれた帯域幅監視およびトラフィック制御モジュールは、事前に設定された静的な上限値を超過したフローを検知して制限するというルールベースの仕組みに一部依存している。トラフィックの変動が極めて激しく予測困難な現実の大規模 SDN 環境においては、このような静的なしきい値への依存が正規の通信を誤って遮断するリスクを持ち、より動的なトラフィッ

ク制御メカニズムの統合が今後の重要な研究課題として残されている。

3. **Hybrid EFL（Va & Rajkumar、2025）**[81]：5G の IoT エッジコンピューティング環境において、DDoS 攻撃を効率的かつプライバシーを保護しつつ検知するため、合成少数派オーバーサンプリング手法（SMOTE）および Tomek リンクを活用したハイブリッドなアンサンブル連合学習フレームワークを提案した。この研究は、限られた計算資源しか持たないエッジデバイスが直面する通信オーバーヘッドとデータ不均衡の課題に着目している。提案手法では、ランダムフォレスト、勾配ブースティング、および XGBoost という三つの異なる軽量な機械学習モデルを各クライアントに配置し、SMOTE と Tomek リンクを組み合わせて少数派の攻撃クラスデータを補完およびダウンサンプリングすることでデータ分布を均等化する。さらに、本フレームワークの最大の特徴として、従来の連合学習のように巨大なモデルの重みパラメータ全体を送信するのではなく、各モデルが出力するクラス確率ベクトルのみを中央サーバへ送信する設計を採用している。サーバ側では、ソフトボーティングおよびロジスティック回帰をメタ学習器として用いるスタッキング手法によってこれらの確率を集約し、最終的な攻撃分類を行う。ラズベリーパイや産業用 5G ルータなどの実機から収集されたデータセットを用いた環境および実機データ収集環境でのシミュレーションにおいて、スタッキングを用いた EFL モデルは 97.90%の精度と F1 スコア、および 0.9993 の AUC を達成した。同時に、モデルの重みではなく確率のみを共有することで通信コストを 62.64 メガバイトに抑え、検知遅延を 190.10 ミリ秒まで短縮するなど、リアルタイム性と資源効率の大幅な向上を実証した。

   一方、現在の評価はクラウドベースの計算プラットフォームである Google Colab 上のシミュレーション環境に完全に依存しており、物理的なエッジデバイス上への本格的な実装や実機ネットワーク間での通信オーバーヘッドの動的検証は今後の課題として残されている。今後ネットワークが 6G へと進化し、デバイスの接続密度やトラフィックの複雑さがさらに爆発的に増大する環境を見据えた場合、現在の静的な集約アプローチでは限界がある。そのため、クライアントの信頼度やネットワーク状況に応じた適応的な重み付け戦略や、異常発生時の自己修復機能を持つより高度な FL メカニズムへの拡張が、実用化に向けた不可欠なステップとして指摘されている。

4. **Al-Ghadi et al. 2025** [82]：Al-Ghadi らによる 2025 年の研究は、IoT ネットワークにおける DoS 攻撃の検知精度向上と計算資源の最適化を目的として、アンサンブル特徴選択と深層学習モデルを組み合わせた。エッジデバイスの限られたメモリや計算能力を考慮し、トラフィックデータをそのままモデルに入力するのではなく、分散分析やカイ二乗検定などの複数の特徴選択手法を用いて重要な通信特徴のみを抽出する前処理段階を導入している。特徴選択によって次元削減されたデータは、畳み込み CNN や LSTM、GRU な

どの深層学習アルゴリズムへ入力され、各クライアント上でローカルモデルが学習される。その後、FedAvg を用いて中央サーバで重みが集約される。Google Colab の CPU 環境を用いて、再帰的特徴消去（RFE）と GRU 分類器の組み合わせが最も高い性能を示し、99.91%という極めて高い精度を達成することで、特徴選択が分類性能の向上とモデルの軽量化に寄与することを実証した。

一方、分散協調によるスケーラビリティを最大の利点として謳いながら、実際の実験はわずか 3 台から 5 台のクライアントを論理的に分割した極小規模な設定に留まっており、何千ものデバイスが参加する現実のネットワークにおける拡張性が全く検証されていない。また、通信オーバーヘッドやデバイスの資源制約を課題として挙げているにもかかわらず、実験はクラウド上の計算環境における論理的な模擬に完全に依存しており、実際のネットワーク帯域幅の消費や、パケット損失に伴う遅延、各デバイスの物理的な消費電力といった運用指標が一切測定されていない。

5. **CAFNet（Tayeen et al.、 2023）**[83]：IoT デバイスの膨大なトラフィックデータによる通信オーバーヘッドと、異常トラフィックのラベル付きデータが極端に不足する不均衡問題に対処するため提案したも。未知のゼロデイ攻撃を検知するために正常なトラフィックのみを用いて訓練を行う教師なし学習アプローチを採用している。さらに、連合学習における最大のボトルネックである通信コストを劇的に削減するため、分散配置されたエッジデバイスと中央サーバ間でモデルのすべての重みパラメータを転送するのではなく、オートエンコーダの潜在空間表現パラメータのみを抽出して送受信し、サーバ側で FedAvg を用いて集約する独自の圧縮手法を設計した。CICDDoS2019、Bot-IoT、および UNSW-NB15 という三つの公開データセットを用いた環境での交差検証実験において、提案フレームワークは従来の標準的な連合学習手法と比較して通信コストを最大 95%削減しつつ、F1 スコアの低下をわずか 1%から 4%に抑え、高い検知性能を維持できることを実証した。

   一方、具体的には、ローカルモデルから潜在表現パラメータのみを抽出して共有する現在の設計では、エンコーダやデコーダの他のパラメータ層にカプセル化されてエンコードされた重要なトラフィック特徴情報が通信の過程で必然的に失われてしまう。この情報の欠落が原因となり、完全なモデルパラメータを共有するベースライン手法に比べて検知精度や F1 スコアの微小な低下を引き起こすことが実験結果からも明らかになっている。

6. **FLIoT（Yang et al.、 2023）**[84]：IoT における分散型ネットワーク侵入検知が直面する悪意のある参加者によるポイズニング攻撃の脅威に対処するため、連合学習フレームワークに軽量な異常ノード検知メカニズムを統合し、従来の集中型アプローチの課題である高遅延やプライバシー漏えいを克

服する FL の利点を認めつつも、各クライアントがモデルパラメータや学習データに直接アクセスできるという FL 固有のセキュリティ脆弱性に着目している。

特に、攻撃者がローカルの学習データのラベルを意図的に反転させるラベルフリッピング攻撃が、協調学習されたグローバルモデルの精度と信頼性を著しく低下させる問題を指摘した。これらの攻撃の悪影響を緩和するため、提案手法では中央サーバにおいて異常な参加者をフィルタリングし、その更新パラメータをグローバルモデルの集約プロセスから完全に除外する仕組みを導入している。具体的には、各参加者がアップロードしたローカルモデルの損失値および学習データセットのサイズに基づいてスコアリングを行い、そのスコアから参加者間のマンハッタン類似度を計算する。

その上、クラスタリングアルゴリズムを用いて類似性分析を行うことで、振る舞いの異なる異常な参加者を特定して隔離する。CIC-IDS-2017 データセットを用いた実験において、本手法はラベルフリッピング攻撃を受ける環境下でも、従来の防御を持たない侵入検知モデルの精度である 84.3%から 97.1%へと劇的な回復と向上を達成し、ネットワークのセキュリティ保護能力を大幅に強化できることを実証した。

一方、著者ら自身の今後の課題に関する記述から、現在の防御システム設計が抱える実践的な限界も明確に読み取れる。本研究の防御メカニズムは主にデータ層での単純なラベルフリッピング攻撃を対象としており、攻撃者がモデルの勾配や重みを直接かつ巧妙に操作する、より高度に設計されたモデルポイズニング攻撃に対する堅牢性は依然として未検証である。さらに著者らは、悪意のあるクライアントの数が動的に変化する状況や、学習の特定の通信ラウンドを狙い澄まして攻撃が開始されるような、時間的および戦略的に高度化された攻撃の有効性についての調査が不足していることを率直に認めている。このような適応型かつ遅延型の攻撃に対しては現在の静的なスコアリングおよびクラスタリング手法が破綻する危険性があり、敵対的かつ予測不可能な現実の IoT 環境において実運用に耐えうる包括的な防御メカニズムの構築が極めて重要な課題として残されている。

7. **Dimolianis et al. 2022** [85]：複数の独立したネットワークドメインにまたがるサイバーインフラストラクチャにおいて、プライバシーを保護しつつ分散サービス妨害攻撃を検知し、さらにドメイン間で協調して攻撃トラフィックを早期に緩和する連合学習フレームワークを提案した。

   この研究は、単一のネットワーク境界のみでの防御には限界があるという課題に着目し、複数の自律システムがトラフィックの機密なデータを共有することなく、多層パーセプトロンなどの軽量な機械学習モデルを共同で訓練するアプローチを採用している。特にドメインネームシステムを悪用したフ

ラッド攻撃やリフレクション攻撃などの大規模な脅威に対して、各ドメインの境界ルータで抽出された特徴量を用いてローカルモデルを更新し、中央のオーケストレータで集約を行う。さらに本フレームワークの最大の貢献は、検知モデルの構築にとどまらず、ボーダーゲートウェイプロトコルのフローマネージメント仕様や、オペレーティングシステムのカーネルをバイパスする高速なパケット処理技術である拡張データパスをデータプレーンに活用し、検知された攻撃のシグネチャや緩和ルールをドメイン間で即座に共有および適用する協調緩和メカニズムを設計した点にある。シミュレーション環境における評価により、この連携防御が大規模な攻撃トラフィックを標的ネットワークに到達する前のエッジネットワークの段階で効率的に遮断できることを示した。

一方、異なるセキュリティポリシーを持つ複数の自律システム間でトラフィックの遮断ルールを動的かつ自動的に共有し適用するためには、ドメイン間の極めて強固な信頼モデルの構築が不可欠である。検知モデルが誤警報を出した場合、ブラックホールルーティングの特質上、広範囲の正規トラフィックまで巻き込んで遮断してしまうという深刻な副作用リスクを常に抱えている。エッジルータにおけるトラフィックの高速な分類と破棄を実現するために特定のパケット処理フレームワークやデータプレーン開発キットに強く依存しており、計算資源やオペレーティングシステムに厳しい制約を持つ一般的なIoT ネットワークのエッジデバイスや、異機種が混在する環境に対して、この高度なハードウェア依存の協調緩和機構をそのままスケーラブルに展開することは、アーキテクチャ上の最大の障壁として残されている。

8. **Trusted Multi-Domain（Yin et al. 2022）** [86]：複数のネットワークドメイン間でデータプライバシーを保護しつつ、包括的なDDoS 攻撃の検知を実現するため、連合学習とブロックチェーンを統合した。従来の検知手法は単一のドメインデータに依存しているため、多様な攻撃ベクトルを網羅できないという課題があった。この問題に対処するため、提案手法では攻撃をネットワーク層やアプリケーション層などの五つのカテゴリに分類し、ドメイン間でモデルパラメータのみを共有して学習を行う。さらに、悪意のあるノードによるデータポイズニング攻撃を防ぐため、スマートコントラクトを用いて各参加者の相互作用、データ、およびリソースを総合的に評価する評判メカニズムを導入している。VMware 上の仮想マシン群を用いた環境における評価では、FATE およびイーサリアムを基盤として実装され、大部分の攻撃カテゴリにおいて 95%以上の精度を達成した。また、評判のしきい値を0.6 に設定することで、悪意のある参加者を効果的に特定および排除できることを実証し、各ドメインの分散知識ベースへ安全に抽出された特徴を保存する仕組みも構築した。

一方、ブロックチェーンとの頻繁なやり取りによる膨大な通信リソースの消

費を避けるため、現在の実装では最終的な反復モデルのみを評価の対象としている。しかしこの妥協により、学習の途中段階における悪意のある振る舞いや巧妙に隠蔽された汚染を見逃すリスクが残されている。また、著者らはクライアント間のデータ分布が異なる Non-iid が検知精度に与える悪影響を完全に排除できておらず、データの偏りに対する緩和策の探求が今後の課題であると率直に認めている。

9. **FELIDS（Friha et al.、 2022）** [87]：農業分野への技術導入に伴うサイバーセキュリティの脅威に対処するため、農業 IoT インフラストラクチャを保護する深層学習ベースの連合学習侵入検知システムである FELIDS を提案した。この研究は、データを中央サーバに集約する従来の手法が機密性の高い農業データに対するプライバシー上の懸念を引き起こすという問題に着目している。提案システムでは、ネットワークのエッジノードがローカルのトラフィックデータを用いて DNN、CNN、RNN の三種類の異なる分類器を独立して学習させる。その後、各エッジデバイスはモデルの更新パラメータのみを安全なチャネルを通じて集約サーバに送信し、グローバルモデルを最適化する。Google Colaboratory および単一の計算機を用いた computation 環境での実験において、CSE-CIC-IDS2018、MQTTset、および InSDN という三つの多様なデータセットを用いて評価が行われた。その結果、データが独立同分布であるかどうかにかかわらず、提案手法はプライバシーを保護しつつ、集中型学習モデルに匹敵するかそれを上回る検知精度を達成した。特に、時間およびエネルギー消費の観点からは DNN が最も効率的であり、実用的なエッジ展開に適していることが実証されている。

   一方、現在のシステム設計は、すべての参加エッジノードが初期状態から完全に信頼できるという前提に立っているが、実際の広大な農業ネットワークにおいては一つまたは複数の屋外ノードが悪意を持った攻撃者に物理的または論理的に乗っ取られるリスクが常に存在する。悪意のあるノードがグローバルモデルの検知を逃れることを目的として、意図的に汚染されたデータを学習プロセスに混入させるデータポイズニング攻撃を実行した場合、現在のメカニズムではそれを特定し防御することができない問題も存在している。そのため、このような悪意のある参加者の影響を自律的に排除し、システムの堅牢性を高めるためのゼロトラストベースのセキュリティ対策の統合が、実稼働に向けた極めて重要な未解決課題として指摘されている。

10. **Shahid et al. 2021** [88]：IoT デバイスを標的とするネットワーク攻撃を検知するために、連合学習を適用できるかを検証した研究で、エッジデバイスの厳しい計算資源の制約を考慮し、ロジスティック回帰（LR）およびマルチラベル分類（MLC）モデルという比較的軽量な機械学習アルゴリズムを検知器として採用している。モデルの評価は、ネットワーク侵入検知の標準的なベンチマークである NSL-KDD データセットを用いて実装された。これ

により、生のトラフィックデータや機密情報を中央サーバに送信することなく、分散したクライアント間でモデルの重みパラメータのみを共有し、プライバシーを保護しながらネットワークの脅威を識別する分散型アプローチの基本的な有効性が示されている。

一方、本研究の評価が純粋なソフトウェア上の計算シミュレーションに完全に依存しており、実際の分散サービス妨害攻撃が発生した際の通信帯域幅の枯渇、パケットの損失、あるいはエッジデバイス間の非同期的な通信遅延といった物理的なネットワークの制約が一切考慮されていない点である。

4.4 主要な研究課題

本節では、表 4.4 および第 4.1 節・第 4.2 節の整理を踏まえ、モデル系統（G1–G7）を横断して共通に扱われる主要課題を整理する。本節の焦点は、検知器（CNN/MLP/RNN/GNN/Transformer 等）のアーキテクチャそのものの差異ではなく、連合学習 を実運用へ適用する上で不可避な制約（データの分布偏在、端末の異機種性、通信制約、遅延制約、攻撃耐性）に対して、各研究がどのように対処しているかにある。

4.4.1 Non-iid と異機種性

多くの研究は、クライアント間でデータ分布が一致しない状況（Non-iid：ラベル分布の偏り、特徴量の偏り、環境差）を前提としている。Non-iid の影響は、大域モデルの局所最適への収束、更新の不安定化（振動）、および未知の攻撃に対する汎化性能の低下として顕在化しやすく、これらへの対策が研究の中心課題となっている。代表的なアプローチは、以下の三つに大別できる。

- 集約における更新の不均衡を抑制する手法：クライアントの信頼度や性能に基づく重み付け平均、外れ値による更新の抑制（クリッピング等）、頑健な集約規則（Robust Aggregation）の導入。
- モデル共有範囲の設計：局所特徴を保持するための層別共有や、潜在表現のみを共有する部分共有など、分布差の大きい要素の影響を局所側へ留める設計。
- 学習プロセスの安定化：正則化項や近接項により大域モデルからの過度な乖離を抑制する手法、あるいは学習率や同期周期の動的調整による収束の安定化。

また、IoT 環境では端末性能の差（CPU、メモリ、電力、通信品質）により、学習・推論の実行能力がクライアントごとに異なる「異機種性」が課題となる。異機種性を前提条件として言及する研究は多い一方で、低性能デバイスを用いた実測評価や、端末性能差がモデルの更新品質・遅延に与える影響を体系的に評価した事例は限定的である（表 4.2）。

4.4.2 通信効率と更新の実行可能性

連合学習 の運用において通信帯域は主要なボトルネックであり、通信量の削減はほぼすべてのモデル系統に共通する設計目標である。具体的には、勾配の量子化・圧縮・疎化、更新頻度の低減、共有対象の限定（潜在層共有など）、パラメータ効率化（LoRA 等）といった方策が提案されている。これらは平常時の通信量削減には有効である一方、DDoS 攻撃下で顕在化する通信混雑・遅延増大・パケット損失といった悪条件下を、評価設計に明示的に組み込んでいる研究は多くない。そのため、攻撃下においても更新プロセスが完遂できるか、すなわち更新の遅延・欠落が生じても学習が破綻しないかについて、横断的に比較することは困難である。

さらに、通信削減策は更新に含まれる情報量を減少させ得るため、検知性能や収束安定性との間にトレードオフが生じる。多くの研究は精度指標（Accuracy/F1 等）を中心に議論するが、通信削減に伴う性能劣化が、どの条件下（Non-iid の強度、端末差、回線品質）で顕著になるかを体系的に示した例は少ない。この点は第 5 章で論じる「帯域幅のパラドックス」および運用条件の妥当性と密接に関係する。

4.4.3 遅延とリアルタイム性

近年、推論遅延（サンプル当たりの推論時間）や CPU 負荷などの計算コスト指標を提示する研究が増加傾向にある。ただし、遅延の定義は研究間で一様ではなく、推論演算のみ、 前処理と推論の合計、 検知から遮断・制御までを含む E2E 応答、のいずれを指すかが不明確な場合がある。実運用において重要なのは E2E 応答時間であるが、対象文献の範囲では E2E 応答を実測報告した研究はほとんど確認できず、多くは必要性の言及や設計上の議論に留まっている。そのため、検知器単体の高速性が、運用上の即応性（遮断・緩和の完了）へどの程度寄与するかは、現状の文献からは検証可能な形で整理しにくい。

また、時間に関する指標が報告されていても、その粒度と定義が統一されていないため、横断比較の妥当性が損なわれている。例えば、同じ「時間」を扱う指標であっても、以下の定義が混在している：

1. 収束までに要する総時間
2. 1 ラウンド当たりの学習時間

3. 1 エポック当たりの計算時間

4. 時間窓（ウィンドウ）当たりの推論時間

5. サンプル（レコード）当たりの推論時間

これらは計測対象（学習過程か推論過程か）や分母（ラウンド／エポック／窓幅／サンプル）が異なるため、数値が記載されていても同一条件下での速度比較には直結しない場合が多い。

さらに、連合学習 のリアルタイム性は推論遅延だけでなく、更新遅延（ラウンド所要時間、同期待ち、欠落更新の影響）にも依存する。しかし、推論側の速度指標と学習側の運用指標が同一の実験条件で統合的に評価される例は限定的であり、運用要件（即応性・可用性）との整合性を議論するには、評価設計の統一（指標定義、計測範囲、分母の明確化）が不可欠である。この点は第 5 章において、評価設計の不統一や現実的なネットワーク条件の欠落と合わせて批判的に検討する。

4.5 まとめ

本章では、連合学習 を適用した IoT ネットワークにおける DoS/DDoS 検知研究を対象に、年次推移（第 4.1 節）とモデル系統（第 4.2 節）を軸として研究動向を整理した。年次分布からは 2024–2025 年に研究が集中しており、当該分野が近年急速に拡大していることが確認された。また、表 4.4 により、対象（端末／エッジ／ネットワーク／制御層）、データセット、検知手法、連合学習 の枠組み、目的、評価指標を同一の観点で比較可能な形式に整理した。

モデル系統の観点では、G1（CNN 系）および G2（MLP/DNN 系）、G3（RNN 系）は比較的成熟した検知器として採用例が多く、精度面での報告が蓄積されている。一方で、G4（GNN 系）や G6（Transformer 系）は構造情報の活用や高い表現力を通じて新たな可能性を提示するものの、グラフ更新コストや計算・通信負荷といった運用上の課題も含めた議論が必要である。さらに、G7（システム／ネットワーク機構主導）は、帯域割当や制御連携など運用制約へ踏み込む点で重要であり、リアルタイム性および更新実行可能性の議論と親和性が高い。

4.4 節では、Non-iid と異機種性、通信効率と更新の実行可能性、遅延とリアルタイム性が主要課題として反復的に現れることを整理した。集計結果（表 4.2）からは、モデル精度は主要指標として現れる。推論遅延や通信量の報告は増加傾向にあるが、通信混雑・損失条件や E2E 応答に関する体系的な報告は依然として限定的であることが確認された。

表 4.4: 連合学習 を用いた IoT ネットワーク DoS/DDoS 研究の整理

年	論文	対象	データセット	検知方法	連合学習	目的	評価
G1. CNN / ResNet / 1D-CNN 系							
2026	AFL-SecNet	ネットワーク/IDS	UNSW-NB15 CICDDoS2019	1D-CNN（個性化）	暗号化安全集約（sk1/sk2 分割）+適応学習	秘匿更新 プライバシー保護 個性化で精度維持	Acc: 97.58/96.60% Latency: 1.15s/0.63s
2025	FLAME	ネットワーク/IDS	CICDDoS2019 CICIDS2017 クロス	1D-CNN	JSD 距離フィルタ KDE 再構成集約	ゼロデイ攻撃への汎化 feature skew 緩和 更新逸脱の抑制	Acc(CICDDoS2019、ゼロデイ=NTP)：97.03–99.81% Acc(CICIDS2017、ゼロデイ=DoS slowloris)：70.24–94.72% Acc(train=CICIDS2017 → test=CICDDoS2019)：LDAP 93–94%+ Latency: 未報告
2025	FedDWC	IoT 端末	IoTID20 CICIoT2023	軽量 CNN	クラスタリング 動的重み付け集約	Non-iid 対応 収束安定 精度維持/向上	Acc: 99.21%(IoTID20) 99.10%(CICIoT2023) Latency: 未報告
2025	Albanbay（Data Scaling Study）	IoT 端末	CICIoT2023	CNN（DNN/CNN+BiLSTM 比較）	FedAvg	規模 (10 → 150) 影響 実機で速度/性能測定 精度-速度の統計基盤	Acc: 95.27%(10) 93.98%(150) Latency: 104–109 ms/sample
2024	FL-DAD	IoT ネットワーク	CICIDS2017	CNN	FedAvg	分散検知 集中点リスク低減 可拡張性（ノード増）	Acc(DDoS): 98.7% 通信量: 7.9MB Latency: 80ms

続く...

年	論文	対象	データセット	検知方法	連合学習	目的	評価
2024	FEDSA-ResnetV2	IoV/エッジ	TON-IoT UNSW-NB15 CIC-IDS2017	SA-ResnetV2 （ResNetV2 改）	FedAvg 最良モデル選択 早期停止 8-bit 量子化	精度向上 計算削減 エッジ展開	Acc: 95.76% 性能向上: 最大+34% （ベースライン比） 効率: 最大 72.6 倍 （INFT/Ms/mFLOPs/Params） Latency: INFT で評価 （絶対値未記載）
2024	NIDS-FGPA	IoT/IIoT	Edge-IIoTset CICIoT2023	2D-CNN+ BiGRU （混合）	GSA 集約 Paillier 安全集約	安全集約 通信輪次削減 精度向上	Acc: 94.5% / 99.2% Latency: 未報告 Comm: 22%〜51%削減 通信量が重い (256-bit 鍵で 363.2MB)
G2. ANN / MLP / DNN 系							
2025	ARAFL-BAD	IoT エッジ	N-BaloT	MLP	非同期 FL 信頼度重み付け 更新クリップ	ゼロデイ検知 耐性強化 (Poisoning) 異機種/遅延対応	F1: 99.85% ASR(対抗攻撃): <7% Inf Time: ∼0.3ms
2024	FLAD	IoT ネットワーク	CICDDOS2019	MLP	適応的 FL (動的参加選択 +計算量割当)	Non-iid 対応 収束高速化 負荷分散	F1: 0.9899 (50 clients) 収束時間: ベース比 ∼1/5 Latency: 未報告
2024	DTFL-CDS	医療 IoT エッジ (デジタルツイン)	CICDDOS2019	ANN	FLDS + ATES (閾値ベース 早期停止)	低遅延化 ゼロデイ検知 計算資源効率化	F1: 0.81–0.83 (peak:0.98($\alpha = 25$)) Latency: クラウド比 33%減 CPU 負荷低減
2023	WFL-SDN- LR-DDOS	SDN 制御層	CAIDA	ANN	WFL (重み付き FL) 係数手動設定	LR-DDoS 検知 プライバシー保護 制御層負荷軽減	Acc: 98.85% Latency: 0.019 ms/rec (サンプル単位)

続く...

年	論文	対象	データセット	検知方法	連合学習	目的	評価
2022	FedDDoS (CIDD)	IoT エッジ	CICDDOS2019	MLP	クラウド集約型 (周期的協調)	テナント間連携 プライバシー保護 新規導入即時化	Acc: 84.8% (未知データ) Latency: 未報告
G3. RNN / LSTM / GRU 系							
2025	SAMFL-SCDCOA	IoT エッジ	CICIDS2017 UNSW-NB15	SCDGRU (Self-Attn +GRU)	FedAvg +最適化 (COA/IPOA)	ゼロデイ検知 精度向上 計算効率化	Acc: 99.14%(CIC) 99.04%(UNSW) F1: 97.76%(CIC) 91.69%(UNSW)% CT: ∼5.08s/7.88s (処理時間)
2025	FedXAI	IoT デバイス	N-BaloT	LSTM	FedAvg +XAI (SHAP)	説明可能性 (XAI) プライバシー保護	Acc: 99.28%(Botnet) F1: 99.54%(Botnet)
2024	FL-SCNN-Bi-LSTM	WSN/IoT ノード	WSN-DS CICIDS2017	SCNN+ Bi-LSTM	FedAvg +層別集約 (ローカル層保持)	Non-iid 対応 DoS 検知 異機種対応	Acc: 99.70%(WSN) 99.93%(CIC)% F1: 99.60%(WSN) 99.93%(CIC)
2022	FL-AD (Mothukuri)	IoT/IIoT GW	Modbus Network	GRU+RF (アンサンブル)	FedAvg +多窓モデル 並行学習	異常検知 プライバシー保護 収束高速化	95.65% (GRU) 学習時間: 集中学習より効率的 Latency: 未報告
2021	FLDDoS	分散 NIDS	CICDDOS2019 CICIDS2017	RNN+AE	FedAvg +階層的集約 (K-Means)	Non-iid 対応 通信削減	Acc: 94.82% 通信回数: 40%削減 (対 FedAvg)
2021	FIDS	分散 NIDS	CICDDOS2019	GRU+TDG	FedAvg+ 重み付け	Non-iid 対応 収束安定化 DDoS 検知	Acc: >97% (多数) 収束速度向上 Loss 変動抑制
G4. GNN 系							

続く...

年	論文	対象	データセット	検知方法	連合学習	目的	評価
2025	FedSTGCN	IoT エッジ	NF-BoT-IoT-v2 NF-TON-IoT-v2	STGNN (LSTM+GNN)	FedAvg ベース +FedCog (グラフ分離)	時空間相関学習 プライバシー保護	Acc: 96.9–97.5% Recall: ∼90% Latency: 未報告
2025	FedGATSage	IoT ネットワーク	NF-TON-IoT CIC-TON-IoT	GAT(Local) +GraphSAGE	階層的集約 (サーバ側 グラフ構築)	構造/時間 パターン保持 通信効率化	Bal-Acc: 78–80% 通信量: 85%削減 (ベースライン比較)
2025	GraphFedAI	IoT デバイス	CICIoT2023	GNN+Attention	FedAvg +通信圧縮 (量子化等)	スケーラビリティ プライバシー保護 適応性向上	Acc: 97.9–98.7% Lat: 24–28ms Comm: 9.3MB/round
2025	FeCoGraph	エッジサーバ	NF-BoT/TON CICIDS2018	GCN+対照学習	Ditto (個性化 FL)	Few-shot 対応 Non-iid 対応 ラベル依存低減	Acc: 96.90–99.89%(二値分類) F1: 91.12–99.11%(二値分類)
2024	Federated DGNN	IoT エッジ	未公開	DGNN	FedAvg	異機種対応 複雑攻撃検知 プライバシー保護	Acc: 95.30% F1: 93.30% Latency: 未報告
2025	Fed-GNN	IoT GW	CAIDA (混合データ)	GNN (時系列)	加重集約 +暗号化 (FH-MIFE)	ポイズニング防御 Non-iid 対応 安全性確保	Acc: 97.86% Latency: 未報告
2024	FLARE	医療 IoT エッジ	CICIDS2017 UNSW-NB15	GNN	Reg-FedAvg (正則化)	性能制約対応 異機種対応 プライバシー保護	F1: 0.935(CIC) 0.927(UNSW) Inf Time: ∼1.9s
2024	AGAT-FL	IDS	N-BaloT CIC-TON-IoT	CNN-GRU (+GAT 信頼度)	信頼度加重 (GAT ベース) +異常除外	ポイズニング防御 ロバスト性向上	Acc: 94.01%(N-BaIoT) 91.02%(ToN-IoT) F1: 93.75%(N-BaIoT) 91.15%(ToN-IoT)
2024	Fed-GAT (Wu et al.)	IDS	NSL-KDD KDD Cup 99	GAT	FedAvg	構造情報学習 プライバシー保護	Acc: >99.9% Latency: 未報告

続く...

年	論文	対象	データセット	検知方法	連合学習	目的	評価
2023	FedGNN (CAN)	車載 IDS	CAN データ	GNN+ Openmax	FedAvg /FedProx	リアルタイム性 ゼロデイ検知 Non-iid 対応	Acc: 100% Latency: 3.2ms (予測時間)
G5. Autoencoder / 表現学習系							
2025	FL-IDS (Pranggono)	IoT デバイス	N-BaloT	Deep AE (Unsupervised)	FedAvgM (Momentum) +RandSearch	プライバシー保護 教師あり/なし比較 異常検知	Acc: 90.39% f1: 93.12%
2025	FedGAN-ID	車載 CAN (ECU/GW)	実車データ（非公開）	GANomaly +AC 分類器 +OOD 検知	FedAvg ベース +DP 合成データ共有	Non-iid 対応 未知攻撃検知 プライバシー保護	Acc: >94% (実車 5 分類) Lat: 1.6–2.0ms Comm: ∼16MB/rd
2025	FL-DAE (Shirvani)	IoT デバイス	N-BaloT	DAE +閾値判定	FedAvgM +部分選択 +再学習	Non-iid 対応 不安定な参加対応 DDoS/Bot 検知	AUC: 93.95% FPR: 2.19% TrainTime: 49s
2025	FedMSE	IoT GW	N-BaloT	SAE-CEN (Shrink AE +Centroid)	MSEAvg (性能ベース 重み付け)	Non-iid 対応 通信効率化 半教師あり学習	AUC: 97.30%(高 Non-iid) 参加率 50%で 性能維持
2024	FL-LSTM-AE (Munaweera)	5G Core/CPS	自作データ (5G Core)	LSTM-AE	FedAvg +堅牢集約比較 (Krum/Bulyan)	5G-CPS 保護 ポイズニング防御 異常検知	Acc: 89.7% (FL) F1: 92.2% Latency: 未報告
2024	CAE-Agri (Khacha)	農業 IoT (Fog/Edge)	CICIoT2023	CAE (Conv AE) +SMOTE	FedAvg	データ不均衡対応 異常検知 プライバシー保護	Recall: 99.99% (IID) ∼86%(Non-iid) F1: 99.40%
2024	AE-IoT (Majeed)	IoT エッジ	N-BaloT	AE (閾値判定)	FedAdam +特徴選択 (MI)	ボットネット検知 プライバシー保護	Acc: 99.83% FPR: 0.69% Latency: 未報告

続く...

年	論文	対象	データセット	検知方法	連合学習	目的	評価
G6. Transformer / ViT 系							
2025	FLoV2T	AIoT エッジ	CICIDS2017 CICIDS2018 (Raw PCAP)	ViT + LoRA (RBV 画像化)	FedAvg +RGPA (正則化集約)	細粒度分類 通信削減 Non-iid 対応	Acc: 97.26% F1: 96.99% Comm: LoRA で 1/64 削減
2025	ResVGG- SwinNet	IoT ネットワーク	CIC-DDoS2019 UNSW-NB15 IoT23	ResVGG- SwinNet	FedAvg	多ラベル検知 拡張性 異機種対応	Acc: 99.0% F1: 98.5% Inf Time: 2.8ms
2025	SecureFogDL	医療 IoT (Fog)	TON-IoT	Transformer +Autoencoder	FedAvg (周期通信)	フォグ層検知 低遅延 プライバシー保護	Acc: 98.35% Latency: 21ms/rec Comm: 8.2MB/rd
2024	securityBERT	5G/IoT エッジ	Edge-IIoTset	securityBERT (軽量 BERT)	FedAvg	エッジ展開 軽量化 Non-iid 対応	Acc: 97.12% Inf Time: 2.6ms(GPU) 450ms(RPi4) model size:30.38Mb
2024	Access-Side PFL (Luo)	エッジサーバー	N-BaIoT	Transformer	PFL (共通/個性化層 +動的調整)	アクセス側検知 Non-iid 対応 個性化	Acc: 99.2%(IID) ∼99.5%(Non-iid) Latency: 未報告
G7. システム/ネットワーク機構主導（帯域/協調緩和/信頼管理など）							
2025	EGDA	MEC	MNIST 等 (FL タスク用) ※攻撃は帯域枯渇	ゲーム理論 +オークション	FL 最適化 (帯域割当)	帯域 DDoS 緩和 FL 完了時間短縮	Latency: 削減確認 (FL ラウンド時間) Social Welfare: 向上
2025	Hybrid EFL	IoT エッジ (5G)	NCSRD-5G-DDoS (公開)	XGB/RF (アンサンブル)	確率共有型 FL (パラメータ不使用)	帯域効率化 プライバシー保護 高精度化	Acc: 97.9% Latency: 190.10ms 通信量削減：62.64MB

続く...

年	論文	対象	データセット	検知方法	連合学習	目的	評価
2025	FedLAD	SDN 制御層	CICDoS2017 CICDDoS2019 InSDN	XGBoost	FedAvg +順位ベース選択	SDN 拡張性 リアルタイム検知 帯域監視連携	Acc: 最高 97.64% CPU 負荷: ∼32% (集中式は 48%)
2025	Al-Ghadi et al. 2025	IoT ネットワーク	IoTID20	DNN/CNN/ RNN+FS	FedAvg	特徴量選択 (FS) 精度向上 実行可能性	Acc: 99.73%(CNN) 97.54%(RNN)
2023	CAFNet	IoT エッジ	CICDDoS2019 Bot-IoT	Autoencoder	FedAvg +潜在層のみ共有	通信効率化 ゼロデイ検知 異常検知	F1: 99% (CICDDoS) 通信量: 95%削減
2023	FLIoT	IoT エッジ	CICIDS2017	DNN(推定)	FedAvg +LSM スコア除外	ポイズニング防御 (Label-flip) ロバスト性	Acc: 97.1% (攻撃下) 84.3%(防御なし) Loss: 正常化
2022	FELIDS	Agri-IoT エッジ	独自生成 (Agri-IoT 環境)	DNN/CNN/ LSTM	FedAvg	Non-iid 対応 実行可能性 プライバシー保護	Acc: 99%+ (IID/Non-iid) Energy: LSTM が高負荷
2022	Dimolianis et al. 2022	多ドメイン/ISP	WIDE (良性) Booters (攻撃)	パケット署名 (XDP 緩和)	FedAvg +署名共有	協調緩和 (XDP) クロスドメイン インフラ防御	TPR/TNR: 良好 Throughput: 線形向上
2022	Trusted-MD	多ドメイン GW	独自生成 (模擬 DDoS22 種)	HomoCNN	安全集約 +ブロックチェーン	ポイズニング防御 信頼性管理 知識共有	Precision: >95% Recall: 高水準 (DRDoS のみ低)
2021	Shahid et al. 2021	IoT デバイス	NSL-KDD	LR/MLC	FedAvg	IoT 適用性 プライバシー保護 モデル比較	Acc: 約 98%

第5章　現実運用に近づけるための論点整理

5.1　運用ギャップ

第 4 章では、対象文献を年次推移と検知器に基づいて分類し、研究動向を俯瞰した。本章ではその整理結果を踏まえ、連合学習を用いた DoS/DDoS 検知研究を実運用へ繋ぐ際に生じる主要な論点を、ギャップとして整理する。

集計結果からは、検知性能（Acc/F1 等）の報告は充実であるが、現実運用に左右する条件や指標の扱いが十分に揃っていないことが分かる。具体的には、推論遅延（13/50）や通信量（12/50）の報告は一部に留まり、攻撃時の通信混雑・損失条件を明示した評価は少数（3/50）である。また、検知から遮断・緩和動作の完了までを含む E2E 応答については、実測値として報告した研究は確認できない（0/50）。このため、通信制約されるときに更新が成立するか、どのぐらいの時間で検出かつ緩和ができるかといった論点を、文献の間では記述内容が異なるため、比較することが難しい。

さらに、時間に関する指標は提示されていても定義が統一されていない場合が多く、学習と推論の計測値は単位が研究ごとに異なる。加えて、実験環境は計算のみ computation が多数を占め、低性能デバイスによる実測評価も限定的である。これらは評価の比較可能性と、運用要件との接続を難しくする要因となる。

5.2　通信悪化時の更新テスト

本節では、DoS/DDoS により帯域制限、通信遅延と損失による悪化する状況で、連合学習のモデル更新がどの程度成立し得るかを論じる。ここで重要なのは、通信量を減らす通信削減の工夫と、攻撃されるときでも更新が届き続けることは同一ではない点である。通信削減は更新の成立性の必要条件になり得るが、帯域が枯渇した状況ではそれだけで十分条件にはならない。

5.2.1　通信削減と更新の成立性の切り分け

連合学習 における通信は、更新するためのデータ量と更新頻度に大別できる。量を減らす手法としては、量子化や圧縮、共有範囲の限定（特定層のみ共有、潜在

表現のみ共有)、パラメータ効率化（LoRA 等）などがある。回数を減らす手法としては、参加設備の選択・非同期化、階層的集約、更新間隔の調整などがある。これらは通常時の通信負荷を抑える上で有効である一方、DoS/DDoS により、ネットワークの帯域幅が飽和し、遅延や損失が顕著になる状況では、「そもそも更新が届くか」「更新が欠落しても学習が破綻しないか」という現実的な問題を直面する。

更新の成立性を議論するには、少なくとも次の 2 点を明示する必要がある：

1. 攻撃されるときに、通信悪化する時の性能や環境指標。特に、利用可能帯域幅、遅延、パケット損失、再送有無、混雑制御の挙動などを明示すること。

2. 更新の欠落・遅延が学習へ与える影響

これらが提示されない場合、通信削減を報告していても「攻撃下で更新が成立する」ことの裏付けには直結しにくい。

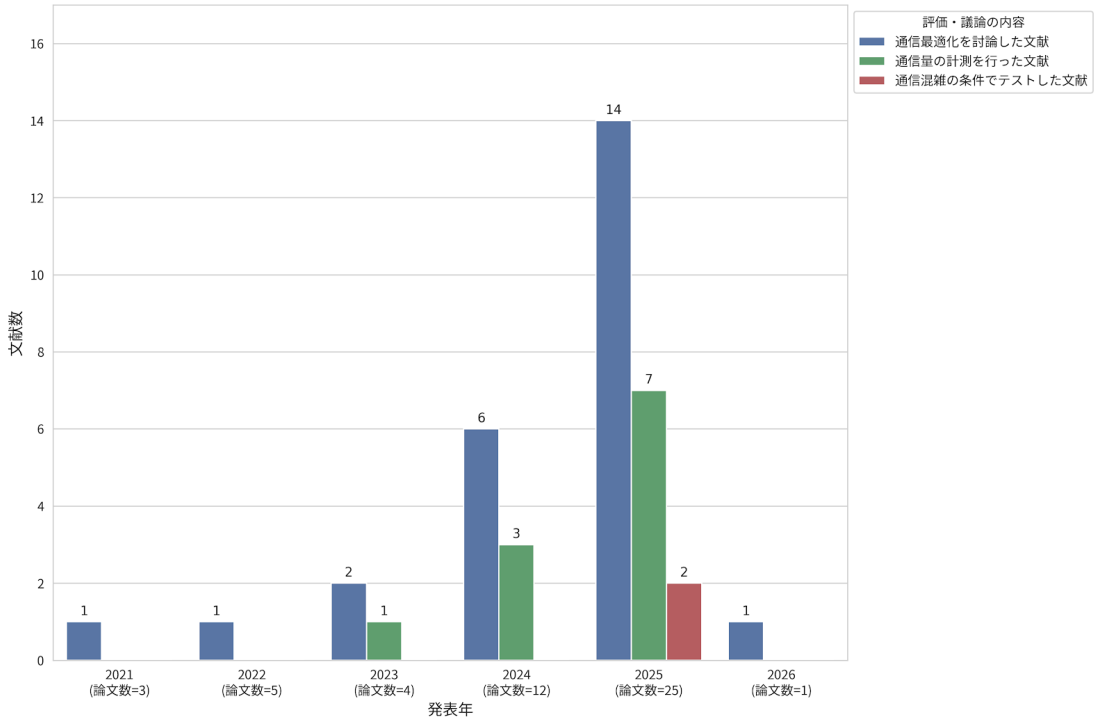

図 5.1: 年次別に見た通信に関連する観点を論じた論文数

5.2.2 対象文献における報告状況

本調査対象文献では、図 5.1 に示す通り、通信最適化に関して討論した研究は 25 本であるが、通信オーバーヘッドを具体的な数値として報告する研究は 11 本に留まり、攻撃下の通信混雑及び損失の条件（帯域枯渇、遅延増大、パケット損失等）

を明示的に組み込んだ評価は 2 本で、少数である。このため、通信削減の効果は示されても、DoS/DDoS 下における更新の成立性までを横断比較できる形で示す研究は限定的である。

図 5.1 に示した通り、2024 年以降は通信最適化を掲げる研究が増加し、2025 年には通信量の数値報告も増える。しかし、混雑・損失条件を組み込んだ評価は 2 件程度に留まり、まだ重視度が低いと考えられる。その結果、多数の研究は攻撃されている時に正常に作動できる保証はないと考える。

表 5.1: 分類別に見た通信関連観点の分布（数（割合%））

分類	文献数	通信最適化の討論	通信量の数値報告	通信混雑・損失条件の評価
G1	7	6（85.7）	3（42.9）	0（0.0）
G2	5	3（60.0）	1（20.0）	0（0.0）
G3	6	4（66.7）	0（0.0）	0（0.0）
G4	10	5（50.0）	2（20.0）	0（0.0）
G5	7	2（28.6）	1（14.3）	0（0.0）
G6	5	3（60.0）	3（60.0）	1（20.0）
G7	10	2（20.0）	1（10.0）	1（10.0）

表 5.1 の示す通り、通信最適化に関して討論した論文の類では、数が最も多いのは G1(CNN 系) であり、7 本の中 6 本で確認される。次いで G3(RNN ／ LSTM 系) が 6 本の中 4 本、G2(DNN/MLP 系) と G6(Transformer 系) はいずれも 5 本の中 3 本である。また、通信量を具体的な数値として提示する研究は G1 と G6 が一番数的に多い。しかしながら、通信混雑や損失といった通信悪化条件を明示的に含めた評価は極めて少なく、G6 と G7(システム／ネットワーク機構主導系) に各 1 本に留まる。すなわち、多くの研究が通信を重要な論点として扱っている一方で、更新の送受信のデータ量を定量化し、さらに不安定な通信条件下で成立性まで示す報告は十分ではない。

5.2.3 通信最適化の典型パターンと限界

対象文献で見られる通信最適化は、主に以下の 3 種類に分けられる：

1. **更新データ量の削減**：量子化、部分更新、共有範囲の限定、LoRA など。
2. **更新頻度や参加を制御**：参加選択、非同期化、階層的集約など。
3. **ネットワーク側の構造と結び付ける**：帯域割当、制御層の連携など。

ただし、これらの手法は、「通常時の通信負荷」を抑える効果は示しやすいが、DoS/DDoS 下の帯域幅が枯渇することを前提とした場合には、更新が欠落する状況をどの程度許容できるかが別問題となる。更新欠落が増えると、サーバ側の集約が遅延し、結果として検知性能の変動や適応の遅れが生じ得る。したがって、通信最適化を「更新の成立性」の議論として成立させるには、通信条件と欠落更新の扱いを評価設計へ組み込む必要がある。

本節の結論として、通信削減が報告される研究は増えつつあるが、攻撃されている時の通信混雑/損失条件を明示し、更新の欠落、遅延まで含めて更新の成立性を検証する研究はまだ限られている。

5.3 リアルタイム性と応答遅延

連合学習を用いた DoS/DDoS 防御では、検知精度だけでなく、検知から対処までの時間制約が運用上の前提となる。第 2 章・第 3 章で整理したように、連合学習は局所学習と更新交換を反復するため、遅延は推論処理だけでなく、通信、集約、同期待ち、更新の再配布により増幅され得る。表 5.2 に示す通り、推論遅延の報告は 50 本中 13 本に留まり、エンドツーエンドの応答時間はほぼ確認されない。この偏りは、通信最適化の提案が増える一方で、リアルタイム性の観点から更新の成立性まで一貫して検証する枠組みが十分に共有されていないことを示している。実運用では、観測と特徴抽出、推論と判定、遮断と緩和（SDN ルール適用、ACL 更新、レート制御など）までを含む E2E 応答が本質的である。一方で 連合学習を用いる場合、これに加えて学習、集約、再配布という更新時間による遅延も運用上の即応性を左右する。推論の高速性と運用の即応性を同一視すると、評価と要件の間にギャップが生じやすい。

本調査で時間関連指標の報告状況を集計すると、推論遅延や学習時間の記載は一定数見られる一方、E2E 応答を実測して報告した研究は確認できない。E2E については必要性の言及や設計上の議論に留まり、検知器が速いことが遮断まで速いことに直結するかは、現状では文献横断で検証できる形になっていない。

表 5.2: 時間関連指標の報告状況

指標	報告数	割合（%）
推論遅延（推論時間）	13	26.0
学習時間（収束/訓練時間）	15	30.0
E2E 応答（検知 $\rightarrow$ 遮断/緩和）	0	0.0
いずれかの時間指標を報告	21	42.0
推論遅延と学習時間の両方を報告	7	14.0

加えて、表5.3の示す通り、たとえ時間指標が報告されていても、定義と計測範囲が統一されていないため、比較が難しい点も存在する。推論遅延ではmsとだけ示され、または1レコード当たり、時間窓当たり、前処理を含むなど、様々な形式で表現されていて、統一されていないという問題が存在する。学習時間でも、収束までの総時間、1ラウンド当たりの時間、1エポック当たりの時間、あるいは時間窓に紐づく処理時間などが混在する。同じ「速さ」を扱っていても、測っている対象や単位が異なれば、数値はそのまま比較できない。

表 5.3: 時間指標の定義が分岐する代表パターン

観点	分母（例）	意味
推論時間	レコード サンプル当たり	遮断の即応性に近いが、前処理（特徴抽出）や制御適用時間を含まない場合が多い。
	時間窓当たり	データ序列、窓幅と不可分であり、窓幅が長いほど「見かけの遅延」は小さくなり得る。
	バッチ当たり 平均推論時間	実装条件（バッチサイズ）に依存し、オンライン運用の遅延推定に直結しない場合がある。
学習時間	ラウンド当たり 総ラウンド時間	通信・同期待ち・欠落更新の影響を受けやすく、攻撃下や混雑下で悪化し得る。
	収束までの総時間 エポック当たり	収束判定や実験条件に依存し、比較には計測範囲と停止条件の明示が必要。

以上より、リアルタイム性の議論を現実運用へ繋ぐには、少なくとも次を同一条件で併記することが望ましい。また、現状ではE2Eの実測が欠落しているため、研究間で「即応性」を比較・議論できる土台が十分ではない。

5.3.1 実験環境と端末性能制約

本節では、対象文献における評価テストについて現実運用への討論を行う。ここでは、第3章で整理した環境カテゴリの定義に従う。unknownに属する文献がないため、結果として、simulation、simulation+real、computation、computation+realの4区分で扱う。

評価環境カテゴリ

図 5.2 に環境カテゴリ別の件数を示す。全体では computation が最多（35/50）であり、次いで simulation（7/50）、computation+real（6/50）、simulation+real（2/50）の順であった。

ここで computation は、連合学習の手順を仮定しつつも、モデル更新の送受信が実験として実行された証拠が十分でない評価を指す。この比率が高いことは、DoS/DDoS の本質である通信悪化の問題を評価設計に組み込みにくい現状と整合的である。とりわけ computation では、通信遅延や欠落更新、同期待ちといった影響が原理的に観測されないため、得られた精度や応答性が攻撃下の運用条件へどこまで外挿可能かを文献横断で議論しにくい。

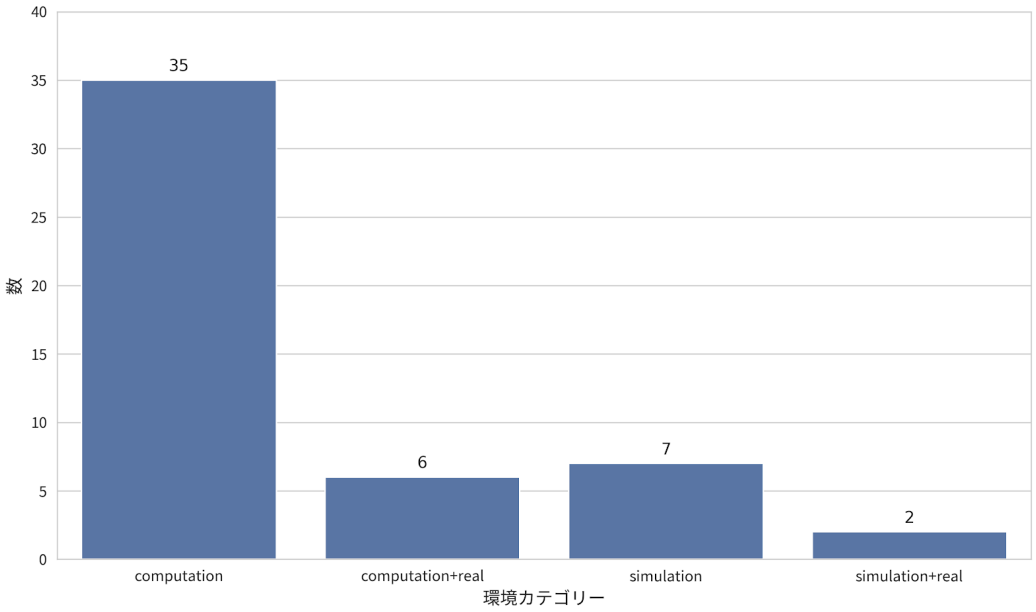

図 5.2: 環境カテゴリ別の件数 (論文数＝ 50)

さらに、computation では通信テストがされず、クライアント間の非同期性や通信状況が評価に反映されない。そのため、集約の安定性や更新の成立性に関する議論が理想条件に偏りやすい。その結果、通信制約下での性能悪化や破綻を比較することが難しく、防御手法の有効性も「通信が正常に成立する」前提に依存して見積もられる恐れがある。DoS/DDoS 攻撃が入ると、通信悪化が学習過程と検知応答に及ぼす相互作用を評価から落とすと、攻撃されている時の検知能力や安定性を把握できないこととなる。

一方、simulation+real は通信動作の実行を前提にしやすく、更新の成立性や通信量、通信混雑の条件など第 5.3 節の論点へ接続しやすい。ただし、real を含む場合でも、実機で検証した範囲が推論のみか、前処理や学習更新、通信まで含むかは研究により異なる。また、通信条件の設定も統一されていない場合があり、結

果の比較可能性は限定される。したがって、環境カテゴリの提示に加えて、計測範囲とネットワーク条件を併記する報告様式が望まれる。

以上より、現状の文献分布はモデルの検知性能には一定の知見が蓄積しているが、通信制約下での成立性や攻撃下の継続運用に関する根拠が不足しやすいことを示している。今後は通信を含む評価へ段階的に拡張し、実運用で支配的となる要因を切り分けて示す必要がある。

低性能端末の評価

IoT を想定する研究では性能制約への言及は多いが、実機上での検証がどの程度行われているかは文献間でばらつきが大きい。そこで本研究では、実機評価を次で整理する：

- 低性能デバイスでの実測（例：Raspberry Pi、Jetson 等）
- 実機評価あるが、低性能デバイスではない
- 実機評価なし

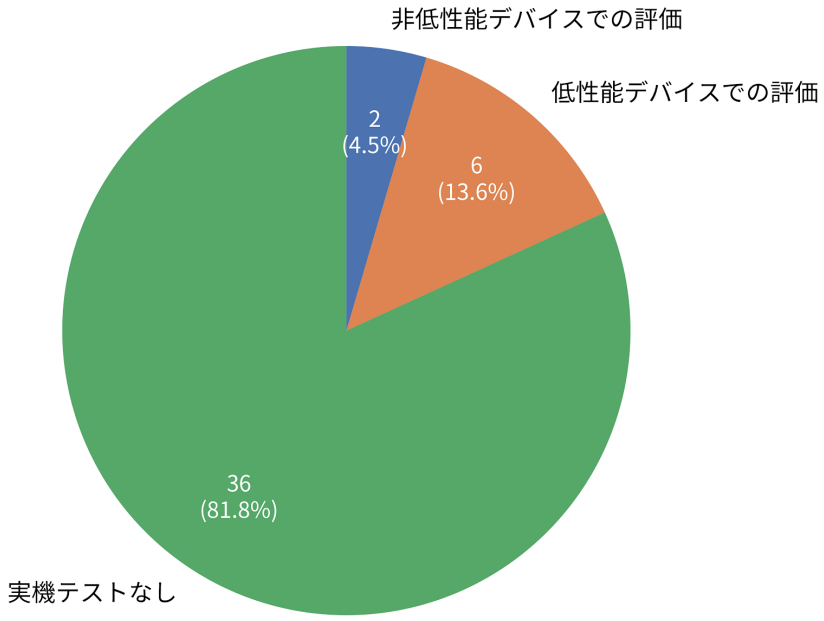

図 5.3: 実機評価のカバレッジ

対象論文のうち IoT 環境を明言した論文は全部で 44 本、そのうち、低性能デバイスでの実測は 6 件に留まり、実機評価はあるものの低性能デバイスではない研究は 2 件であった。したがって、実機評価が確認できない研究約 81.8%のと多数を占める。残りの 6 本は、IoT 環境の適用を明言してはいないが、対象設備実機テス

トは実施されていない結果となった。この分布は、エッジで動作可能やリアルタイムといった主張を運用要件へつなげる際に、可用性が疑われる。IoT ネットワークの適用範囲が拡大している現状では、例え低性能デバイスでなくても、対象設備への実機テストは必須だと考える。

また、実機評価がある場合でも、計測範囲や通信条件は研究ごとに異なる。DoS/DDoS 下の運用要件に照らすと、推論時間だけでなく、学習更新の生成と送受信を含めた成立性の説明が重要である。少なくとも計測範囲、単位、通信条件を併記し、比較可能な形で報告することが望まれる。

5.4 公開データセットの年次分布

本節では、既存研究が用いた公開データセットを対象に、論文年次の分布を整理する。図 5.4 は、横軸を論文年次、凡例をデータセット年次として、各論文が使用した公開データセットの出現回数を集計した結果である．なお、年次ラベルは可視化の都合上、**2026 年分を 2025 年に統合し「2025–2026」**として扱う．

5.4.1 結果

図 5.4 の示した通り、全体として、新しい論文年次であっても古いデータセットへの依存が強い傾向が確認できる。特に 2025–2026 では、2015–2019 年のデータセットが依然として中心であり、出現回数ベースで約 68%が 2019 年以前のデータセットに属する。内訳を見ると、2018 年（例：N-BaIoT、BoT-IoT、CSE-CIC-IDS2018）や 2017 年（CIC-IDS2017）、2019 年（CIC-DDoS2019）といった 2017–2019 年帯の定番ベンチマークが支配的である．また、2015 年（UNSW-NB15）も複数回用いられており、10 年前後のベンチマークが現役である点が示唆される．

一方で、2023 年以降の比較的新しいデータセット（例：CICIoT2023、CIC-ToN-IoT など）も一部に採用が見られるが、2025–2026 における出現回数は限られる。「新しいデータセットへ置き換わった」というよりは、従来ベンチマークの補助として追加されている段階に留まる。

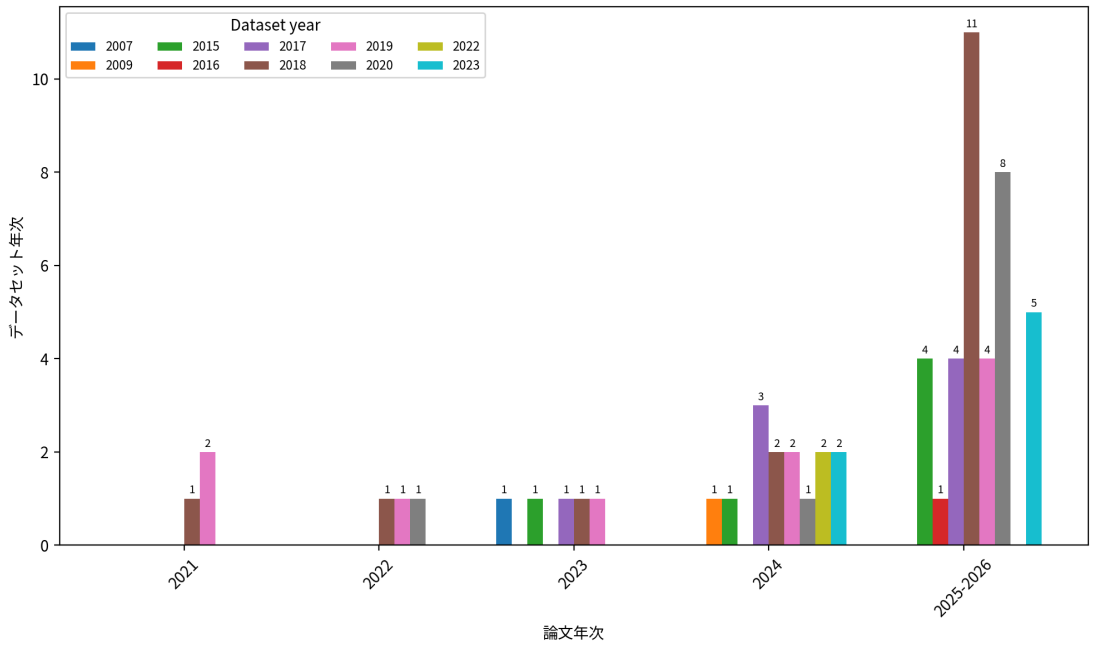

図 5.4: 論文年次ごとの公開データセット年次分布

5.4.2 古いデータセット依存問題

DDoS/DoS は攻撃手法、ボットネットの構造、プロトコル運用が継続的に変化するため、古いデータセットで高精度を示しても、近年の実トラフィックに対する汎化を保証しない可能性がある．加えて、古いベンチマークは、以下のリスクを持つ：

- データの環境やラベリング方針が現在の IoT/エッジ運用とずれがある。
- 攻撃多様性や暗号化、通信混雑条件などの反映が不十分。
- 同一データセット内評価に偏りやすく過大評価を招く。

したがって、近年データセットの採用拡大に加え、クロスデータセットテストや概念ドリフトを意識した検証設計が今後の実験妥当性の向上に重要である。

5.5 今後の研究方向性

第 5 章では、連合学習を用いた IoT 環境下の DoS/DDoS 防御の研究について、運用要件へ繋げる際に生じやすいギャップを、通信関連の評価観点、実験環境や条件の観点から整理した。本節では、これらの整理結果を踏まえ、今後の研究が満たすべき要件を「アルゴリズムの新規性」や「精度」の観点ではなく、「現実運用への条件配慮」や「比較可能な評価設計」という観点を整理する。

本調査では、精度向上を目的とした学習アルゴリズムの議論が先行しやすい一方で、DoS/DDoS 攻撃では通信悪化と性能制約が同時に顕在化し、学習そのものが成立しない、あるいは検知・緩和が遅延して実運用に耐えない状況が生じ得る。したがって重要なのは、以下の 3 点を、評価指標と条件設定を伴って示すことである：

1. 通信が悪化しても更新が届き、集約が進むか
2. 低性能端末を含む実機条件で実行できるか
3. 未知攻撃や環境変動に対してどの程度適応できるか

以上を踏まえ、本章では今後の研究方向を整理する。

5.5.1 通信効率とロバスト性の両立

連合学習は複数クライアントが局所学習を行い、更新を集約することでモデルを改善する枠組みである。しかし DoS/DDoS 下では、帯域枯渇や遅延増大、損失により、更新が欠落しやすく、同期待ちやタイムアウトが発生して学習が停滞し得る。このとき、通信量を削減する工夫があっても、「更新が届く」「集約が進む」ことは自動的には保証されない。したがって、通信効率を論じる際には、少なくとも更新の成立性を独立の論点として扱い、通信が劣化した条件で学習がどの程度継続できるかを明示する必要がある。

本節では、まず更新の成立性を評価設計として扱うための基本要素を整理し、次に成立性を損なわずに通信効率を高める設計の方向性を述べる。

更新の成立性を明確化するための評価設計　更新の成立性を議論するには、「どの条件で更新が成立しなくなるか」を示す必要がある。そのため評価条件として、少なくとも以下の三要素を明示することが望ましい。

- **ネットワーク条件**：遅延・損失・帯域制限を固定値で与えるだけでなく、攻撃時のみ悪化する条件や、時間変動する条件として与える。
- **運用方式**：同期型か非同期型かを明示し、シミュレーション、タイムアウト、欠落更新の扱い（破棄・再送・次ラウンドへ繰り越し）を定義する。
- **動的参加による影響**：部分参加、脱落と再参加を含め、参加率の低下が学習と検知に与える影響を評価する。

これらを固定した単一点の評価では、破綻点が見えにくい。したがって、条件を段階的に厳しくしたスイープを行い、どの条件で学習が停滞し、どの条件で検知が劣化するかを示すことが、実運用への可能性を高める。

実現性を損なわない通信効率化の方向性　通信効率化は、単に送る量を減らすだけでなく、通信悪化下でも更新が届き、集約が進むように設計されるべきである。今回の調査資料内も以下の内容が代表的なものである：

- **通信量改善**：量子化・スパース化・差分更新などにより通信量を削減する。ただし、更新の情報量低下は欠落更新の影響を増幅し得るため、精度だけでなく更新欠落率や収束時間と併せて評価する必要がある。
- **参加制御**：帯域状況や重要度に基づき参加者を選択し、混雑時に学習を破綻させない運用へ切り替える。ただし、参加が偏ると分布の偏りが増幅し得るため、参加制御の基準と公平性、偏りの影響を併記することが望ましい。
- **非同期化とタイムアウト設計**：同期待ちを避けることで停滞を緩和する、ただし、古い更新が混入することで不安定化する場合がある。よって、更新の安定性の根拠を示す必要がある。
- **階層化と局所集約**：端末–エッジ–センターの多段集約により、上位帯域の負荷を緩和し、局所的な実現性を確保する。ただし、集約点が攻撃対象となる可能性があるため、集約点の過負荷や障害時の振る舞いまで含めて評価することが望ましい。

ロバスト性の範囲拡張　更新の外れ値への耐性だけでなく、通信悪化に起因する欠落更新、遅延、偏った参加がもたらす不安定化を含めて扱うべきである。DoS/DDoSでは、回線品質に依存して参加者が偏ることで、分布の偏りが増幅され、学習が特定条件へ過適合しやすい。したがって、通信条件と参加の偏りを明示し、防御や集約戦略が「通信が成立する前提」に依存していないかを検証する必要がある。

評価指標の改善　通信効率と更新の成立性を同時に議論する最低ラインとして、以下の報告を推奨する：

- **通信量**：更新 1 回または 1 ラウンド当たりの送受信量
- **更新**：更新成功率または欠落率、タイムアウト発生率
- **収束の時間**：所定性能に到達するまでの時間またはラウンド数
- **応答性**：推論遅延、可能であれば検知から緩和までの遅延
- **条件の明示**：遅延・通信損失・帯域制限、参加率、運用方式

これらを併記することで、通信削減の効果と実現性への影響を切り分けられ、どの条件で成立し、どの条件で破綻するかを比較可能な形で示しやすくなる。

5.5.2　ハードウェア実証テスト

連合学習ではリアルタイム、エッジで動作可能といった主張が見られるが、実機での計測が伴わない場合、その主張がどの条件まで成立するかを判断しにくい。とりわけ IoT では端末性能が制約となりやすく、推論のみを測定していても、更新生成や送受信を含めると成立しない場合があり得る。したがって、実機条件で、どの処理までを計測対象としたかを明示し、運用要件と接続できる形で報告することが重要である。

最小構成のテストベッド設計　比較可能な検証のためには、役割分担を明示するように、端末・エッジ・集約サーバの 3 層構成を基本とし、次の要素を最小セットとして用意するべきだと考えられる：

- **低性能端末**：演算性能が限定された端末を用意し、推論・更新生成の実行可能性を測定する。
- **エッジ**：端末の集約点として、推論・学習・前処理の分担を切り替えられる構成とする。
- **集約サーバ**：集約とモデル配布を担い、同期・非同期、部分参加を切り替えられるようにする。
- **ネットワーク条件付与**：遅延・損失・帯域制限を与え、攻撃時に条件が悪化するシナリオも再現する。
- **負荷生成**：データ面の混雑だけでなく、端末性能や制御面の過負荷を想定した負荷も与える。

これにより、端末主導、エッジ主導、ネットワーク／制御層主導といった展開シナリオを切り替え、脅威モデルに対応した評価が実施しやすくなる。

計測項目：推論だけでなく更新と通信を含める　実機検証で重要なのは、推論速度だけでなく、更新生成と送受信を含む成立性を示すことである。最小報告項目として、少なくとも以下の計測を推奨する：

- **推論遅延**：前処理を含むか否かを明示する。
- **更新時間**：ローカル学習の計算コストを測定する。
- **更新送受信時間**：通信込みの遅延と欠落の発生を測定する。
- **エンドツーエンド応答**：検知から緩和動作までの遅延を測定する。

- **計算資源指標**：CPU 使用率、メモリ、電力、温度、スロットリングの有無を併記する。

また、計測範囲が混在すると比較が困難になるため、推論のみ、前処理込み、更新生成込み、通信込み、緩和動作込み、のように範囲を明示し、結果を範囲別に提示することが望ましい。

展開シナリオと脅威モデルの整合：評価負荷の対応付け　実機評価では、どの展開シナリオを仮定するかを明示し、脅威モデルが想定する枯渇点と、評価で与えた負荷を対応付ける必要がある。例えば、帯域枯渇を想定するなら遅延・損失・帯域制限を与え、端末資源枯渇を想定するなら学習更新を含めた計算資源の測定を行う。この対応関係を明示することで、「検知できる」だけでなく「攻撃下で運用できる」ことを根拠付きで議論しやすくなる。

そして実機検証の価値を高めるには、再現性と比較可能性が不可欠である。少なくとも、データ分割、学習設定、ネットワーク条件、計測範囲、および実装の構成を遡れるように明示することが望ましい。また、ログの形式や計測手順を統一することで、異なる研究間で結果を比較しやすくなる。

5.5.3　未知の攻撃への適応能力

実運用では、攻撃手口やボットネットの振る舞いが時間とともに変化し、学習時に観測していないゼロデイ攻撃が起こりうる。さらに IoT 環境は設置場所・機器構成が多様であり、Non-iid 課題も存在するため、単一データセット上の精度だけではゼロデイの検知を判断しにくい。DoS/DDoS 下では通信条件そのものが悪化し得る以上、攻撃下でも学習が成立するかを含めた議論が必要である。

この観点で特に気になるのが、評価に用いられるデータセットの鮮度である。対象文献では CIC-IDS2017 や CIC-DDoS2019 など、広く流通した公開データセットへの依存が強い一方、近年の攻撃トレンドや運用環境の変化トラフィック構成の変動、機器の更新、防御側の設定差などを十分に反映できていない可能性がある。データセットが古いほど、モデルが学習しているのは当時の特徴量分布やラベルに最適化されたものになりやすく、ゼロデイや亜種への検知能力を過大評価するリスクが残る。

また、データセットのクロスバリデーションや横断評価が少ない点も、未知条件の議論を弱くする。単一のホールドアウト分割では、偶然の分割に依存した結果になりやすく、真に未知条件へ汎化できているのかが見えにくい。

以上を踏まえると、未知攻撃への適応能力は「高い精度が出た」という一点ではなく、以下の 3 点をセットで示して説得力を上げる必要がある：

1. データセットが現在の環境へ適用できることや使用理由の妥当性。

2. クロスバリデーションによる性能テスト。

3. 攻撃時や通信悪化を含む学習成立性。

特に連合学習では、データ分布の偏りと通信制約が同時に発生するため、未知の条件の評価設計を丁寧に置くことが、研究結果を現実運用議論へ接続する上での要件になる。

5.6 本章のまとめ

本章では、連合学習を用いた IoT 環境下の DoS/DDoS 検知研究について、第 4 章で整理運用ギャップの観点から解釈し、どの仮定が現実運用を難しくしているのかを明確化した。まず 5.1 節では、既存研究が精度中心の議論に偏りやすく、通信、応答時間、デバイス、データセットといった運用要件が評価に十分反映されない点を全体像として整理した。

5.2 節から、論文の情報を統計し、通信削減の工夫がそのまま更新の成立性の保証を意味しないことを示し、帯域枯渇下では更新欠落や遅延が学習過程に直接影響するという前提に立ち、評価で何を分離して測るべきかを明確にした。

5.3 節では、運用ギャップを生む主要因を三つの軸に集約した。第一に、リアルタイム性を論じるには、推論時間だけでなく、検知から緩和までの応答遅延として計測範囲を揃える必要がある。また、評価環境と実機評価の不足は、攻撃時での成立条件を曖昧で、computation 中心の評価では通信悪化と学習、検知の相互作用が観測されにくい。その上、展開シナリオと脅威モデルが対応していない場合、「検知できるか」と「攻撃下で運用できるか」が分離して議論されやすく、結果が過度に楽観的に解釈される恐れがある。

5.4 節では、研究で使用するデータセットは古めの問題存在する。2020 年以前のデータセットを使用することは、現在のネットワーク環境に適していない可能性が存在する。以上より、本分野は検知性能に関する知見を蓄積しつつある一方で、帯域枯渇・性能制約・制御面過負荷といった運用上の枯渇点に対して、どの条件で成立し、どの条件で破綻するかを比較可能な形で示す報告が限られている。

5.5 節では、今後の研究は、通信効率と成立性を分けて評価し、未知攻撃と環境変動への適応をデータセットに影響されることを意識するべきだと、その上、低性能端末を含む実機条件での計測をすることを明示する必要があるという観点を整理した。

第6章　結論

本調査では、IoT ネットワークにおける DoS/DDoS 検知を対象として、連合学習を用いる文献を 50 本整理し、検知モデルの系統と、現実運用へ繋ぐ評価観点を文献横断で比較した。その結果、連合学習はデータを外部へ集約せずに検知器を更新できるという点で有望であり、GNN などの考え方を加えると分散的のネットワークを構造化して性能を上げることができる一方、DoS/DDoS の本質である通信悪化、CPU やメモリの枯渇が、学習更新と検知応答に同時に影響するという前提を評価設計へ落とし込めていない研究が多く、精度中心の議論だけでは現実運用の要件「更新の成立、応答の即時性、攻撃下での継続性」を説明し切れないことが明らかになった。

技術整理　第 4 章の整理から、検知器の高度化（CNN/DNN/RNN/GNN/Autoencoder/Transformer）に加え、Non-iid への対処、個人化、ロバスト集約、通信効率化など、連合学習 特有の難しさに向き合う研究は増加している。一方で、これらの工夫はしばしば「検知精度の改善」として提示される一方、DoS/DDoS 下で問題となる「通信が乱れた状態でも更新が成立するか」「成立しないときにどう振る舞うか」という運用論点とは分離して評価されがちである。

運用ギャップ　第 5 章の整理から、運用ギャップは大きく次の形で現れる。第一に、環境カテゴリの分布として computation が多数を占める点である（図 5.2）。computation では、モデル更新の送受信を実行した根拠が薄く、通信遅延・欠落更新・同期待ちといった影響が原理的に観測されない。そのため、得られた精度や応答性が攻撃下の運用条件へどこまで外挿可能かを文献横断で議論しにくい。とりわけ、クライアント間の非同期性や回線品質の揺らぎが評価に反映されない場合、集約の安定性や更新成立性に関する議論が理想条件に偏り、破綻点（どの条件で学習が止まるか／性能が崩れるか）を比較できなくなる。

その上、実機評価のカバレッジが限定的である点である（図 5.3）。IoT を想定しつつも、低性能デバイスでの実測は少数に留まり、多くは実機評価なし（あるいは計測範囲が限定的）である。その結果、「エッジで動作可能」「リアルタイム」といった主張を、CPU・メモリ・電力・通信条件といった分母を伴って裏付けることが難しい。さらに、実機評価が存在する場合でも、推論のみ／前処理込み／学習更新・通信込みといった計測範囲が統一されていなければ、数値の比較可能性は限定される。

また、展開シナリオと脅威モデルの対応付けが弱い点である。「どこで検知し、どこで学習・集約し、どこで緩和動作が起きるか」が曖昧なまま評価されると、攻撃が枯渇させる対象（帯域、制御面、端末性能）と、評価で実際に与えた負荷・観測した指標が噛み合わない。この不整合は、「検知できる」ことと「攻撃下で運用できる」ことが分離して扱われる温床となり、実運用を論じる際の解釈を過度に楽観的にする。

結論　以上を踏まえると、IoT における DoS/DDoS を対象とした連合学習型検知の価値を運用へ接続するには、精度向上だけでなく、通信と配置を含むシステム設計と、比較できる報告様式の整備または統一が不可欠である。少なくとも、以下の要件を明示する必要があると考える：

- 計測範囲（推論／前処理／学習更新／通信モデルの大きさなど、どこまで含めたか）
- 通信条件（遅延・損失・帯域幅制限、更新欠落の扱い）
- 更新成立性（何が成立の条件で、どの条件で破綻するか）
- 展開シナリオ（検知・学習・集約・緩和の位置）と脅威モデル（何が枯渇するか）の対応
- ハードウェア条件（端末・エッジの性能上限と実測）

その上、精度指標と並べて提示することが、攻撃されている時の運用議論の出発点となる。

今後に向けて　本調査で整理した運用ギャップは，裏を返せば研究余地の大きさを示している。今後は，通信効率とロバスト性の両立，未知攻撃への適応，評価指標の統一，低性能端末を含む実証，およびデータセット更新を含めた評価設計の高度化が重要である。加えて，研究の重心を「高精度」から「現実運用へ繋ぐ検証・設計」へ移していく必要があると考える。

第7章　謝辞

本調査を進めるにあたり、終始多大なるご指導と貴重なご助言を賜りましたBEURAN Razvan 准教授に、心より感謝申し上げます。調査の過程では、課題点をご指摘いただくとともに、研究の方向性について丁寧にご指導いただき、多くを学ぶことができました。また、BEURAN 研究室の皆様には、議論や助言を通じて多くの支援をいただきました。サイバーセキュリティ分野に不慣れな私にとって、大変心強い支えとなりました。さらに、論文審査ご審査をご担当いただく丹康雄教授、リム勇仁教授、宇多仁准教授に、深く御礼申し上げます。最後に、これまで温かく支えてくれた家族および友人に、心より感謝いたします。

（※生成 AI に関して）本研究の遂行にあたり、生成 AI を以下の目的に活用させていただきました：

- 文献の翻訳と理解の補助
- 情報、データの整理
- 文章表現の改善

※使用ツール：ChatGPT-5.2

参考文献

[1] O. Yoachimik and J. Pacheco、“Cloudflare's 2025 Q3 DDoS threat report – including Aisuru、 the apex of botnets、” Cloudflare Quarterly DDoS Threat Report、 Q3 2025、 2025.

[2] O. Yoachimik and J. Pacheco、“Hyper-volumetric DDoS attacks skyrocket: Cloudflare's 2025 Q2 DDoS threat report、” Cloudflare Quarterly DDoS Threat Report、 Q2 2025、 2025.

[3] Nokia、*Threat Intelligence Report 2024: What's next for telecom: Emerging trends and technologies*、2024.

[4] Bitdefender、*The 2024 IoT Security Landscape Report*、Jun. 28、 2024.

[5] Dragos、*OT/ICS Cybersecurity Report: 8th Annual Year in Review*、2025.

[6] VicOne、*Shifting Gears: VicOne 2025 Automotive Cybersecurity Report*、2025.

[7] R. Singh and S. S. Gill、“Edge AI: A survey、” *Internet of Things and Cyber-Physical Systems*、vol. 3、 pp. 71–92、 2023、 doi:10.1016/j.iotcps.2023.02.004.

[8] P. Sun、 S. Shen、 Y. Wan、 Z. Wu、 Z. Fang、 and X.-Z. Gao、“A Survey of IoT Privacy Security: Architecture、 Technology、 Challenges、 and Trends、” *IEEE Internet of Things Journal*、vol. 11、 no. 21、 Nov. 2024、 doi:10.1109/JIOT.2024.3372518.

[9] M. O. Ojo、 D. Adami、 and S. Giordano、“A SDN-IoT Architecture with NFV Implementation、” in *Proc. 2016 IEEE Globecom Workshops (GC Wkshps)*、Washington、 DC、 USA、 Dec. 2016、 pp. 1–6、 doi:10.1109/GLOCOMW.2016.7848825.

[10] M. Chen, Y. Miao, Y. Hao, and K. Hwang, “Narrow Band Internet of Things,” *IEEE Access*, vol. 5, pp. 20557–20577, 2017, doi:10.1109/ACCESS.2017.2751586.

[11] J. Xu, J. Yao, L. Wang, Z. Ming, K. Wu, and L. Chen, "Narrowband Internet of Things: Evolutions, Technologies, and Open Issues," *IEEE Internet of Things Journal*, vol. 5, no. 3, pp. 1449–1462, Jun. 2018, doi:10.1109/JIOT.2017.2783374.

[12] A. Khraisat、 I. Gondal、 P. Vamplew、 and J. Kamruzzaman、"Survey of intrusion detection systems: techniques、 datasets and challenges、" *Cybersecurity*、 vol. 2、 art. 20、 2019、 doi: 10.1186/s42400-019-0038-7.

[13] CAIDA. *DDoS Attack 2007 Dataset*. 2007. `https://www.caida.org/catalog/datasets/ddos-attack-2007/`.

[14] Canadian Institute for Cybersecurity (CIC)、 University of New Brunswick. *Intrusion Detection Evaluation Dataset (CIC-IDS2017)*. 2017. `https://www.unb.ca/cic/datasets/ids-2017.html`.

[15] Canadian Institute for Cybersecurity (CIC)、 University of New Brunswick. *CSE-CIC-IDS2018 (Intrusion Detection Evaluation Dataset)*. 2018. `https://www.unb.ca/cic/datasets/ids-2018.html`.

[16] Canadian Institute for Cybersecurity (CIC)、 University of New Brunswick. *CIC-DoS2017 Dataset* (public mirror). 2017. `https://www.kaggle.com/datasets/dhoogla/cicdos2017`.

[17] Canadian Institute for Cybersecurity (CIC)、 University of New Brunswick. *DDoS Evaluation Dataset (CIC-DDoS2019)*. 2019. `https://www.unb.ca/cic/datasets/ddos-2019.html`.

[18] Canadian Institute for Cybersecurity (CIC)、 University of New Brunswick. *CICIoT2023 Dataset*. 2023. `https://www.unb.ca/cic/datasets/iotdataset-2023.html`.

[19] UNSW Canberra Cyber. *UNSW-NB15 Dataset*. 2015. `https://research.unsw.edu.au/projects/unsw-nb15-dataset`.

[20] UNSW Canberra Cyber. *Bot-IoT Dataset*. 2018. `https://research.unsw.edu.au/projects/bot-iot-dataset`.

[21] UNSW Canberra Cyber. *ToN_IoT Datasets*. 2020. `https://research.unsw.edu.au/projects/toniot-datasets`.

[22] Sebastian Garcia、Agustin Parmisano、and Maria Jose Erquiaga. *IoT-23: A labeled dataset with malicious and benign IoT network traffic (Version 1.0.0)*. Zenodo、 2020. `https://doi.org/10.5281/zenodo.4743746`.

[23] Mohammad A. Ferrag et al. *Edge-IIoTset: A New Comprehensive Realistic Cyber Security Dataset of IoT and IIoT Applications for Centralized and Federated Learning*. IEEE Access、 2022. (Dataset is distributed via public data repository referenced by the paper.)

[24] ASEA/ADOS、 University College Dublin. *InSDN: A Novel SDN Intrusion Dataset* (download site). 2020. `https://aseados.ucd.ie/datasets/SDN/`.

[25] IoTID20 Project. *IoTID20 Dataset*. 2020. `https://sites.google.com/view/iotid20`.

[26] I. Vaccari、 G. Chiola、 M. Aiello、 M. Mongelli、 and E. Cambiaso. MQTTset、 a new dataset for machine learning techniques on MQTT. *Sensors*、 20(22):6578、 2020. `https://www.mdpi.com/1424-8220/20/22/6578`.

[27] UCI Machine Learning Repository. *N-BaIoT: Network-based Detection of IoT Botnet Attacks Dataset*. 2018. `https://archive.ics.uci.edu/dataset/442/n+ba+iot`.

[28] A. Almomani et al. WSN-DS: A dataset for intrusion detection systems in wireless sensor networks. *Journal of Sensors*、 2016. `https://onlinelibrary.wiley.com/doi/10.1155/2016/4731953`. (Public mirrors exist; verify the exact version you used.)

[29] M. Tavallaee、 E. Bagheri、 W. Lu、 and A. A. Ghorbani. A detailed analysis of the KDD CUP 99 data set. In *Proceedings of IEEE CISDA*、 2009. (NSL-KDD is a refined version; public mirrors such as Figshare host KDDTrain+/KDDTest+.)

[30] NIDS Datasets Repository (CIC/UNSW derived NetFlow datasets). *CIC-ToN-IoT (NetFlow/standard feature version)*. 2023. `https://researchdata.edu.au/cic-ton-iot/2784357`.

[31] NIDS Datasets Repository. *NF-ToN-IoT-v2*. 2023. `https://researchdata.edu.au/nf-ton-iot-v2/2784371`.

[32] NIDS Datasets Repository. *NF-BoT-IoT-v2*. 2023. `https://researchdata.edu.au/nf-bot-iot-v2/2784366`.

[33] NIDS Datasets Repository. *NF-CSE-CIC-IDS2018-v2*. 2023. `https://researchdata.edu.au/nf-cse-cic-ids2018-v2/2784381`.

[34] NCSRD. *NCSRD-DS-5GDDDoS Dataset*. 2024. `https://zenodo.org/records/11578711`.

[35] Hacking and Countermeasure Research Lab (HCRL). *Car-Hacking Dataset*. 2016. https://ocslab.hksecurity.net/Datasets/car-hacking-dataset.

[36] Hacking and Countermeasure Research Lab (HCRL). *OTIDS Dataset*. 2018. https://ocslab.hksecurity.net/Datasets/OTIDS.

[37] Hacking and Countermeasure Research Lab (HCRL). *CAN Signal Extraction and Translation Dataset*. (Year not explicitly stated on the download page). https://ocslab.hksecurity.net/Datasets/can-signal-extraction-and-translation-dataset.

[38] WebNLG Community. *WebNLG Dataset (releases)*. 2017–. https://github.com/fuzihaofzh/webnlg-dataset.

[39] A. Naz、 I. Ullah、 M. Uzair、 M. F. Khokhar、 A. Sabir、 and R. U. Khan、 "AFL-SecNet: An adaptive federated learning framework for secure and privacy-preserving network traffic analysis、" *Peer-to-Peer Networking and Applications*、 vol. 19、 art. no. 25、 2026、 doi:10.1007/s12083-025-02155-w.

[40] D. Chiaro、 P. Qi、 E. Prezioso、 A. Guzzo、 and F. Piccialli、 "FLAME: Federated learning for attack mitigation and evasion、" in *Proc. IEEE International Parallel and Distributed Processing Symposium (IPDPS)*、 2025、 doi:10.1109/IPDPS64566.2025.00060.

[41] Y. K. Beshah、 S. L. Abebe、 and H. M. Melaku、 "Dynamic weight clustered federated learning for IoT DDoS attack detection、" *Scientific Reports*、 vol. 15、 art. no. 34036、 2025、 doi:10.1038/s41598-025-13204-y.

[42] N. Albanbay、 Y. Tursynbek、 K. Graffi、 R. Uskenbayeva、 Z. Kalpeyeva、 Z. Abilkaiyr、 and Y. Ayapov、 "Federated learning-based intrusion detection in IoT networks: Performance evaluation and data scaling study、" *Journal of Sensor and Actuator Networks*、 vol. 14、 art. no. 78、 2025、 doi:10.3390/jsan14040078.

[43] Y. Alhasawi and S. Alghamdi、 "Federated learning for decentralized DDoS attack detection in IoT networks、" *IEEE Access*、 vol. 12、 pp. 42357–42368、 2024、 doi:10.1109/ACCESS.2024.3378727.

[44] Z. Qu and Z. Cai、 "FEDSA-ResnetV2: An efficient intrusion detection system for vehicle road cooperation based on federated learning、" *IEEE Internet of Things Journal*、 vol. 11、 no. 18、 pp. 29852–29863、 Sep. 2024、 doi:10.1109/JIOT.2024.3410871.

[45] J. Wang、 K. Yang、 and M. Li、"NIDS-FGPA: A federated learning network intrusion detection algorithm based on secure aggregation of gradient similarity models、" *PLOS ONE*、 vol. 19、 no. 10、 e0308639、 2024、 doi:10.1371/journal.pone.0308639.

[46] C. Lohani and A. L. Sangal、"ARAFL-BAD: adversarial resilient asynchronous federated learning framework for zero-day IoT botnet attack detection、" *Sadhana*、vol. 50、 Art. no. 302、 2025、 doi:10.1007/s12046-025-02958-w.

[47] M. M. Salim、 D. Camacho、 and J. H. Park、"Digital Twin and Federated Learning Enabled Cyberthreat Detection System for IoT Networks、" *Future Generation Computer Systems*、vol. 161、 pp. 701–713、 2024、 doi:10.1016/j.future.2024.07.017.

[48] R. Doriguzzi-Corin and D. Siracusa、"FLAD: Adaptive Federated Learning for DDoS attack detection、" *Computers & Security*、vol. 137、 Art. no. 103597、 2024、 doi:10.1016/j.cose.2023.103597.

[49] M. N. Ali、 M. Imran、 M. Salah ud Din、 and B.-S. Kim、"Low Rate DDoS Detection Using Weighted Federated Learning in SDN Control Plane in IoT Network、" *Applied Sciences*、vol. 13、 no. 3、 Art. no. 1431、 2023、 doi:10.3390/app13031431.

[50] E. C. Pinto Neto、 S. Dadkhah、 and A. A. Ghorbani、"Collaborative DDoS Detection in Distributed Multi-Tenant IoT using Federated Learning、" in *Proc. 2022 19th Annual International Conference on Privacy、 Security & Trust (PST)*、2022、 doi:10.1109/PST55820.2022.9851984.

[51] R. Kalakoti、 J. Shetty、 and H. Bahsi. LSTM-based federated explainable AI for intrusion detection systems. *Computer Networks*、 vol. 270、 Art. no. 111479、 2025. doi: 10.1016/j.comnet.2025.111479.

[52] M. A. Alohali、H. Dafaalla、M. Baihan、S. Alahmari、A. Ben Miled、O. Alrusaini、 A. Alqazzaz、 and H. Alkhudhayr. Leveraging self attention driven gated recurrent unit with crocodile optimization algorithm for cyberattack detection using federated learning framework. *Scientific Reports*、 vol. 15、 Art. no. 23805、 2025. doi: 10.1038/s41598-025-99452-4.

[53] M. Bukhari、 B. Rababah、 M. A. Javed、 H. Wang、 N. Kumar、 and A. Al-Fuqaha. Secure and privacy-preserving intrusion detection in wireless sensor networks with federated learning: A BiLSTM-based

SCNN approach. *Ad Hoc Networks*、 vol. 155、 Art. no. 103407、 2024. doi: 10.1016/j.adhoc.2024.103407.

[54] V. Mothukuri、 P. Khare、 R. M. Parizi、 S. Pouriyeh、 A. Dehghantanha、 and G. Srivastava. Federated learning-based anomaly detection for IoT security attacks. *IEEE Internet of Things Journal*、 vol. 9、 no. 4、 pp. 2545–2554、 2022. doi: 10.1109/JIOT.2021.3077803.

[55] J. Li、 Z. Zhang、 Y. Li、 X. Guo、 and H. Li. FIDS: Detecting DDoS through federated learning based method. In *Proceedings of the 2021 IEEE 20th International Conference on Trust、 Security and Privacy in Computing and Communications (TrustCom)*、 pp. 856–862、 2021. doi: 10.1109/TrustCom53373.2021.00121.

[56] J. Zhang、 P. Yu、 L. Qi、 S. Liu、 H. Zhang、 and J. Zhang. FLDDoS: DDoS attack detection model based on federated learning. In *Proceedings of the 2021 IEEE 20th International Conference on Trust、 Security and Privacy in Computing and Communications (TrustCom)*、 pp. 635–642、 2021. doi: 10.1109/TrustCom53373.2021.00095.

[57] Y. Wang、 J. Li、 Z. Han、 P. Cheng、 and R. Kumar、 "FedSTGCN: A novel approach to network intrusion detection based on federated spatiotemporal graph learning for the IoT、 " *Frontiers of Information Technology & Electronic Engineering*、 vol. 26、 no. 7、 pp. 1164–1179、 2025、 doi: 10.1631/FITEE.2400932.

[58] F. Al Tfaily、 Z. Ghalmane、 M. E. Brahmia、 H. Hazimeh、 A. Jaber、 and M. Zghal、 "Federated Graph Attention and graphSAGE-based learning for intrusion detection、 " *Scientific Reports*、 vol. 15、 Art. no. 41264、 2025、 doi: 10.1038/s41598-025-25175-1.

[59] J. Li and Y. Qi、 "Security-Aware Federated Graph Neural Network with Attention-Based Aggregation and Function-Hiding Encryption、 " *IEICE Transactions on Fundamentals of Electronics、 Communications and Computer Sciences*、 Early Access、 2025、 doi: 10.1587/transfun.2025EAP1052.

[60] X. Ma、 J. Hu、 S. Liang、 and Y. Wu、 "Federated Learning and Resource-Aware Graph Neural Network for Intrusion Detection in 6G-IoT Driven Healthcare System、 " *IEEE Internet of Things Journal*、 Early Access、 2025、 doi: 10.1109/JIOT.2025.3564722.

[61] M. Anjum、 A. K. Dutta、 A. Elrashidi、 S. Shahab、 A. Aldrees、 Z. A. Shaikh、 and A. Aljohani、"GraphFedAI framework for DDoS attack detection in IoT systems using federated learning and graph based artificial intelligence、" *Scientific Reports*、vol. 15、 Art. no. 28050、 2025、 doi: 10.1038/s41598-025-10826-0.

[62] Q. Mao、X. Lin、Y. Qi、G. Li、J. Li、W. Xu、and X. Su、"FeCoGraph: Label-Aware Federated Graph Contrastive Learning for Few-Shot Network Intrusion Detection、" *IEEE Transactions on Information Forensics and Security*、vol. 20、 pp. 2266–2280、 2025、 doi: 10.1109/TIFS.2025.3541890.

[63] S. Sanjalawe、S. M. Hossain、M. A. Rahman、A. Hayakawa、S. J. Kwon、and K.-S. Kwak、"Adaptive graph attention-based federated learning for IoT intrusion detection against model poisoning attacks、" *PeerJ Computer Science*、vol. 11、 e3281、 2025、 doi: 10.7717/peerj-cs.3281.

[64] M. K. Islam、 M. S. Othman、 M. I. Gheyath、 and T. M. Ghazal、"Federated Deep Graph Neural Network Algorithm for Real-Time Intrusion Detection in Heterogeneous IoT Systems、" *Journal of Internet Services and Information Security*、vol. 15、no. 4、pp. 484–497、2025、doi: 10.58346/JISIS.2025.I4.035.

[65] J. Wu、 G. Qiu、 C. Wu、 W. Jiang、 and J. Jin、"Federated learning for network attack detection using attention-based graph neural networks、" *Scientific Reports*、vol. 14、 Art. no. 19088、 2024、 doi: 10.1038/s41598-024-70032-2.

[66] X. Zhang、Y. Wang、S. Xie、Y. Han、and S. Wang、"Federated Graph Neural Network for Fast Anomaly Detection in Controller Area Networks、" *IEEE Transactions on Information Forensics and Security*、vol. 18、pp. 1566–1579、2023、 doi: 10.1109/TIFS.2023.3240291.

[67] M. Chen、 B. Wu、 H. Sun、 and Z. Wang、"FedGAN-ID: Federated-Learning-Based Intrusion Detection for In-Vehicle Network Using GANs、" *IEEE Internet of Things Journal*、vol. 12、 no. 17、 pp. 36155–36167、Sep. 2025、 doi:10.1109/JIOT.2025.3580537.

[68] G. Shirvani、 S. Ghasemshirazi、 and M. Abdollahi Azgomi、"Mitigating DDoS attacks on the internet of things using federated learning、" *Peer-to-Peer Networking and Applications*、vol. 18、 Art. no. 198、 2025、doi:10.1007/s12083-025-02001-z.

[69] B. Olanrewaju-George and B. Pranggono、“Federated learning-based intrusion detection system for the IoT using unsupervised and supervised deep learning models、” *Cyber Security and Applications*、vol. 3、Art. no. 100068、2025、doi:10.1016/j.csa.2024.100068.

[70] V. T. Nguyen and R. Beuran、“FedMSE: Semi-supervised federated learning approach for IoT network intrusion detection、” *Computers & Security*、vol. 151、Art. no. 104337、2025、doi:10.1016/j.cose.2025.104337.

[71] P. Munaweera、S. Prasad、T. Hewa、Y. Siriwardhana、and M. Yliant-tila、“Federated Learning-powered DDoS Attack Detection for Securing Cyber Physical Systems in 5G and Beyond Networks、” in *Proceedings of the 14th International Conference on the Internet of Things (IoT '24)*、2024、pp. 1–12、doi:10.1145/3671047.3709015.

[72] A. Khacha、R. Saadouni、M. R. Ouhadi、and E. M. Laadissi、“Federated Learning and Convolutional Autoencoder for Robust Anomaly Detection in Agricultural IoT、” in *2024 International Conference of the African Federation of Operational Research Societies (AFROS)*、2024、pp. 1–6、doi:10.1109/AFROS62115.2024.11037027.

[73] R. Majeed and A. L. Sangal、“Enhancing IoT Security: Federated Learning with Autoencoder Model for IoT Attacks Detection、” in *2024 International Conference on Advances in Modern Age Technologies for Health and Engineering Science (AMATHE)*、2024、pp. 1–6、doi:10.1109/AMATHE61652.2024.10582233.

[74] F. Zeng、Q. Li、J. Wen、and H. Li、“FLoV2T: Federated learning based IoT traffic classification using lightweight vision transformer、” *Computer Communications*、vol. 242、Art. no. 108288、2025、doi:10.1016/j.comcom.2024.108288.

[75] A. A. Alshdadi、A. A. Almazroi、N. Ayub、M. D. Lytras、E. Alsolami、F. S. Alsubaei、and R. Alharbey、“Federated Deep Learning for Scalable and Privacy-Preserving Distributed Denial-of-Service Attack Detection in Internet of Things Networks、” *Future Internet*、vol. 17、no. 2、Art. no. 88、2025、doi:10.3390/fi17020088.

[76] F. Zheng and C. Pu、“SecureFogDL: A federated Transformer framework for secure and intelligent fog computing in healthcare IoT、” *Internet of Things*、vol. 30、Art. no. 101566、2025、doi:10.1016/j.iot.2024.101566.

[77] Y. Luo、 X. Chen、 H. Sun、 X. Li、 N. Ge、 W. Feng、 and J. Lu、"Securing 5G/6G IoT Using Transformer and Personalized Federated Learning: An Access-Side Distributed Malicious Traffic Detection Framework、" *IEEE Open Journal of the Communications Society*、vol. 5、pp. 1325–1339、2024、doi:10.1109/OJCOMS.2024.3365976.

[78] E. S. Adjewa、G. Beaucagin、C. Grecos、and P. K. Oppong、"Efficient Federated Intrusion Detection in 5G Ecosystem Using Optimized BERT-Based Model、" in *Proc. IEEE International Conference on Wireless and Mobile Computing、 Networking and Communications (WiMob)*、 2024、 pp. 518–523、 doi:10.1109/WiMob64165.2024.10822130.

[79] Y. Xu、 S. Zhang、 C. Lyu、 J. Liu、 Y. Shen、 and N. Shiratori、"Mitigating Distributed DoS Attacks on Bandwidth Allocation for Federated Learning in Mobile Edge Networks、" *IEEE Transactions on Dependable and Secure Computing*、vol. 22、 no. 3、 pp. 1941–1960、 May/Jun. 2025、 doi: 10.1109/TDSC.2024.3474299.

[80] A. Fotse、 P. Memarmoshrefi、 S. Barikbin、 and H. Karl、"FedLAD: Federated Learning for Low-Rate and Asynchronous DDoS Attack Detection in Multi-controller SDN、" *IEEE Transactions on Computers*、vol. 74、 no. 1、 pp. 101–115、 Jan. 2025、doi: 10.1109/TC.2024.3474180.

[81] L. V and S. Rajkumar、"A Hybrid Edge Federated Learning-based Detection and Mitigation of DDoS Attacks in Internet of Things Networks、" *Results in Engineering*、vol. 28、 Art. no. 107601、 2025、 doi: 10.1016/j.rineng.2025.107601.

[82] T. Q. Al-Ghadi、 S. Manickam、 and I. D. M. Widia、"Leveraging Federated Learning for DoS Attack Detection in IoT Networks Based on Ensemble Feature Selection and Deep Learning Models、" *Cyber Security and Applications*、 vol. 3、 Art. no. 100098、 2025、doi: 10.1016/j.csa.2024.100098.

[83] A. Rezapour、 A. Khodadadi、 S. K. Das、 and D. Grosu、"CAFNet: Compressed Autoencoder-based Federated Network for Anomaly Detection、" in *Proc. IEEE MILCOM 2023 (MILCOM)*、2023、doi: 10.1109/MILCOM58377.2023.10356377.

[84] R. Yang、 H. He、 Y. Wang、 B. Luo、 and Y. Tang、"Dependable Federated Learning for IoT Intrusion Detection Against Poisoning Attacks、" *Computers & Security*、vol. 132、 Art. no. 103381、 2023、 doi: 10.1016/j.cose.2023.103381.

[85] M. Dimolianis、D. K. Kalogeras、N. Kostopoulos、and V. Maglaris、“DDoS Attack Detection via Privacy-aware Federated Learning and Collaborative Mitigation in Multi-domain Cyber Infrastructures、” in *Proc. 2022 IEEE 11th International Conference on Cloud Networking (CloudNet)*、 Oct. 2022.

[86] Z. Yin、 K. Li、 and H. Bi、“Trusted Multi-Domain DDoS Detection Based on Federated Learning、” *Sensors*、vol. 22、 no. 20、 Art. no. 7753、 2022、 doi: 10.3390/s22207753.

[87] O. Friha、 M. Ammar、 A. Abdelkader、 E. Bouraoui、 and J. M. S. Boubaker、“Federated Learning-Based Intrusion Detection in IoT Networks: Performance Evaluation and Data Scaling Study、” *Journal of Parallel and Distributed Computing*、vol. 165、 pp. 17–31、 2022、 doi: 10.1016/j.jpdc.2022.04.005.

[88] O. Shahid、 S. ul Islam、 A. Almogren、 and others、“Detecting Network Attacks Using Federated Learning for IoT Devices、” in *Proc. 2021 IEEE 29th International Conference on Network Protocols (ICNP)*、 2021.