RADemics

Federated Learning for Secure and Decentralized Anomaly Detection Across Networked Systems

P. Krishnamoorthy
Sasi Institute of Technology & Engineering.

13. Federated Learning for Secure and Decentralized Anomaly Detection Across Networked Systems

P. Krishnamoorthy, Associate Professor , Department of Computer Science and Engineering, Sasi Institute of Technology & Engineering , Tadepalligudem, West Godavari District, Andhra Pradesh - 534 101.krishnancse0206@gmail.com .

Abstract

Federated Learning (FL) has emerged as a promising approach for decentralized and privacy-preserving anomaly detection across distributed systems, particularly in large-scale networks. The integration of FL for scalable and efficient anomaly detection, addressing key challenges such as network and communication constraints, edge device limitations, and the scalability of machine learning models. Emphasis is placed on optimizing model aggregation strategies, reducing communication overhead, and leveraging local training to enhance performance. The chapter explores advanced techniques like asynchronous updates, model compression, and hierarchical aggregation to overcome data synchronization issues. Additionally, it discusses dynamic federation strategies that adapt to system load and the importance of data management for improved scalability. By addressing these critical aspects, the chapter provides a comprehensive framework for implementing FL-based anomaly detection in real-world, resource-constrained environments. The combination of innovative methodologies and practical insights presented here paves the way for deploying FL in diverse applications, ranging from IoT systems to large-scale industrial networks, ensuring robust and efficient anomaly detection without compromising security or scalability.

Keywords: Federated Learning, Anomaly Detection, Scalability, Edge Devices, Model Aggregation, Communication Efficiency.

Introduction

Federated Learning (FL) has gained significant attention in recent years due to its ability to enable decentralized machine learning across distributed systems [1]. In traditional machine learning paradigms, data is collected and stored in a centralized location before training models [2]. With the advent of technologies such as IoT, edge computing, and industrial sensor networks, data is often distributed across numerous devices [3]. This creates challenges in terms of data privacy and security, particularly when dealing with sensitive or proprietary information [4]. FL addresses these challenges by allowing models to be trained collaboratively across edge devices without the need to share raw data [5]. This is particularly useful in applications like anomaly detection, where timely identification of outliers or system failures is critical to maintaining the integrity of large-scale systems [6].

Anomaly detection is a crucial task across various domains, including cybersecurity, industrial systems, healthcare, and finance [7]. Detecting deviations from normal behavior can help identify potential threats, faults, or inefficiencies [8]. In traditional anomaly detection methods, data from different sources is collected and centralized, which may expose the system to privacy risks and increased latency [9]. Federated Learning offers a decentralized approach to anomaly detection, ensuring that sensitive data remains on local devices while models are updated collaboratively [10]. As the number of devices in these systems grows, significant challenges arise, such as computational constraints, communication overhead, and difficulties in model synchronization [11]. These factors need to be carefully managed to ensure that FL-based anomaly detection remains effective and scalable in large-scale systems [12].

One of the primary challenges in deploying Federated Learning for anomaly detection is the scalability of the system [13]. As the number of participating devices increases, so does the complexity of model training and aggregation [14]. In a typical FL setup, each device trains a local model using its own data, and periodically, these local models are aggregated to form a global model [15]. As more devices join the system, the communication costs associated with sharing model updates become substantial [16]. Large datasets and complex models can strain the computational resources of edge devices, leading to delays in training and inference [17]. To overcome these challenges, techniques such as model compression, pruning, and federated optimization algorithms have been proposed to enhance the scalability of FL systems for anomaly detection [18]. These methods aim to reduce the size of the models, optimize the training process, and minimize the amount of data exchanged between devices and central servers.

In addition to scalability, communication efficiency is another critical concern when implementing FL-based anomaly detection systems. In large-scale systems, the frequent exchange of model parameters between edge devices and central servers can create significant network traffic, especially in resource-constrained environments where bandwidth may be limited [19]. To address this, several communication-efficient strategies have been developed, such as sparsification and gradient quantization. These methods aim to reduce the size of the data transmitted, which can significantly decrease communication costs [20]. Additionally, techniques like local training and data aggregation can help minimize the frequency of communication, enabling edge devices to perform multiple training iterations locally before synchronizing with the global model. By improving communication efficiency, FL systems can become more viable for deployment in large-scale, distributed anomaly detection applications, where latency and bandwidth limitations are crucial factors [21].

Edge devices involved in FL-based anomaly detection systems are often limited by computational and storage resources, which can hinder the training of complex machine learning models [23]. These devices are typically constrained by factors such as processing power, memory, and battery life, which can limit their ability to handle large datasets or run resource-intensive models [24]. To overcome these challenges, it is essential to design lightweight models that can operate efficiently on edge devices while maintaining high detection performance. Techniques like model distillation, transfer learning, and incremental learning can be employed to reduce the computational load on edge devices without sacrificing the accuracy of the anomaly detection process. The use of hybrid architectures that offload more computationally expensive tasks to cloud servers or powerful local machines can help ensure that edge devices can contribute effectively to the anomaly detection process without being overburdened [25].