ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

PRIVACY AND SECURITY ENHANCEMENTS OF FEDERATED LEARNING

M.Sc. THESIS

Şükrü ERDAL

Department of Applied Informatics

Cybersecurity Engineering and Cryptography Programme

JULY 2024

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

PRIVACY AND SECURITY ENHANCEMENTS OF FEDERATED LEARNING

M.Sc. THESIS

Şükrü ERDAL
(707211008)

Department of Applied Informatics

Cybersecurity Engineering and Cryptography Programme

Thesis Advisor: Prof. Dr. Enver ÖZDEMİR
Co-Advisor: Dr. Ferhat KARAKOÇ

JULY 2024

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

FEDERE ÖĞRENME UYGULAMALARINDA MAHREMİYET VE GÜVENLİK GELİŞTİRMELERİ

YÜKSEK LİSANS TEZİ

Şükrü ERDAL
(707211008)

Bilişim Uygulamaları Anabilim Dalı

Bilgi Güvenliği Mühendisliği ve Kriptografi Programı

Tez Danışmanı: Prof. Dr. Enver ÖZDEMİR
Eş Danışman: Dr. Ferhat KARAKOÇ

TEMMUZ 2024

Şükrü ERDAL, a M.Sc. student of ITU Graduate School student ID 707211008 successfully defended the thesis entitled "PRIVACY AND SECURITY ENHANCEMENTS OF FEDERATED LEARNING", which he prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

Thesis Advisor :	**Prof. Dr. Enver ÖZDEMİR** Istanbul Technical University	.....
Co-advisor :	**Dr. Ferhat KARAKOÇ** Ericsson	.....
Jury Members :	**Prof. Dr. Kemal BIÇAKÇI** Istanbul Technical University	.....
	Doç. Dr. Ergün YARANERİ Istanbul Technical University	.....
	Dr. Ramin FULADİ Ericsson	.....

Date of Submission : **24 May 2024**
Date of Defense : **12 July 2024**

To my spouse Hadiye,
This thesis is dedicated to you, my partner in life and my greatest supporter. Your patience, understanding, and many sacrifices have made it possible for me to pursue my academic goals. I am grateful for your unwavering love and encouragement.
With all my love.

FOREWORD

I am deeply thankful to my advisors, Prof. Dr. Enver Özdemir and Dr. Ferhat Karakoç, for their guidance, expertise, and support. As I embark on the next phase of my professional journey, I carry with me the lessons learned, the challenges overcome, and the memories shared.

July 2024 Şükrü ERDAL

TABLE OF CONTENTS

	Page
FOREWORD	**ix**
TABLE OF CONTENTS	**xi**
ABBREVIATIONS	**xiii**
LIST OF TABLES	**xv**
LIST OF FIGURES	**xvii**
SUMMARY	**xix**
ÖZET	**xxi**
1. INTRODUCTION	**1**
1.1 Purpose of Thesis	4
1.2 Literature Review	6
2. CHALLENGES, SOLUTIONS, AND APPLICATIONS OF FEDERATED LEARNING	**9**
2.1 Privacy and Security Challenges in Federated Learning	9
2.2 Privacy and Security Solutions in Federated Learning	12
2.3 Applications of Federated Learning	19
3. FEDERATED LEARNING SURVEYS	**23**
3.1 Roles and Benefits of Surveys in the Federated Learning Domain	23
3.1.1 Roles of surveys in FL	23
3.1.2 Benefits of surveys in FL	24
3.2 Survey Analysis in Federated Learning	25
4. METRICS AND STATE OF THE ART FEDERATED LEARNING STUDIES	**39**
5. CONCLUSIONS	**43**
5.1 Privacy and Security Challenges in Federated Learning	43
5.2 Privacy and Security Solutions in Federated Learning	43
5.3 Applications of Federated Learning	44
5.4 Surveys of Federated Learning	44
5.5 Comparison Metrics and State of the Art Federated Learning Studies	45
REFERENCES	**47**
CURRICULUM VITAE	**53**

ABBREVIATIONS

FL	**:** Federated Learning
IoT	**:** Internet of Things
ML	**:** Machine Learning
SMPC	**:** Secure Multi-Party Computation
HE	**:** Homomorphic Encryption
DP	**:** Differential Privacy
CC	**:** Confidential Computing
GDPR	**:** General Data Protection Regulation
EU	**:** European Union
IID	**:** Independent and Identically Distributed
FEDLOC	**:** Federated Localization
DeepSC	**:** Deep learning-enabled Semantic Communication
HIPAA	**:** Health Insurance Portability and Accountability Act
PET	**:** Privacy-Enhancing Technology
FDL	**:** Federated Deep Learning

LIST OF TABLES

Page

Table 4.1 : Metrics for Comparison of Security & Privacy Solutions **40**
Table 4.2 : Comparison of Security & Privacy Solutions..... **41**

LIST OF FIGURES

Page

Figure 2.1 : FL Challenges..... **12**
Figure 2.2 : FL Solutions..... **13**

PRIVACY AND SECURITY ENHANCEMENTS OF FEDERATED LEARNING

SUMMARY

Considering the deficiencies of machine learning, federated learning has been introduced as an alternative. FL promises and brings data privacy and security hardening mechanisms. In contrast to conventional machine learning models, which necessitate data aggregation on centralized servers, thereby increasing the susceptibility to data breaches and privacy violations, federated learning distributes the model training process across many decentralized edge devices. This approach maintains the confidentiality of raw data by limiting its storage and processing to localized devices, alleviating the potential risks that come with centralized data.

The motivation for this thesis arises from the increasing need for privacy and security improvements in FL applications. With more stringent data privacy regulations and heightened public awareness of data security, there is a significant demand for robust Federated Learning frameworks that can safeguard sensitive information while maintaining high model performance. The capability of Federated Learning to utilize the power of end nodes, such as smart endpoints, presents a promising solution for various fields: the Internet of Things, mobile communications, the health industry, and financial services.

The primary objectives of this thesis are threefold:
1. To comprehensively survey existing research on privacy-enhanced FL, synthesizing key concepts, methodologies, and findings.
2. To identify gaps, limitations, and open research questions in the current literature on privacy-enhanced FL.
3. To evaluate and compare different privacy-enhancing techniques and methodologies used in FL, assessing their effectiveness, scalability, and trade-offs.

FL inherently mitigates several privacy risks by keeping data local to clients. However, it introduces new challenges, particularly related to inference attacks and model update poisoning. Inference attacks exploit model updates to extract sensitive information, while model update poisoning involves malicious clients injecting false updates to corrupt the global model. These challenges require robust solutions to maintain FL system integrity and privacy.

Non-IID data and communication overheads further complicate FL implementation. Non-IID data, where data distributions vary across clients, can hinder model convergence and performance. Additionally, frequent and substantial data exchanges between clients and servers result in significant communication overheads, which can strain network resources.

Various strategies have been devised to tackle these privacy and security challenges. An example of this is differential privacy, which inserts noise into data updates in order to preserve the confidentiality of individual contributions. Protocols incorporating cryptographic signatures and Secure Multiparty Computation techniques bolster the security of model updates and maintain data integrity. Additionally, co-utility frameworks, which encourage mutual benefits for servers and clients, along with robust aggregation methods, are crucial in protecting Federated Learning systems.

Innovative methodologies such as Flamingo and SafeFL leverage advanced cryptographic techniques to provide secure aggregation and enhance privacy preservation. These solutions collectively improve the robustness, efficiency, and security of FL frameworks, enabling their application in real-world scenarios.

FL has been applied successfully in various domains, demonstrating its versatility and effectiveness. In wireless communication, FL enhances vehicular communication, localization, and semantic communication by enabling collaborative model training without data centralization. In the IoT sector, FL improves privacy and reduces data transfer costs, with significant applications in smart homes and industrial IoT.

Healthcare is another critical area where FL has made substantial impacts. By allowing institutions to collaboratively train models on medical imaging and predictive analytics without sharing patient data, FL addresses stringent privacy regulations while improving model accuracy and generalizability. Studies have shown that FL can maintain high diagnostic accuracy and support personalized medicine.

In the financial sector, FL addresses privacy and regulatory challenges by enabling collaborative credit risk assessment and fraud detection. By leveraging data from multiple institutions without centralizing it, FL-based models achieve higher accuracy and adaptability, enhancing the detection of fraudulent activities and improving credit scoring models.

Surveys play indispensable roles and offer numerous benefits within the FL domain. They serve as comprehensive repositories of existing research, providing newcomers with a foundational understanding while guiding experienced researchers toward unexplored frontiers. By scrutinizing and synthesizing a plethora of literature, surveys identify emerging trends, highlight successful applications, and outline future research directions.

Federated Learning offers an innovative approach to machine learning by enabling decentralized data processing, which mitigates the significant privacy and security concerns associated with traditional centralized models. This thesis examines various aspects of Federated Learning, particularly focusing on the challenges and solutions related to privacy and security, as well as its wide-ranging applications across different sectors.

Emerging trends in Federated Learning research, such as advancements in cryptographic techniques, federated learning frameworks, and compliance mechanisms, highlight the consistent innovation need with interdisciplinary collaboration. As Federated Learning progresses, it has the potential to revolutionize secure communication systems and promote a culture of security awareness and privacy by design in machine learning technologies.

FEDERE ÖĞRENME UYGULAMALARINDA MAHREMİYET VE GÜVENLİK GELİŞTİRMELERİ

ÖZET

Federe Öğrenme, makine öğrenimi alanında veri gizliliği ve güvenliği ile ilgili önemli endişelere çözüm getiren devrim niteliğinde bir yaklaşım olarak ortaya çıkmıştır. Geleneksel merkezi makine öğrenimi modelleri, verilerin merkezi sunucularda toplanmasını gerektirir, bu da veri ihlalleri ve gizlilik ihlalleri açısından büyük riskler oluşturur. Federe Öğrenme ise model eğitim sürecini birden fazla merkezi olmayan uç cihaz arasında dağıtarak, ham veriyi yerel tutar ve merkezi veri depolama ve işleme ile ilgili gizlilik risklerini azaltır.

Bu tezin motivasyonu, Federe Öğrenme uygulamalarında gizlilik ve güvenliği artırma ihtiyacından kaynaklanmaktadır. Veri gizliliği düzenlemeleri sıkılaştıkça ve kamuoyunun veri güvenliği konusundaki farkındalığı arttıkça, hassas bilgileri koruyabilen ve aynı zamanda yüksek model performansını sürdürebilen güçlü FL çerçevelerine olan talep de artmaktadır. Federe Öğrenme, akıllı telefonlar ve Nesnelerin Interneti cihazları gibi uç cihazların hesaplama gücünden yararlanarak, sağlık, finans ve Nesnelerin Interneti gibi çeşitli alanlarda umut verici bir çözüm sunar.

Bu tezin başlıca amaçları üç aşamada ele alınmıştır:
1. Gizlilik artırılmış Federe Öğrenme konusundaki mevcut araştırmaları kapsamlı bir şekilde inceleyerek ana kavramları, metodolojileri ve bulguları sentezlemek.
2. Mevcut gizlilik artırılmış Federe Öğrenme literatüründe boşlukları, sınırlamaları ve açık araştırma sorularını belirlemek.
3. Federe Öğrenme'de kullanılan farklı gizlilik artırma tekniklerini ve metodolojilerini değerlendirip karşılaştırarak, bunların etkinliğini, ölçeklenebilirliğini ve ödünleşimlerini analiz etmek.

Federe Öğrenme, verileri istemcilerde yerel tutarak birçok gizlilik riskini doğal olarak azaltır. Ancak, çıkarım saldırıları ve model güncelleme zehirlemesi gibi yeni zorluklar da beraberinde getirir. Çıkarım saldırıları, model güncellemelerini kullanarak hassas bilgileri çıkarma çabalarını içerirken, model güncelleme zehirlemesi, kötü niyetli istemcilerin yanlış güncellemeler yaparak küresel modeli bozmasını içerir. Bu zorluklar, Federe Öğrenme sürecinin bütünlüğünü ve gizliliğini sağlamak için sağlam çözümler gerektirir.

Ayrıca, Federe Öğrenme uygulamasını karmaşık hale getiren diğer faktörler arasında non-IID veri ve iletişim yükleri bulunmaktadır. Non-IID veri, veri dağılımlarının istemciler arasında farklılık göstermesini ifade eder ve bu durum model yakınsamasını ve performansını olumsuz etkileyebilir. Ek olarak, istemciler ve sunucu arasında sık ve büyük veri değişimlerinin gerekmesi, ağ kaynaklarını zorlayarak önemli iletişim yüklerine neden olabilir.

Bu gizlilik ve güvenlik zorluklarını ele almak için çeşitli stratejiler geliştirilmiştir. Diferansiyel gizlilik, veri güncellemelerine gürültü ekleyerek bireysel katkıların gizliliğini sağlar. Kriptografik imzalar ve Güvenli Çoklu Hesaplama tekniklerini içeren protokoller, model güncellemelerinin güvenliğini artırır ve veri bütünlüğünü korur. Ortak fayda çerçeveleri, sunucular ve istemciler arasında karşılıklı faydayı teşvik ederken, sağlam toplama yöntemleri Federe Öğrenme sistemlerinin güvenliğini ve gizliliğini güçlendirir.

Flamingo ve SafeFL gibi yenilikçi metodolojiler, gelişmiş kriptografik teknikler kullanarak güvenli toplama ve gizlilik korumasını sağlar. Bu çözümler, Federe Öğrenme çerçevelerinin gerçek dünya senaryolarında daha güvenli, verimli ve gizlilik dostu olmasını mümkün kılar.

Federe Öğrenme, çeşitli alanlarda başarılı bir şekilde uygulanarak çok yönlülüğünü ve etkinliğini göstermiştir. Kablosuz iletişimde, Federe Öğrenme araçlar arası iletişimi, konum belirlemeyi ve anlamsal iletişimi veri merkezileştirmeden geliştirir. Nesnelerin Interneti sektöründe, Federe Öğrenme gizliliği artırır ve veri aktarım maliyetlerini düşürür; akıllı evler ve endüstriyel Nesnelerin Interneti uygulamaları için önemli çözümler sunar.

Sağlık alanında, Federe Öğrenme, kurumların hasta verilerini paylaşmadan tıbbi görüntüleme ve tahmine dayalı analizler üzerinde işbirlikçi modeller eğitmesini sağlar. Bu sayede, katı gizlilik düzenlemelerine uyulurken model doğruluğu ve genelleştirilebilirlik artırılır. Çalışmalar, Federe Öğrenme'nin yüksek tanı doğruluğunu koruyabildiğini ve kişiselleştirilmiş tıbbı destekleyebildiğini göstermiştir.

Finans sektöründe, Federe Öğrenme gizlilik ve düzenleyici zorlukları ele alarak işbirlikçi kredi riski değerlendirmesi ve sahtekarlık tespiti sağlar. Birden fazla kurumdaki verileri merkezileştirmeden değerlendirerek, Federe Öğrenme tabanlı modeller daha yüksek doğruluk ve uyum kabiliyeti ile sahtecilik faaliyetlerini tespit eder ve kredi skorlaması modellerini iyileştirir.

Araştırmalar, Federe Öğrenme alanında vazgeçilmez bir rol oynar ve birçok fayda sağlar. Mevcut odaklı çalışmaların kapsamlı birer deposu olarak hizmet ederler, yeni başlayanlara temel bir anlayış sunarken deneyimli araştırmacıları keşfedilmemiş sınırlar yönünde yönlendirirler. Çok sayıda literatürü dikkatlice inceleyip sentezleyerek, ortaya çıkan eğilimleri belirler, başarılı uygulamaları vurgular ve gelecekteki araştırma yönelimlerini özetlerler.

Federe Öğrenme, verilerin merkezi olmayan bir şekilde işlenmesini sağlayarak geleneksel merkezi modellerdeki kritik gizlilik ve güvenlik sorunlarını ele alan dönüştürücü bir makine öğrenimi yaklaşımı sunar. Bu tez, Federe Öğrenme'nin gizlilik ve güvenlik ile ilgili çeşitli yönlerini, karşılaşılan zorlukları ve çözümleri, ayrıca farklı sektörlerdeki çeşitli uygulamalarını derinlemesine incelemektedir.

Federe Öğrenme araştırmalarında ortaya çıkan eğilimler, kriptografik tekniklerdeki gelişmeler, federe öğrenme çerçeveleri ve düzenleyici uyum mekanizmaları gibi sürekli yenilik ve disiplinler arası işbirliği ihtiyacını vurgulamaktadır. Federe Öğrenme, gelişmeye devam ettikçe, güvenli iletişim sistemlerini devrim niteliğinde değiştirme ve makine öğrenimi teknolojilerinde güvenlik farkındalığı ile gizliliğin tasarım aşamasında düşünülmesini teşvik etme potansiyeline sahiptir.

Bu çalışma, Federe Öğrenme sistemlerinin güvenliğini sağlama mekanizmaları hakkında içgörüler sunarak, araştırmacılar, uygulayıcılar ve politika yapıcılar için değerli bilgiler sağlamayı amaçlamaktadır. Kriptografik tekniklerin, güvenli toplama protokollerinin veya anomali tespit mekanizmalarının geliştirilmesi, güvenlik ihlallerini tespit edip önlemeye yardımcı olur. Ayrıca, sunucu ile katılımcı istemciler arasında güven tesis etmenin önemine değinir, bu da model güncellemelerinin güvenilirliğini ve doğruluğunu sağlar.

Sonuç olarak, Federe Öğrenme'nin veri gizliliğini ve güvenliğini artırarak merkezi olmayan bir makine öğrenimi yaklaşımı olarak önemli bir rol oynadığı vurgulanmaktadır. Bu tez, Federe Öğrenme alanındaki mevcut durumun kapsamlı bir değerlendirmesini sunarken, gelecekteki araştırmalar için de yol gösterici nitelikte önerilerde bulunmaktadır. Federe Öğrenme, verilerin yerel olarak işlenmesini sağlayarak, hem bireysel gizliliği korur hem de işbirlikçi öğrenme süreçlerini destekler. Bu bağlamda, tez, Federe Öğrenme'nin potansiyelini ve uygulanabilirliğini geniş bir bakış açısıyla ele almakta ve bu alandaki önemli katkıları ortaya koymaktadır. Disiplinler arası işbirliği ve kolektif çabalarla, Federe Öğrenme'nin daha güvenli ve gizlilik dostu bir dijital toplum oluşturma potansiyeline sahip olduğu sonucuna varılmaktadır.

1. INTRODUCTION

In an era where data is hailed as the new currency driving innovation and progress, the preservation of privacy stands as a cornerstone of ethical and legal considerations. As society becomes increasingly interconnected through digital platforms and smart devices, data types and their quantity generated are unprecedented. Yet, the pervasive gathering and yielding of increased data amount to profound weaknesses related to individual privacy rights, autonomy, and security.

Privacy has evolved from a fundamental human right to a critical aspect of modern governance and technological development. The digital age has brought about a paradigm shift in how personal information is collected, stored, analyzed, and shared. From targeted advertising to personalized services, the commodification of data has ushered in an era of unprecedented convenience and efficiency, but not without sacrificing privacy at times. [1]

Moreover, recent high-profile data breaches, surveillance revelations, and controversies surrounding data misuse have catalyzed public discourse on privacy. Individuals, policymakers, and technologists alike are grappling with the complexities of balancing the benefits of data-driven innovation with the imperative to safeguard privacy. [2]

In the era of big data and ubiquitous connectivity, the traditional centralized model of machine learning faces formidable challenges in scalability, privacy, and resource constraints. The exponential growth of data generated by diverse sources, including mobile devices, IoT sensors, and online platforms, has outpaced the capacity of centralized servers for data processing and storage. Moreover, concerns regarding data privacy, security breaches, and regulatory compliance have become increasingly prevalent, prompting a paradigm shift towards more privacy-preserving and decentralized machine learning approaches. [3]

FL has emerged as a promising solution to address these challenges by distributing the model training process across a network of decentralized edge devices. This novel paradigm enables collaborative learning while keeping raw data localized, thereby mitigating privacy risks associated with centralized data aggregation. FL harnesses the processing capabilities of end nodes, including smart devices and Internet of Things (IoT) gadgets, to conduct local model training and aggregation, thereby reducing the necessity for data transmission to centralized servers. This decentralized approach has the additional benefit of safeguarding data privacy while simultaneously reducing bandwidth consumption, latency, and computational overhead. [4]

FL has gained traction across various domains, including IoT, mobile communication, the health industry, and financial services where data privacy and security are paramount concerns. In healthcare, FL enables collaborative model training on sensitive patient data while adhering to strict privacy regulations. Similarly, in finance, FL facilitates secure model training on financial transactions while ensuring compliance with regulatory requirements. Additionally, in IoT applications, FL enables edge devices to collaboratively learn from sensor data while preserving user privacy and optimizing resource utilization. [5]

Despite its potential, FL presents several challenges, including communication overhead, model heterogeneity, and Byzantine failures. Communication overhead arises from the need to transmit model updates between end nodes and centralized servers, leading to increased bandwidth consumption and latency. Model heterogeneity refers to variations in data distribution, device capabilities, and network conditions across edge devices, which can affect the convergence and performance of federated models. Byzantine failures, such as malicious attacks or device failures, pose additional challenges to the security and reliability of FL systems. [6]

Recent advancements in FL research have focused on addressing these challenges through optimization techniques, privacy-preserving mechanisms, and robust communication protocols. Optimization techniques, such as federated averaging and model distillation, aim to improve convergence speed and model performance in FL systems. Privacy-preserving mechanisms, including differential privacy and secure aggregation,

provide strong guarantees against data leakage and privacy violations in federated settings. Robust communication protocols, such as gossip-based communication and adaptive aggregation, enhance the resilience and scalability of FL systems in dynamic environments. [7]

1.1 Purpose of Thesis

The purpose of this thesis is to conduct a comprehensive survey on privacy-enhanced federated learning, aiming to achieve the following objectives:

- The thesis seeks to synthesize existing research and literature on privacy-enhanced federated learning, providing a holistic view of the progress level of the field. By reviewing and summarizing key concepts, methodologies, and findings from a wide range of sources, this survey aims to consolidate the collective knowledge and insights in the domain of privacy-enhanced FL.
- Through a systematic review of the literature, the thesis aims to identify gaps, limitations, and open research questions in the field of privacy-enhanced federated learning. By critically analyzing existing approaches and methodologies, the survey intends to pinpoint areas where further research and development are needed, particularly in addressing emerging privacy challenges and scalability issues in FL designs.
- One of the central aims of the thesis is to evaluate and compare different privacy-enhancing techniques and methodologies used in federated learning. By evaluating the effectiveness, scalability, and trade-offs of different approaches such as differential privacy, secure aggregation, and homomorphic encryption, we aim to gain insight into the pros and cons of each technique and its applicability in different scenarios.
- In addition to evaluating privacy techniques, the thesis will explore the application of privacy-enhanced federated learning in specific domains and use cases. By examining real-world applications in areas such as IoT, mobile communications, financial services, and the health industry the survey seeks to identify domain-specific privacy challenges and considerations and assess how federated learning techniques can address them effectively.
- The thesis will provide practical guidance and recommendations for researchers in implementing privacy-enhanced federated learning systems. By synthesizing best

practices, lessons learned, and emerging trends from the literature, the survey aims to offer actionable insights and guidelines for designing, deploying, and managing federated learning systems that prioritize privacy protection while maximizing utility and performance.

- Finally, the comparison of different state-of-the-art studies will be carried out using several metrics that constitutes as baselines for different FL solutions. This section aims to bind many approaches to FL from a general view.

1.2 Literature Review

Several surveys have been conducted that focus on various aspects of FL so far. These surveys form the basis of this thesis.

In a recent survey, Al-Huthaifi et al. address privacy concerns in smart cities and how security and privacy can be improved using FL. [8] The applications of FL in smart cities are mainly focused on three categories: communications, healthcare, and transportation.

Another survey by Jie Wen et al. outlines the distributed training mode of FL and the security aggregation mechanism that makes it suitable for applications with strict privacy requirements. [9] Also, the paper systematically reviews FL research across five aspects: basics of FL, privacy and security mechanisms, challenges of communication overhead, and heterogeneity.

Soykan et al. emphasize the role of collaborative ML approaches like FL in addressing these privacy concerns, enabling the use of sensitive data while protecting critical features. [10] The paper provides a detailed analysis of the current state of collaborative ML methods from a privacy perspective, including a thorough examination of threat models and security considerations for each approach. It also delves into Privacy Enhancing Technologies such as secure multi-party computation, homomorphic encryption, differential privacy, and confidential computing within the context of collaborative ML.

Sirohi et al. analyze FL as a collaborative and distributed approach to training ML models, enabling the utilization of unlimited data and distributed computing power while preserving privacy by default. However, FL faces challenges related to data heterogeneity, scalability, and security threats. [11] The paper proposes a structured vulnerability and risk assessment for successful FL deployment, analyzing threats from different application areas and reviewing defensive algorithms and strategies. It categorizes applications into four main areas: space, air, ground, and underwater communications, comparing methodologies based on approach, base model, datasets, evaluation metrics, and achievements.

Almanifi et al. highlight the significance of FL in the era of big data and Artificial Intelligence due to its role in safeguarding data privacy and reducing the need to transfer and process massive amounts of data while retaining the benefits of ML. FL involves collaboratively training statistical models by exchanging learned parameter updates, contrasting with centralized training processes. However, widespread adoption of FL is impeded by communication and computation overhead, primarily due to the computational cost of training and large-sized parameter updates. This issue is exacerbated in applications involving the IoT due to the limited computational capabilities of edge and fog devices, bandwidth constraints, and internet connection capacities. [12] The paper aims to bridge this gap by systematically reviewing recent efforts to improve communication and computation efficiency in FL.

Mothukuri et al. focus on FL use cases where security and privacy are paramount, and understanding these risk factors is crucial for both implementers and researchers. [13] The study finds that while there are fewer privacy-specific threats, security threats such as communication bottlenecks, poisoning, and backdoor attacks are prevalent, with inference-based attacks posing the most significant privacy risk.

Blanco-Justicia et al.'s study examines security and privacy attacks on FL and surveys proposed solutions to mitigate each attack. It also discusses the challenge of achieving both security and privacy simultaneously. [14]

Truong et al. study challenges in managing data from various sources and complying with strict data protection regulations like the EU General Data Protection Regulation. Traditional centralized ML approaches pose privacy risks due to potential data leakage, misuse, and abuse. FL has emerged as a promising solution, enabling distributed collaborative learning without revealing original training data. However, FL systems still face privacy challenges, as model parameters exchanged among participants may contain sensitive information vulnerable to privacy attacks. Consequently, FL-based systems may not naturally comply with GDPR requirements. [15] The article surveys state-of-the-art privacy-preservation techniques in FL and their relation to GDPR requirements.

Enthoven and Al-Ars analyze existing vulnerabilities in FL and conduct a literature review of potential attack methods targeting FL privacy protection. These attack methods are categorized using a basic taxonomy. Additionally, the paper reviews recent defensive strategies and algorithms for FL aimed at mitigating these attacks, categorizing them by their underlying defense principles. [16]

Al-Quraan et al. provide a detailed examination of various applications of FL in wireless networks, highlighting their challenges and limitations. The efficacy of FL is explored in the context of emerging technologies beyond 5G and 6G communication systems. [17]

Ferrag et al. examine edge learning as an emerging approach to training models across distributed clients while ensuring data privacy, which can address challenges like resource management and behavior prediction in future network infrastructures like 6G. The article provides a comprehensive review of existing literature focusing on vulnerabilities and defenses related to edge learning for 6G-enabled IoT systems. The article further surveys research on attacks against machine learning, categorizing threat models into different categories, including backdoor attacks, adversarial examples, poisoning attacks, and inference attacks. [18]

2. CHALLENGES, SOLUTIONS, AND APPLICATIONS OF FEDERATED LEARNING

2.1 Privacy and Security Challenges in Federated Learning

FL is a machine learning framework in which numerous clients, such as mobile devices or entire organizations, collaboratively train a model coordinated by a central server, while maintaining decentralized training data. This approach emphasizes focused data collection and minimization, helping to alleviate many systemic privacy risks and costs associated with traditional, centralized machine learning. [19]

With many studies in the FL domain, many challenges have been exposed so far.

Inference attacks are a significant privacy threat in FL. In these attacks, adversaries aim to extract sensitive information from the shared model updates without direct access to the training data. The work by Melis et al. provides a detailed exploration of these vulnerabilities. Inference attacks exploit the fact that model updates, even if they do not include raw data, can still leak information about the underlying training data. These updates, shared between clients and the central server, reflect the characteristics of the local datasets used to train the model. By analyzing these updates, an adversary can infer sensitive information about the data. [20] There are two types of Inference Attacks: Membership Inference Attacks and Property Inference Attacks. Membership Inference Attacks determine whether a specific data point was part of a client's training set. By analyzing model updates, an attacker can distinguish between data points that influenced the model (i.e., were part of the training set) and those that did not. Property Inference Attacks aim to infer aggregate properties of the data used by clients, such as the presence of certain features or the distribution of data points with specific characteristics.

Model update poisoning is a significant security threat in FL, where malicious clients deliberately inject false or manipulated updates to corrupt the global model. [21] Mechanism of Model Update Poisoning:

- Participation as a Malicious Client: The adversary participates in the FL process as one of the clients. This could be an external attacker gaining access or an insider with malicious intent.
- Crafting Poisoned Updates: The adversary trains their local model on a poisoned dataset or directly manipulates the model updates before sending them to the server. These updates are designed to skew the global model towards undesirable behavior.
- Submission of Malicious Updates: The poisoned updates are transferred to the server during the model aggregation phase. Assuming all updates are benign, the server aggregates these malicious updates with the genuine ones from other clients.
- Impact on Global Model: The inclusion of poisoned updates affects the global model's parameters, leading to degraded performance, specific targeted misclassifications, or the embedding of backdoors that the adversary can exploit later.

The researchers explored different attack scenarios, including Targeted Attacks and Untargeted Attacks. Targeted Attacks: Where the adversary aims to misclassify specific inputs (e.g., causing a security system to misclassify intrusions as normal behavior). Untargeted Attacks: Where the goal is to degrade the overall performance of the global model, causing it to perform poorly on general tasks.

Zhao et al. address one of the key challenges in FL: dealing with non-independent and identically distributed data. [22] Traditional machine learning assumes that data is IID, meaning each data point is drawn from the same distribution and is independent of other data points. However, in practical FL applications, data on each client can vary significantly due to differing user behaviors, environments, and other factors, leading to non-IID data. This non-IID nature poses significant challenges for model convergence and performance.

Almanifi et al. give an extended review of the challenges to communication overhead in FL. [12] This overhead arises primarily from the need for frequent and large data exchanges between clients and the central server. Key Points of Communication Overhead:

- Large Data Transfers: Each client in FL must send model updates (weights and gradients) to the server frequently. The size of these updates can be substantial, especially for complex models, leading to significant data transfer volumes.
- Frequent Communication Rounds: FL involves multiple rounds of communication for iterative model updates, which can overwhelm network resources, particularly in environments with limited bandwidth.
- Bandwidth Constraints: Many clients operate over bandwidth-limited networks making it challenging to handle frequent and large data transmissions without experiencing delays or connectivity issues.

Shokri et al. express the aggregation privacy challenge that pertains to the process of aggregating model updates from various client devices or servers while ensuring the privacy of individual contributions. When multiple clients participate in FL by training models on their local data and sending updates to a central server for aggregation, there is a risk of leaking sensitive information from these updates. One of the main risks with aggregation in FL is the potential for inference attacks, where an adversary may infer sensitive information about individual training data from the aggregated model updates. Even if individual updates are anonymized or encrypted, patterns in the aggregated updates could reveal information about the underlying data distribution or individual contributions. [23]

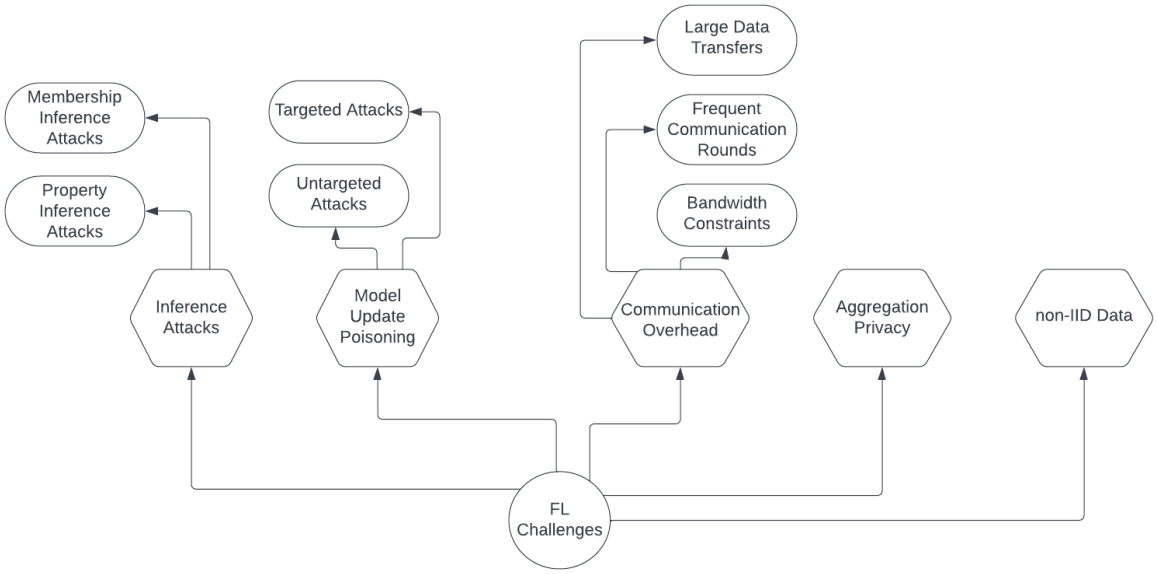

Figure 2.1 : FL Challenges.

2.2 Privacy and Security Solutions in Federated Learning

In Section 2.1, several privacy and security challenges of FL have been explained. In this section, different FL solutions that address these challenges will be explained.

Differential privacy is a mathematical technique for privacy-preserving data analysis of FL. It offers a strict mathematical framework for quantifying the privacy guarantees provided by data analysis algorithms. In FL, differential privacy is critical to protecting sensitive information about individual contributions.

At its core, differential privacy aims to provide plausible deniability for individual data points by assuring that the existence or non-existence of any single data point does not substantially affect the outcome of the assessment. This is achieved by inserting carefully calculated noise into the data or its analysis results in a way that obscures individual contributions while nevertheless enabling valid findings to be drawn from the aggregated data. [24]

In their study, Karakoc et al. present a protocol that is designed to anonymize local model updates transmitted to the server, thereby safeguarding the privacy of individual

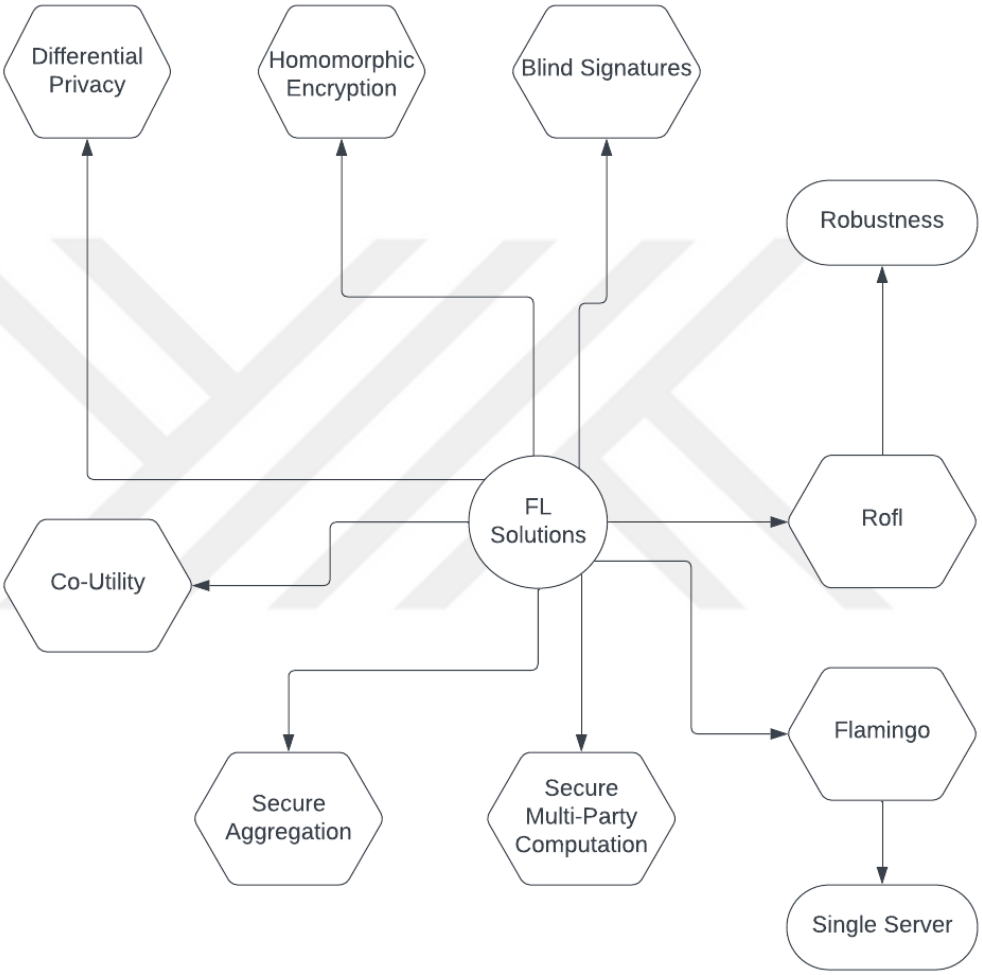

Figure 2.2 : FL Solutions.

contributors. This approach enables the server to scrutinize incoming data for potential poison or backdoor attacks. However, a notable limitation of this method is the absence of data source authentication, which exposes vulnerabilities exploitable by malicious clients situated between legitimate clients and the server. To mitigate this risk, they integrate cryptographic signatures for source authentication, ensuring the integrity of data transmission without compromising anonymity. Leveraging blind signatures, they obscure client identities, thereby preventing the server from discerning individual contributors. They incorporate partially blind signatures to embed common information, such as the current FL round number, into the signatures. The protocol entails the generation of a public-private key pair by each client, followed by a partially blind signature exchange with the server to obtain a blindly signed public key for each FL round. Subsequently, clients engage in local training, sign their model updates, and transmit them, along with signatures and relevant data, to designated forwarders. Upon reception, the server verifies signatures and aggregates updates. Within their trust model, both the server and clients are regarded as potential adversaries, necessitating stringent security measures. The protocol effectively addresses security requirements, including safeguarding individual data owners' privacy, detecting attacks on model training, and preventing unauthorized alterations to updates. Moreover, it ensures that each client submits only one update per FL round. While introducing computational and communication overhead, the protocol's performance remains manageable for both server and clients. [25]

The paper by Domingo-Ferrer et al. offers a comprehensive examination of FL with a particular focus on addressing critical concerns related to security and privacy. Central to the paper is the concept of co-utility, which advocates for mutual benefit and cooperation between the server orchestrating FL and participating client devices. This principle guides the development of a novel FL framework, encompassing secure protocols for model aggregation, privacy-preserving techniques for client data, and trust mechanisms to foster collaboration while mitigating adversarial behavior. [26]

The paper proposes innovative algorithms and methodologies tailored to the co-utility framework, aiming to optimize both model performance and data privacy. Empirical evaluations and simulations are conducted to validate the efficacy and performance of

the proposed solution, benchmarking it against existing FL approaches and assessing key metrics such as model accuracy and convergence speed. Throughout the paper, rigorous consideration is given to security and privacy implications, with analyses of potential vulnerabilities and countermeasures to safeguard against security threats.

The study by Lycklama et al. delves into a comprehensive examination of the robustness aspects within secure FL frameworks. The focus of this paper is to investigate and enhance the robustness of FL systems against various threats and attacks that may compromise the security and integrity of the learning process. To achieve this objective, an analysis of the vulnerabilities and challenges faced by secure FL systems has been conducted. These vulnerabilities may include potential attack vectors such as model poisoning attacks, data manipulation by malicious clients, or privacy breaches during model aggregation. Understanding these threats is crucial for devising effective defense mechanisms to mitigate them. The paper likely proposes novel algorithms, protocols, or techniques aimed at bolstering the security and resilience of FL frameworks. [27]

One of the primary contributions of the paper is likely to provide insights into the mechanisms for securing federated learning systems while ensuring their robustness against adversarial attacks. This may involve exploring cryptographic techniques, secure aggregation protocols, or differential privacy mechanisms to protect sensitive data and models during the FL process. Additionally, the paper may discuss the importance of trust establishment mechanisms between the server and participating clients to mitigate the risk of adversarial behavior. The paper serves as a significant contribution to the ongoing research efforts aimed at enhancing the robustness and security of federated learning systems, offering valuable insights and solutions to address the evolving challenges in this domain.

Chowdhury et al. present a comprehensive investigation into the integrity assurance mechanisms within FL frameworks. The focus of this paper is to explore and enhance mechanisms for ensuring the integrity of FL systems, which is essential for maintaining the reliability and trustworthiness of the learning process. To achieve this objective, the vulnerabilities and challenges associated with maintaining integrity in FL systems

have been examined. These vulnerabilities may include potential threats such as data tampering, model poisoning attacks, or malicious manipulation of model updates during the aggregation process. Understanding these threats is critical for devising effective integrity assurance mechanisms to detect and mitigate them. The paper likely proposes novel algorithms, protocols, or techniques aimed at bolstering the integrity of FL frameworks. [28]

One of the primary contributions of the paper is likely to provide insights into the mechanisms for securing federated learning systems while ensuring their integrity against adversarial attacks. This may involve the development of cryptographic techniques, secure aggregation protocols, or anomaly detection mechanisms to detect and prevent integrity violations. Additionally, the paper may discuss the importance of establishing trust between the server and participating clients to ensure the reliability and authenticity of model updates.

Gehlhar et al. provide an in-depth examination of a novel framework designed to facilitate secure and privacy-preserving FL through the use of SMPC techniques. The primary focus of this paper is to investigate and address the challenges associated with privacy and robustness in FL frameworks, particularly in the context of secure computation. [29]

The paper likely begins by discussing the motivations and objectives behind the development of the Safefl framework, highlighting the increasing demand for privacy-preserving solutions in FL scenarios and the limitations of existing approaches in achieving robustness and scalability. A thorough analysis of the vulnerabilities and threats faced by FL systems has been conducted, including privacy breaches, data leakage, and adversarial attacks, emphasizing the need for robust and privacy-preserving mechanisms to mitigate these risks.

To address these challenges, the Safefl framework is proposed, which leverages MPC techniques to enable secure and privacy-preserving computations across distributed parties. MPC ensures confidentiality and integrity during the FL process by allowing multiple parties to jointly compute a function over their private inputs without revealing sensitive information to each other. The paper presents the technical details of the

Safefl framework, including its architecture, protocols and algorithms, and discusses how MPC techniques are integrated into the FL workflow to enhance privacy and robustness.

The study by Ma et al. presents an in-depth exploration of a novel methodology called Flamingo. This methodology is designed to facilitate secure aggregation in federated learning (FL) scenarios with a single server across multiple rounds, thereby ensuring privacy and integrity throughout the learning process. The primary focus of the study is to investigate, identify and solve the issues related to secure aggregation in FL frameworks, particularly in scenarios involving multiple rounds of communication with a single server. [30]

The paper begins with a discussion of the motivations and objectives behind the development of the Flamingo methodology, highlighting the increasing demand for privacy-preserving solutions in FL scenarios and the limitations of existing approaches in achieving secure and efficient aggregation. A thorough analysis of the vulnerabilities and threats faced by FL systems has been conducted, including privacy breaches, data leakage, and adversarial attacks, emphasizing the need for robust and scalable mechanisms to ensure privacy and integrity in multi-round FL settings.

To address these challenges, the proposed Flamingo methodology leverages cryptographic techniques and secure aggregation protocols to enable privacy-preserving and integrity-assured computations across distributed parties. The methodology likely incorporates advanced cryptographic primitives such as homomorphic encryption, secret sharing, and secure multiparty computation to facilitate secure aggregation while preserving the confidentiality of individual data contributions. The paper likely presents the technical details of the Flamingo methodology, including its architecture, protocols, and algorithms, and discusses how it can be applied to FL scenarios involving multiple rounds of communication with a single server.

The work by Park and Lim explores a methodological approach to FL that prioritizes data privacy through the application of homomorphic encryption techniques. The focal point of the work is to investigate and propose mechanisms for enhancing privacy in FL scenarios, particularly through the utilization of HE.

To address the vulnerabilities and privacy risks, the paper proposes the use of homomorphic encryption, a cryptographic technique that enables computation on encrypted data without the need for decryption. By applying HE to FL, the study aims to facilitate secure and privacy-preserving computations during model aggregation and updates, thereby safeguarding the confidentiality of individual data while still allowing for collaborative learning across distributed devices. [31]

2.3 Applications of Federated Learning

In Sections 2.1 and 2.2, several privacy and security challenges and solutions of FL have been explained respectively. In this section, the application areas of FL will be explained.

FL has been successfully applied to various wireless communication scenarios, including vehicular communications, localization, and semantic communications. Here, detailed examples of these applications are provided. [32]

Vehicular communications aim to enhance safety, efficiency, and environmental sustainability in daily vehicular operations, contributing to the development of intelligent driving systems. Vehicles must communicate not only with infrastructure but also with nearby vehicles to exchange critical information such as safety messages. This requires ultra-reliable low-latency communication to ensure reliability and minimal latency.

In a given environment, the radio features of a mobile device can uniquely determine its location, enabling localization based on these features. However, the relationship between a mobile device's location and its radio features is often complex. Deep learning models can map radio features to specific locations, but training these models requires extensive data on radio features and corresponding locations, which can be difficult to collect.

One approach is to gather data on radio features and mobile locations from all devices in an area and use it to train the deep learning model. However, this method raises privacy issues and incurs substantial communication overhead. Federated learning offers a solution known as federated localization. In FEDLOC, each mobile device acts as a local client, collecting data on radio features and locations, updating the model locally, and sending the updated parameters to a central server. The base station or fusion center, acting as the central server, aggregates these local updates to form a global model.

Semantic communication represents a significant evolution from conventional communication by interpreting information at the semantic level. Instead of focusing on accurately recovering data, semantic communications are application-oriented, jointly designing communication and applications with an emphasis on the meaning of source messages. Only useful, relevant, and important information is transmitted, based on the application's requirements.

A deep learning-enabled semantic communication system has been developed for text transmission, where the meaning of text is extracted and compressed at the transmitter and then recovered at the receiver. DeepSC is particularly effective in low signal-to-noise ratio environments. Implementing DeepSC in a distributed manner for IoT networks involves the cloud/edge platform handling model training and updating, while IoT devices focus on data collection and transmission based on the trained model.

FL has significant applications in the IoT, enhancing privacy and reducing data transfer costs by keeping data local to devices. One prominent application is in smart homes, where FL enables devices to collaboratively learn and improve models for energy management, security, and user preferences without sharing raw data. This approach addresses privacy concerns and reduces network bandwidth usage, crucial in environments with limited connectivity and diverse device capabilities. [33]

In industrial IoT, FL is applied for predictive maintenance and anomaly detection. By aggregating insights from various machinery and equipment without centralizing sensitive operational data, FL helps in early fault detection and maintenance scheduling, thereby increasing efficiency and reducing downtime. The hierarchical aggregation framework, which involves local, edge, and cloud-based processing, optimizes communication and computational resources, making FL feasible in resource-constrained industrial IoT environments.

FL has found significant applications in healthcare, providing a solution to the critical issues of data privacy and security while enabling the development of robust machine learning models. By allowing multiple healthcare institutions to collaboratively train

models without sharing sensitive patient data, FL helps in overcoming the constraints posed by regulations like HIPAA in the US and GDPR in Europe.

One notable application of FL in healthcare is in medical imaging, where it aids in enhancing the accuracy of diagnostic tools. For example, FL is used to train models on datasets from multiple hospitals for tasks such as tumor detection and classification in MRI scans and X-rays. This approach helps in pooling diverse data, which improves the model's generalizability and performance across different populations and imaging devices. Studies have demonstrated that FL can maintain high levels of accuracy in these tasks while ensuring that patient data remains decentralized and secure. [34]

Another significant application of FL in healthcare is in predictive analytics and personalized medicine. FL allows for the aggregation of patient data from various sources to develop predictive models for disease outbreaks, patient readmission rates, and personalized treatment plans. For instance, models trained using FL can predict the likelihood of a patient developing complications after surgery by analyzing data from multiple institutions, thus providing more reliable predictions without compromising patient privacy. Additionally, FL has been utilized in genomic research to develop models that can predict genetic predispositions to certain diseases, leveraging data from various research centers without the need for data centralization. [35]

Overall, the implementation of federated learning in healthcare not only addresses the critical concerns of data privacy and security but also enhances the quality of patient care by enabling the development of more accurate and robust machine-learning models through collaborative efforts.

FL has significant applications in the financial sector, primarily driven by the need for privacy-preserving techniques in handling sensitive financial data. Traditional machine learning models require centralized data collection, which poses severe privacy risks and compliance challenges for financial institutions. Federated Learning addresses these issues by enabling the collaborative training of models across multiple decentralized devices or servers without the need to share raw data, thus preserving privacy and complying with regulatory requirements.

One notable application of FL in finance is in credit risk assessment. Financial institutions can use FL to build robust models for credit scoring by leveraging data from various banks and credit agencies while keeping the data on each institution's premises. This collaborative approach enhances the model's accuracy by incorporating a broader range of data without violating privacy norms. Recent research has shown that employing secure aggregation techniques within FL can significantly mitigate communication overhead and enhance the efficiency of the training process, which is critical given the large datasets and complex models typically used in financial applications. [36]

Another critical application is in fraud detection, where FL allows for the integration of diverse datasets from multiple financial institutions to improve the detection of fraudulent activities. By combining insights from different sources, FL-based models can achieve higher detection rates and adapt more quickly to new fraud patterns. Additionally, innovative methods like gradient sparsification and secure aggregation have been developed to reduce the communication costs associated with FL, making it more feasible for real-world deployment.

These advancements highlight FL's potential to revolutionize financial services by providing secure, efficient, and collaborative machine-learning solutions that respect user privacy and regulatory constraints.

3. FEDERATED LEARNING SURVEYS

3.1 Roles and Benefits of Surveys in the Federated Learning Domain

Surveys play a critical role in the FL domain, serving multiple purposes that drive research, development, and implementation forward. Here, the key roles and benefits of surveys in this field are discussed.

3.1.1 Roles of surveys in FL

Surveys compile and synthesize existing research, providing a holistic view of many highly referenced studies in FL. This includes summarizing key concepts, methodologies, challenges, and solutions, which helps new researchers quickly get up to speed with the field.

By critically analyzing existing literature, surveys help identify gaps and unresolved challenges in the FL domain. This can guide future research directions by highlighting areas that require more attention, such as privacy-preserving techniques, model aggregation methods, or dealing with non-IID data distributions. For instance, surveys often point out the need for more robust solutions to inference attacks and model update poisoning.

Surveys often provide comparative analyses of different techniques and approaches within FL. This includes evaluating the performance, scalability, and security of various algorithms and frameworks, which helps practitioners choose the most appropriate methods for their specific applications. Such comparisons are invaluable for both academic research and practical deployments.

By consolidating diverse research findings, surveys contribute to the establishment of best practices and standard protocols in FL. This standardization is crucial for ensuring compatibility and interoperability between different FL systems and applications.

3.1.2 Benefits of surveys in FL

Surveys help accelerate research and development by providing a ready reference of existing knowledge and advancements. Researchers can build on summarized findings without needing to read through an extensive array of individual papers. This efficiency fosters faster innovation and progress in the field.

For practitioners and decision-makers, surveys offer a consolidated source of information that supports informed decision-making. Whether implementing FL in a corporate environment or developing new algorithms, having access to comprehensive survey data helps stakeholders make better strategic and technical decisions.

Surveys serve as excellent educational resources for students and educators in the field of FL. They provide a structured overview of the domain, making it easier to teach and learn about complex topics. Educational institutions often use surveys as part of their curriculum to introduce advanced concepts in FL.

By highlighting current trends and active research areas, surveys facilitate collaboration among researchers and institutions. Knowing what others are working on and the challenges they face encourages collaborative efforts to address these issues, leading to more robust and holistic solutions.

3.2 Survey Analysis in Federated Learning

The survey by Al-Huthaifi et al. presents an in-depth investigation into the application of federated learning techniques within the context of smart cities, focusing particularly on the critical aspects of privacy and security. The emergence of smart cities, characterized by extensive data collection through various sensor networks and IoT devices, presents both opportunities and challenges in leveraging machine learning algorithms for optimizing urban operations. The paper delves into the unique requirements and challenges posed by smart city environments concerning privacy and security. Given the sensitive nature of the data collected in urban settings, including personally identifiable information and sensitive infrastructure details, ensuring robust privacy protections is paramount to foster trust among citizens and stakeholders. Moreover, the distributed nature of data sources in smart cities introduces additional security vulnerabilities, necessitating robust mechanisms for authentication, access control, and secure communication protocols to safeguard against malicious attacks and data breaches.

In order to address these concerns, the paper presents a comprehensive review of the existing literature and research efforts focused on improving the privacy and security of federated learning systems deployed in smart city applications. This survey encompasses a diverse range of methodologies, techniques, and best practices proposed by researchers and practitioners to address the complex challenges inherent to data privacy and security in federated learning environments. These include cryptographic techniques such as HE, SMPC, and differential privacy mechanisms, which are designed to preserve data confidentiality while enabling collaborative model training.

Moreover, the paper investigates the function of federated learning in enabling compliance with regulatory frameworks and standards pertaining to the protection and privacy of data, including GDPR and HIPAA. By harmonising federated learning practices with regulatory requirements, stakeholders in smart cities can guarantee legal compliance while cultivating responsible data stewardship practices.

Furthermore, the paper explores the implications of federated learning on the broader socio-economic landscape of smart cities, including its potential to empower local communities, promote data sovereignty, and foster innovation while addressing concerns related to digital divide and disparities in access to technology.

All in all, the study makes a significant contribution to the existing body of literature by providing a systematic overview of the current best practices, difficulties, and trends in leveraging federated learning to strengthen privacy and security in smart city environments. By clarifying the sophisticated interconnections between technological innovations, regulatory obligations, and ethical concerns, the study seeks to provide policymakers, urban planners, and researchers with the insights necessary to make well-informed decisions regarding the implementation of federated learning solutions in smart city applications. This, in turn, will add to the collective objective of developing more inclusive, resilient, and sustainable urban ecosystems. [8]

The survey by Wen et al. presents a detailed examination of federated learning, elucidating its challenges and applications across various domains. It delineates the key components of federated learning systems, including client devices, central aggregators, and communication protocols, thereby establishing a conceptual framework for the subsequent discussion.

Building upon this foundational understanding, the paper explores the myriad challenges inherent in federated learning, spanning technical, operational, and regulatory domains. Technical challenges encompass issues related to model convergence, communication efficiency, and heterogeneity of data distributions across client devices, necessitating the development of novel optimization algorithms and federated learning frameworks tailored to address these constraints. Operational challenges revolve around the practical deployment and management of federated learning systems at scale, including resource allocation, fault tolerance, and coordination among participating entities, highlighting the importance of robust infrastructure and governance mechanisms to ensure seamless operation.

Furthermore, the paper delves into the regulatory and ethical considerations surrounding federated learning, emphasizing the need for compliance with data

protection regulations, such as GDPR and HIPAA, to safeguard user privacy and mitigate legal risks. It discusses the implications of federated learning on data ownership, consent management, and accountability frameworks, underscoring the importance of transparent and accountable practices to engender trust among stakeholders.

In addition to addressing challenges, the paper surveys the diverse applications of federated learning across domains such as healthcare, finance, telecommunications, and smart cities. In healthcare, federated learning enables collaborative model training on sensitive patient data distributed across healthcare institutions, facilitating the development of personalized treatment models while preserving patient privacy. In finance, federated learning is utilized for fraud detection, credit scoring, and risk assessment, leveraging insights gleaned from diverse data sources while complying with regulatory requirements.

Moreover, the paper explores the role of federated learning in addressing challenges related to data scarcity, privacy concerns, and regulatory constraints in emerging technologies such as edge computing, IoT, and blockchain. By enabling collaborative model training at the network edge, federated learning empowers edge devices to leverage local data while minimizing communication overhead and preserving user privacy. Similarly, in IoT ecosystems, federated learning facilitates joint learning from heterogeneous sensor data streams distributed across edge devices, enabling real-time insights and decision-making while mitigating privacy risks associated with centralized data aggregation.

Overall, the paper contributes to the scholarly discourse on federated learning by providing a comprehensive synthesis of its challenges and applications across diverse domains. By elucidating the technical, operational, and regulatory dimensions of federated learning, the paper aims to inform researchers, practitioners, and policymakers about the opportunities and considerations associated with deploying federated learning solutions in real-world scenarios. Through interdisciplinary collaboration and concerted efforts, federated learning has the potential to drive

innovation, foster collaboration, and advance the collective goal of building more secure, privacy-preserving, and equitable machine learning systems. [9]

Soykan and et al.'s study provides a comprehensive examination of privacy-enhancing technologies in a collaborative machine learning context. Collaborative machine learning, which involves multiple parties jointly training a machine learning model on their respective datasets, presents inherent privacy risks due to the potential exposure of sensitive information. PETs aim to mitigate these risks by incorporating privacy-preserving mechanisms into the collaborative learning process. The paper begins with an overview of collaborative machine learning paradigms, including federated learning, multi-party computation, and secure aggregation, highlighting their respective strengths and limitations in addressing privacy concerns.

Building upon this foundation, the paper systematically surveys existing PETs and their applicability to collaborative machine learning scenarios. This survey encompasses a wide range of techniques, including differential privacy, homomorphic encryption, secure enclaves, and data anonymization, each offering distinct approaches to preserving privacy while enabling collaborative model training. Differential privacy, for instance, quantifies the privacy guarantees of a machine learning algorithm by measuring the impact of individual data points on the model's output, thereby ensuring that no single data point can unduly influence the learning process.

Homomorphic encryption enables computations to be performed directly on encrypted data without the need for decryption, thereby protecting sensitive information throughout the computation pipeline. Secure enclaves, such as Intel SGX and ARM TrustZone, provide hardware-based isolation mechanisms to safeguard computations and data within a trusted execution environment, mitigating the risk of unauthorized access or tampering. Data anonymization techniques, including k-anonymity and differential privacy, transform raw data into a form that preserves privacy while retaining utility for machine learning tasks.

Furthermore, the paper provides practical guidelines for selecting and deploying PETs in collaborative machine-learning settings, taking into account factors such as computational overhead, communication overhead, and the level of privacy

protection afforded by each technique. It discusses considerations around data preprocessing, model selection, and evaluation metrics to guide researchers and practitioners in designing privacy-preserving machine learning pipelines. Additionally, the paper explores the ethical and regulatory implications of deploying PETs in real-world applications, emphasizing the importance of transparency, accountability, and informed consent in handling sensitive data.

Through a synthesis of theoretical principles, practical considerations, and ethical guidelines, the paper aims to empower stakeholders to navigate the complex landscape of privacy-enhancing technologies and make informed decisions regarding their adoption and implementation. By fostering collaboration and knowledge sharing among researchers, practitioners, and policymakers, PETs have the potential to drive innovation, promote responsible data stewardship, and uphold individuals' rights to privacy in an increasingly data-driven society. [10]

The survey by Sirohi et al. presents an exhaustive exploration of the application of FL in secure communication systems enabled by 6G technology. With the advent of 6G networks promising unprecedented data rates, ultra-low latency, and massive connectivity, guaranteeing the security and privacy of communication becomes paramount.

The survey delves into the unique requirements and challenges posed by secure communication systems in the 6G era, including stringent security and privacy requirements, dynamic network conditions, and heterogeneity of edge devices. By utilizing federated learning, secure communication systems can benefit from the accumulated intelligence of participating nodes while mitigating the risks associated with centralised data aggregation and processing. The paper systematically surveys the existing literature and research efforts focused on applying federated learning techniques to enhance the security, privacy, and efficiency of communication systems in the 6G landscape.

This survey encompasses a broad spectrum of methodologies, including differential privacy, secure aggregation, and federated learning-based intrusion detection, tailored to address the multifaceted challenges encountered in 6G-enabled communication

systems. Differential privacy techniques enable the quantification and enforcement of privacy guarantees during the model training process, ensuring that individual data contributions remain confidential. Secure aggregation mechanisms protect the integrity and confidentiality of aggregated model updates transmitted from edge devices to the central server, mitigating the risk of eavesdropping or tampering.

Furthermore, the paper explores the role of federated learning in facilitating secure and privacy-preserving communication in various application scenarios, including IoT, smart cities, vehicular networks, and healthcare. In IoT ecosystems, federated learning enables collaborative model training on sensor data streams generated by a myriad of connected devices, facilitating real-time insights and decision-making while preserving data sovereignty and user privacy. Similarly, in smart cities, federated learning empowers local authorities to leverage distributed data sources for urban planning, traffic management, and environmental monitoring, without compromising individual privacy rights.

Moreover, the paper discusses the implications of federated learning on the broader socio-economic landscape of 6G-enabled communication systems, including its potential to foster innovation, stimulate economic growth, and address societal challenges. By enabling secure and privacy-preserving communication, federated learning contributes to building trust among stakeholders, promoting responsible data stewardship practices, and advancing the collective goal of creating a more secure and inclusive digital society.

In conclusion, the paper contributes to the scientific discourse by providing a thorough synthesis of the existing approaches, challenges and future directions in the use of federated learning for secure communication systems in the 6G era. By elucidating the complex interplay between technological advancements, regulatory requirements, and societal implications, the paper aims to inform researchers, practitioners, and policymakers about the opportunities and considerations associated with deploying federated learning solutions in real-world communication environments. Through interdisciplinary collaboration and concerted efforts, federated learning has the

potential to revolutionize secure communication systems, ushering in a new era of connectivity, resilience, and trust in the 6G landscape. [11]

The survey by Almanifi et al. provides a comprehensive examination of the efficiency aspects of FL, focusing particularly on communication and computation efficiency. The paper delves into the critical importance of communication efficiency in FL systems, given the bandwidth constraints, latency considerations, and energy consumption associated with transmitting model updates between participating nodes and the server. The paper systematically surveys existing literature and research efforts aimed at optimizing communication efficiency in FL, encompassing techniques such as model compression, quantization, and adaptive communication strategies.

Model compression techniques aim to reduce the size of model updates transmitted between edge devices and the central server, thereby minimizing communication overhead without sacrificing model performance. Quantization methods quantize model parameters to lower precision representations, enabling more compact representation and efficient transmission over communication channels. Adaptive communication strategies dynamically adjust communication schedules, bandwidth allocation, and aggregation frequencies based on network conditions and edge device capabilities to optimize communication efficiency.

Furthermore, the paper explores the challenges and opportunities associated with computation efficiency in FL systems, considering factors such as computational complexity, resource constraints, and heterogeneity of edge devices. It surveys existing approaches for enhancing computation efficiency in FL, including federated optimization algorithms, model parallelism, and edge computing techniques.

Federated optimization algorithms adapt traditional optimization techniques to the decentralized and asynchronous nature of FL, enabling efficient model training across distributed edge devices while mitigating computational overhead. Model parallelism strategies partition model parameters and computations across multiple edge devices, allowing for parallel execution and efficient utilization of computational resources. Edge computing technologies leverage the computational capabilities of edge devices

to offload computation-intensive tasks from the central server, reducing latency and improving scalability in FL systems.

Moreover, the paper discusses the interplay between communication and computation efficiency in FL, emphasizing the need for holistic optimization strategies that balance trade-offs between communication and computation overhead. It explores the implications of efficiency considerations on model performance, convergence speed, and scalability in FL systems, highlighting the importance of interdisciplinary research and collaboration to address these challenges. [12]

The survey by Mothukuri et al. offers a comprehensive examination of the security and privacy aspects within the realm of FL. The paper delves into the critical importance of security and privacy in FL systems, given the sensitivity of data handled by decentralized edge devices and the potential risks of malicious attacks or unauthorized access. The paper systematically surveys existing literature and research endeavors aimed at enhancing the security and privacy of FL, encompassing a wide spectrum of techniques, methodologies, and best practices.

Security considerations in FL systems include threat models, attack vectors, and defense mechanisms tailored to mitigate risks such as data poisoning, model inversion attacks, and membership inference attacks. Privacy-preserving techniques, including differential privacy, federated encryption, and secure aggregation, aim to safeguard sensitive information during model training and inference, ensuring that individual data contributions remain confidential.

Furthermore, the paper explores the implications of security and privacy concerns on FL system design, architecture, and operation, highlighting the need for interdisciplinary research and collaboration to address these challenges effectively. It discusses the role of regulatory frameworks, industry standards, and best practices in guiding the development and deployment of secure and privacy-preserving FL systems.

In addition to technical solutions, the paper examines the socio-economic and ethical implications of security and privacy concerns in FL, considering factors such as user trust, fairness, and accountability. It underscores the importance of

transparency, accountability, and informed consent in handling sensitive data and making algorithmic decisions that impact individuals' privacy rights.

Also, the paper discusses emerging trends and future directions in the security and privacy of FL, including advancements in cryptographic techniques, federated learning frameworks, and regulatory compliance mechanisms. It emphasizes the need for continuous research and innovation to stay abreast of evolving threats and vulnerabilities in FL systems, fostering a culture of security awareness and privacy by design in machine learning technologies. [13]

The survey by Blanco-Justicia et al. offers a comprehensive investigation into the strategies, challenges, and future directions concerning security and privacy within FL systems. The survey underscores the critical significance of security and privacy in FL systems, given the sensitive nature of data processed by decentralized edge devices and the potential threats posed by malicious actors or unauthorized access. The paper systematically surveys existing literature and research endeavors aimed at enhancing the security and privacy of FL, encompassing a broad spectrum of techniques, methodologies, and best practices.

Security considerations in FL systems encompass threat models, attack vectors, and defense mechanisms tailored to mitigate risks such as data poisoning, model inversion attacks, and membership inference attacks. Privacy-preserving techniques, including differential privacy, federated encryption, and secure aggregation, aim to safeguard sensitive information during model training and inference, ensuring that individual data contributions remain confidential.

Furthermore, the paper delves into the implications of security and privacy concerns on FL system design, architecture, and operation, emphasizing the need for interdisciplinary research and collaboration to effectively address these challenges. It discusses the role of regulatory frameworks, industry standards, and best practices in guiding the development and deployment of secure and privacy-preserving FL systems.

In addition to technical solutions, the paper examines the socio-economic and ethical implications of security and privacy concerns in FL, considering factors such as user trust, fairness, and accountability. It underscores the importance of

transparency, accountability, and informed consent in handling sensitive data and making algorithmic decisions that impact individuals' privacy rights.

What's more, the paper discusses emerging trends and future directions in achieving security and privacy in FL systems, including advancements in cryptographic techniques, federated learning frameworks, and regulatory compliance mechanisms. It emphasizes the need for continuous research and innovation to stay ahead of evolving threats and vulnerabilities in FL systems, fostering a culture of security awareness and privacy by design in machine learning technologies. [14]

The survey by Truong et al. provides a comprehensive examination of privacy preservation within the context of federated learning (FL), with a specific focus on compliance with GDPR. It highlights the critical importance of privacy preservation in FL systems, given the sensitivity of data processed by decentralized edge devices and the legal obligations imposed by GDPR regarding the collection, processing, and storage of personal data. The paper systematically surveys existing literature and research endeavors aimed at enhancing privacy preservation in FL systems from the GDPR perspective, encompassing a wide spectrum of techniques, methodologies, and best practices.

Privacy preservation techniques in FL systems encompass differential privacy, federated encryption, data anonymization, and consent management mechanisms, designed to safeguard individual privacy rights while enabling collaborative model training. Differential privacy quantifies and limits the privacy risks associated with individual data contributions, ensuring that no single data point can unduly influence the learning process.

Federated encryption techniques protect sensitive information during model training and inference, ensuring that data remains encrypted throughout the computation pipeline. Data anonymization methods transform raw data into a form that preserves privacy while retaining utility for machine learning tasks, thereby minimizing the risk of re-identification or unauthorized access.

Furthermore, the paper delves into the implications of GDPR compliance on FL system design, architecture, and operation, emphasizing the need for robust governance

frameworks, transparency, and accountability mechanisms. It discusses the role of data protection impact assessments, data minimization principles, and privacy-enhancing technologies in achieving GDPR compliance while leveraging the benefits of federated learning.

In addition to technical solutions, the paper examines the socio-economic and ethical implications of privacy preservation in FL, considering factors such as user trust, fairness, and algorithmic accountability. It underscores the importance of informed consent, purpose limitation, and data subject rights in ensuring GDPR compliance and protecting individual privacy rights.

In conclusion, the paper discusses emerging trends and future directions in privacy preservation in FL systems, including advancements in privacy-preserving machine learning techniques, federated learning frameworks, and regulatory compliance mechanisms. It emphasizes the need for interdisciplinary research and collaboration to address the complex interplay between technological advancements, legal requirements, and ethical considerations in the context of FL and GDPR compliance. [15]

The survey by Enthoven and Al-Ars provides a detailed examination of privacy attacks and defensive strategies within the realm of federated deep learning. It underscores the critical significance of privacy preservation in FDL systems, given the sensitive nature of data processed by decentralized edge devices and the potential risks posed by adversarial attacks or unauthorized access. The paper systematically surveys existing literature and research endeavors aimed at understanding and mitigating privacy attacks in FDL, encompassing a broad spectrum of techniques, methodologies, and best practices.

Privacy attacks in FDL systems encompass various threat models, attack vectors, and exploitation techniques aimed at compromising the confidentiality, integrity, and availability of sensitive information. Adversarial attacks such as model inversion, membership inference, and poisoning attacks exploit vulnerabilities in FDL systems to infer sensitive information about individual data contributions, manipulate model parameters, or compromise model performance.

Furthermore, the paper delves into the implications of privacy attacks on FDL system design, architecture, and operation, emphasizing the need for robust defense mechanisms, anomaly detection techniques, and adversarial resilience strategies. It discusses the role of differential privacy, federated encryption, and secure aggregation in mitigating privacy attacks and preserving confidentiality while enabling collaborative model training.

In addition to defensive strategies, the paper examines the socio-economic and ethical implications of privacy attacks in FDL, considering factors such as user trust, fairness, and algorithmic accountability. It underscores the importance of transparency, accountability, and informed consent in handling sensitive data and making algorithmic decisions that impact individuals' privacy rights.

Also, the paper discusses emerging trends and future directions in privacy attacks and defensive strategies in FDL, including advancements in adversarial machine learning techniques, federated learning frameworks, and regulatory compliance mechanisms. It emphasizes the need for continuous research and innovation to stay ahead of evolving threats and vulnerabilities in FDL systems, fostering a culture of security awareness and privacy by design in deep learning technologies. [16]

The survey by Al-Quraan et al. provides a comprehensive examination of the integration of edge-native intelligence and federated learning (FL) techniques within the context of 6G communications. With the emergence of 6G networks promising unprecedented data rates, ultra-low latency, and massive connectivity, there arises a need to leverage edge computing capabilities and FL methodologies to meet the demands of diverse applications. The paper highlights the critical importance of edge-native intelligence and federated learning in enabling distributed intelligence and collaborative model training across edge devices while preserving data privacy and minimizing communication overhead. The paper systematically surveys existing literature and research endeavors aimed at exploring the trends and challenges in leveraging edge-native intelligence and FL for 6G communications, encompassing a wide spectrum of techniques, methodologies, and best practices.

Edge-native intelligence techniques enable computation and decision-making to be performed closer to the data source, reducing latency, bandwidth requirements, and reliance on centralized infrastructure. Federated learning methodologies facilitate collaborative model training across distributed edge devices while preserving data privacy through techniques such as federated encryption, secure aggregation, and differential privacy.

Furthermore, the paper delves into the implications of integrating edge-native intelligence and FL on the design, architecture, and operation of 6G communication systems, emphasizing the need for scalable, interoperable, and resilient solutions. It discusses the role of edge computing frameworks, federated learning platforms, and communication protocols in enabling efficient and secure communication, computation, and learning at the network edge.

In addition to technical considerations, the paper examines the socio-economic and ethical implications of edge-native intelligence and FL in 6G communications, considering factors such as data sovereignty, fairness, and algorithmic accountability. It underscores the importance of transparency, accountability, and informed consent in handling sensitive data and making algorithmic decisions that impact individuals' privacy rights.

Moreover, the paper discusses emerging trends and future directions in edge-native intelligence and FL for 6G communications, including advancements in edge computing architectures, federated learning algorithms, and regulatory compliance mechanisms. It emphasizes the need for interdisciplinary research and collaboration to address the complex challenges and opportunities in this domain, fostering innovation, resilience, and inclusivity in future communication networks. [17]

The survey by Ferrag et al. offers a thorough examination of the integration of edge learning techniques within the context of 6G-enabled IoT systems. The paper pinpoints the critical importance of edge learning in enabling distributed intelligence, real-time decision-making, and adaptive resource management at the network edge, thereby enhancing the efficiency, reliability, and scalability of IoT deployments. The paper systematically surveys existing literature and research endeavors aimed at exploring

the vulnerabilities, datasets, and defense mechanisms in edge learning for 6G-enabled IoT, encompassing a wide spectrum of techniques, methodologies, and best practices.

Vulnerabilities in edge learning for 6G-enabled IoT systems encompass various threat vectors, attack surfaces, and exploitation techniques aimed at compromising the confidentiality, integrity, and availability of IoT devices, data, and services. These vulnerabilities include data poisoning attacks, model inversion attacks, adversarial examples, and privacy breaches, which exploit weaknesses in edge learning algorithms, data collection mechanisms, and communication protocols.

Furthermore, the paper scrutinizes the implications of vulnerabilities on the design, architecture, and operation of 6G-enabled IoT systems, emphasizing the need for robust defense mechanisms, anomaly detection techniques, and adversarial resilience strategies. It discusses the role of secure edge computing frameworks, robust machine learning models, and encrypted communication protocols in mitigating vulnerabilities and preserving the security and privacy of IoT deployments.

In addition to defensive strategies, the paper analyses the availability of datasets for training and evaluating edge learning models in the context of 6G-enabled IoT systems, considering factors such as data diversity, size, and quality. It underscores the importance of benchmark datasets, simulation environments, and real-world deployments in facilitating reproducible research and benchmarking performance across different edge learning algorithms and applications.

Finally, the paper discusses emerging trends and future directions in edge learning for 6G-enabled IoT, including advancements in edge computing architectures, federated learning frameworks, and regulatory compliance mechanisms. It emphasizes the need for interdisciplinary research and collaboration to address the complex challenges and opportunities in this domain, fostering innovation, resilience, and inclusivity in future IoT deployments. [18]

4. METRICS AND STATE OF THE ART FEDERATED LEARNING STUDIES

Many studies have appeared in FL field so far. These studies come with different views such as application domains, privacy enhancements, security enhancements, and regulatory compliance. Diversity of research areas is a niche thing however, the increasing number of well-known studies results in complexity that will lead to incomprehensibility. To overcome this issue and bind several studies, thirteen metrics that fall into five main categories are offered.

In this section, eleven state-of-the-art studies have been included. Most of the examined studies here are published within the last two years. In Table 4.2 these state-of-the-art solutions are analyzed and combined using the metrics explained in Table 4.1.

Table 4.1 : Metrics for Comparison of Security & Privacy Solutions

Metric	Acronym	Definition
Single server	SS	The solution does not require more than one server in the protocol
Multiple server	MS	The solution requires more than one server where at least one of the servers is semi-honest (i.e., all the servers cannot collude)
Semi-honest server	SHS	The solution is secure against the server that follows the protocol steps for the computation of the aggregation result
Malicious server	MCS	The solution is secure against the server that may not follow the protocol steps for the computation of the aggregation result
Aggregation integrity	AI	The solution ensures that the aggregation result is not altered by the server after the computation of the aggregation result.
Input privacy	IP	The solution does not leak information about the input of clients to the server(s)
Client drop-outs	CDO	The solution is robust against the client drop-outs (i.e., the secure aggregation can be computed even if some of the clients drops-out)
Input integrity	II	The solution ensures that the inputs of the clients are in a pre-defined input range
Semi-honest clients	SHC	The solution is secure against the client who follows the protocol steps after starting the protocol by providing their inputs
Malicious clients	MCC	The solution is secure against the clients who may not follow the protocol steps after starting the protocol by providing their inputs
Star topology	ST	The clients only need to communicate with the server
P2P topology	P2PT	The clients need to communicate with each other in addition to the server
Scalable	S	The solution computation and communication are linear both in the number of clients

Table 4.2 : Comparison of Security & Privacy Solutions

Solution	SS	MS	SHS	MCS	AI	IP	CDO	II	SHC	MCC	ST	P2PT	S
ELSA [37]		✓	✓		✓	✓	✓	✓		✓		✓	✓
Co-utility [26]	✓		✓			✓	✓	✓		✓		✓	✓
Rofl [27]	✓		✓			✓	✓	✓		✓	✓		✓
Byzantine-Resilient [38]	✓		✓			✓		✓	✓			✓	✓
EIFFeL [28]	✓			✓	✓	✓	✓	✓		✓	✓		✓
SAFEFL [29]		✓		✓	✓	✓	✓	✓		✓	✓		✓
DP-BREM [39]	✓		✓			✓		✓	✓			✓	✓
Flamingo [30]	✓			✓		✓	✓	✓		✓	✓		✓
zPROBE [40]	✓		✓			✓		✓		✓	✓		✓
Prio [41]		✓	✓			✓		✓		✓	✓		✓
Karakoç et al. [25]	✓		✓			✓	✓	✓		✓		✓	✓

5. CONCLUSIONS

Federated Learning presents a transformative approach to machine learning by enabling decentralized data processing, which addresses critical privacy and security concerns inherent in traditional centralized models. This thesis explored various facets of FL, particularly focusing on the challenges and solutions related to privacy and security, as well as its diverse applications across different sectors.

5.1 Privacy and Security Challenges in Federated Learning

FL inherently mitigates several privacy risks by keeping data local to clients. However, it introduces new challenges, particularly related to inference attacks and model update poisoning. Inference attacks exploit model updates to extract sensitive information, while model update poisoning involves malicious clients injecting false updates to corrupt the global model. These challenges necessitate robust solutions to ensure the integrity and privacy of the FL process.

Non-IID data and communication overheads further complicate FL implementation. Non-IID data, where data distributions vary across clients, can hinder model convergence and performance. Additionally, frequent and substantial data exchanges between clients and servers result in significant communication overheads, which can strain network resources.

5.2 Privacy and Security Solutions in Federated Learning

Several strategies have been developed to address these privacy and security challenges. Differential privacy introduces noise to data updates, ensuring that individual contributions remain confidential. Protocols that incorporate cryptographic signatures and SMPC techniques further enhance the security of model updates and ensure data integrity. Co-utility frameworks, which promote mutual benefit

between servers and clients, and robust aggregation methods also play vital roles in safeguarding FL systems.

Innovative methodologies such as Flamingo and SafeFL leverage advanced cryptographic techniques to provide secure aggregation and enhance privacy preservation. These solutions collectively improve the robustness, efficiency, and security of FL frameworks, enabling their application in real-world scenarios.

5.3 Applications of Federated Learning

FL has been applied successfully in various domains, demonstrating its versatility and effectiveness. In wireless communication, FL enhances vehicular communication, localization, and semantic communication by enabling collaborative model training without data centralization. In the IoT sector, FL improves privacy and reduces data transfer costs, with significant applications in smart homes and industrial IoT.

Healthcare is another critical area where FL has made substantial impacts. By allowing institutions to collaboratively train models on medical imaging and predictive analytics without sharing patient data, FL addresses stringent privacy regulations while improving model accuracy and generalizability. Studies have shown that FL can maintain high diagnostic accuracy and support personalized medicine.

In the financial sector, FL addresses privacy and regulatory challenges by enabling collaborative credit risk assessment and fraud detection. By leveraging data from multiple institutions without centralizing it, FL-based models achieve higher accuracy and adaptability, enhancing the detection of fraudulent activities and improving credit scoring models.

5.4 Surveys of Federated Learning

Surveys play indispensable roles and offer numerous benefits within the FL domain. They serve as comprehensive repositories of existing research, providing newcomers with a foundational understanding while guiding experienced researchers toward unexplored frontiers. By scrutinizing and synthesizing a plethora of literature, surveys pinpoint gaps and challenges, steering the trajectory of future investigations towards

areas demanding attention, such as privacy preservation and robust model aggregation methods. Furthermore, they facilitate informed decision-making for practitioners and policymakers, accelerate research and development efforts, and foster collaborative endeavors within the FL community. As FL continues to evolve, surveys will remain instrumental in shaping its trajectory, fostering innovation, and facilitating the dissemination of knowledge to propel the field toward greater efficiency, security, and scalability.

Each survey dives into various aspects of privacy, security, efficiency, and application domains, offering valuable insights into the challenges, methodologies, and future directions within their respective domains.

5.5 Comparison Metrics and State of the Art Federated Learning Studies

Several studies have been performed since the initial definition and research on FL. Highly referenced studies are declared as state-of-the-art solutions in FL field. The state-of-the-art solutions bring diverse metrics that can be used as a baseline. As mentioned earlier too many state-of-the-art solutions have appeared and this brings many complexities which makes it difficult to digest FL approaches. That's why, a comparison of security and privacy solutions using baseline metrics is a must.

In this thesis, thirteen metrics that fall into six categories are used for the comparison of well known studies in FL field. This approach will help grasp FL solutions and their pros and cons.

REFERENCES

[1] **Solove, D.** (2010). *Understanding Privacy*.

[2] **Dwork, C.** (2008). Differential privacy: A survey of results, *International Conference on Theory and Applications of Models of Computation*, Springer, pp.1–19.

[3] **McMahan, H.B.** *et al.* (2017). *Federated Learning: Collaborative Machine Learning without Centralized Training Data*, `https://ai.googleblog.com/2017/04/federated-learning-collaborative.html`.

[4] **Konečný, J.** *et al.* (2021). Federated Learning: Strategies for Improving Communication Efficiency, *arXiv preprint arXiv:1610.05492*.

[5] **Yang, Q.** *et al.* (2019). Federated Machine Learning: Concept and Applications, *ACM Transactions on Intelligent Systems and Technology (TIST)*, *10*(2), 1–19.

[6] **Bonawitz, K.** *et al.* (2019). Towards Federated Learning at Scale: System Design, *arXiv preprint arXiv:1902.01046*.

[7] **Li, T.** *et al.* (2020). Federated Learning: Challenges, Methods, and Future Directions, *IEEE Signal Processing Magazine*, *37*(3), 50–60.

[8] **Al-Huthaifi, R.**, **Li, T.**, **Huang, W.**, **Gu, J. and Li, C.** (2023). Federated learning in smart cities: Privacy and security survey, *Inf. Sci.*, *632*, 833–857, `https://doi.org/10.1016/j.ins.2023.03.033`.

[9] **Wen, J.**, **Zhang, Z.**, **Lan, Y.**, **Cui, Z.**, **Cai, J. and Zhang, W.** (2023). A survey on federated learning: challenges and applications, *Int. J. Mach. Learn. Cybern.*, *14*(2), 513–535, `https://doi.org/10.1007/s13042-022-01647-y`.

[10] **Soykan, E.U.**, **Karaçay, L.**, **Karakoç, F. and Tomur, E.** (2022). A Survey and Guideline on Privacy Enhancing Technologies for Collaborative Machine Learning, *IEEE Access*, *10*, 97495–97519, `https://doi.org/10.1109/ACCESS.2022.3204037`.

[11] **Sirohi, D.**, **Kumar, N.**, **Rana, P.S.**, **Tanwar, S.**, **Iqbal, R. and Hijji, M.** (2023). Federated learning for 6G-enabled secure communication systems: a comprehensive survey, *Artif. Intell. Rev.*, *56*(10), 11297–11389, `https://doi.org/10.1007/s10462-023-10417-3`.

[12] **Almanifi, O.R.A.**, **Chow, C.**, **Tham, M.**, **Chuah, J.H. and Kanesan, J.** (2023). Communication and computation efficiency in Federated Learning: A survey, *Internet Things*, *22*, 100742, https://doi.org/10.1016/j.iot.2023.100742.

[13] **Mothukuri, V.**, **Parizi, R.M.**, **Pouriyeh, S.**, **Huang, Y.**, **Dehghantanha, A. and Srivastava, G.** (2021). A survey on security and privacy of federated learning, *Future Gener. Comput. Syst.*, *115*, 619–640, https://doi.org/10.1016/j.future.2020.10.007.

[14] **Blanco-Justicia, A.**, **Domingo-Ferrer, J.**, **Martínez, S.**, **Sánchez, D.**, **Flanagan, A. and Tan, K.E.** (2021). Achieving security and privacy in federated learning systems: Survey, research challenges and future directions, *Eng. Appl. Artif. Intell.*, *106*, 104468, https://doi.org/10.1016/j.engappai.2021.104468.

[15] **Truong, N.B.**, **Sun, K.**, **Wang, S.**, **Guitton, F. and Guo, Y.** (2021). Privacy preservation in federated learning: An insightful survey from the GDPR perspective, *Comput. Secur.*, *110*, 102402, https://doi.org/10.1016/j.cose.2021.102402.

[16] **Enthoven, D. and Al-Ars, Z.** (2020). An Overview of Federated Deep Learning Privacy Attacks and Defensive Strategies, *CoRR*, *abs/2004.04676*, https://arxiv.org/abs/2004.04676, 2004.04676.

[17] **Al-Quraan, M.**, **Mohjazi, L.S.**, **Bariah, L.**, **Centeno, A.**, **Zoha, A.**, **Arshad, K.**, **Assaleh, K.**, **Muhaidat, S.**, **Debbah, M. and Imran, M.A.** (2023). Edge-Native Intelligence for 6G Communications Driven by Federated Learning: A Survey of Trends and Challenges, *IEEE Trans. Emerg. Top. Comput. Intell.*, *7*(3), 957–979, https://doi.org/10.1109/TETCI.2023.3251404.

[18] **Ferrag, M.A.**, **Friha, O.**, **Kantarci, B.**, **Tihanyi, N.**, **Cordeiro, L.C.**, **Debbah, M.**, **Hamouda, D.**, **Al-Hawawreh, M. and Choo, K.R.** (2023). Edge Learning for 6G-Enabled Internet of Things: A Comprehensive Survey of Vulnerabilities, Datasets, and Defenses, *IEEE Commun. Surv. Tutorials*, *25*(4), 2654–2713, https://doi.org/10.1109/COMST.2023.3317242.

[19] **Kairouz, P.**, **McMahan, H.B.**, **Avent, B.**, **Bellet, A.**, **Bennis, M.**, **Bhagoji, A.N.**, **Bonawitz, K.A.**, **Charles, Z.**, **Cormode, G.**, **Cummings, R.**, **D'Oliveira, R.G.L.**, **Eichner, H.**, **Rouayheb, S.E.**, **Evans, D.**, **Gardner, J.**, **Garrett, Z.**, **Gascón, A.**, **Ghazi, B.**, **Gibbons, P.B.**, **Gruteser, M.**, **Harchaoui, Z.**, **He, C.**, **He, L.**, **Huo, Z.**, **Hutchinson, B.**, **Hsu, J.**, **Jaggi, M.**, **Javidi, T.**, **Joshi, G.**, **Khodak, M.**, **Konečný, J.**, **Korolova, A.**, **Koushanfar, F.**, **Koyejo, S.**, **Lepoint, T.**, **Liu, Y.**, **Mittal, P.**, **Mohri, M.**, **Nock, R.**, **Özgür, A.**, **Pagh, R.**, **Qi, H.**, **Ramage, D.**, **Raskar, R.**, **Raykova, M.**, **Song, D.**, **Song, W.**, **Stich, S.U.**, **Sun, Z.**, **Suresh, A.T.**, **Tramèr, F.**, **Vepakomma, P.**, **Wang, J.**, **Xiong, L.**, **Xu, Z.**, **Yang, Q.**, **Yu, F.X.**, **Yu, H. and Zhao,**

S. (2021). Advances and Open Problems in Federated Learning, *Found. Trends Mach. Learn.*, *14*(1-2), 1–210, https://doi.org/10.1561/2200000083.

[20] **Melis, L.**, **Song, C.**, **Cristofaro, E.D. and Shmatikov, V.** (2019). Exploiting Unintended Feature Leakage in Collaborative Learning, *2019 IEEE Symposium on Security and Privacy, SP 2019, San Francisco, CA, USA, May 19-23, 2019*, IEEE, pp.691–706, https://doi.org/10.1109/SP.2019.00029.

[21] **Bhagoji, A.N.**, **Chakraborty, S.**, **Mittal, P. and Calo, S.B.** (2019). Analyzing Federated Learning through an Adversarial Lens, ***K. Chaudhuri and R. Salakhutdinov**, editors, Proceedings of the 36th International Conference on Machine Learning, ICML 2019, 9-15 June 2019, Long Beach, California, USA*, volume 97 of *Proceedings of Machine Learning Research*, PMLR, pp.634–643, http://proceedings.mlr.press/v97/bhagoji19a.html.

[22] **Zhao, Y.**, **Li, M.**, **Lai, L.**, **Suda, N.**, **Civin, D. and Chandra, V.** (2018). Federated Learning with Non-IID Data, *CoRR*, *abs/1806.00582*, http://arxiv.org/abs/1806.00582, 1806.00582.

[23] **Shokri, R.**, **Stronati, M.**, **Song, C. and Shmatikov, V.** (2017). Privacy-preserving aggregation of federated learning models, *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, ACM, pp.1175–1191.

[24] **McMahan, H.B.**, **Ramage, D.**, **Talwar, K. and Zhang, L.** (2018). Learning Differentially Private Recurrent Language Models, *6th International Conference on Learning Representations, ICLR 2018, Vancouver, BC, Canada, April 30 - May 3, 2018, Conference Track Proceedings*, OpenReview.net, https://openreview.net/forum?id=BJ0hF1Z0b.

[25] **Karakoç, F.**, **Karaçay, L.**, **Cnudde, P.Ç.D.**, **Gülen, U.**, **Fuladi, R. and Soykan, E.U.** (2023). A security-friendly privacy-preserving solution for federated learning, *Comput. Commun.*, *207*, 27–35, https://doi.org/10.1016/j.comcom.2023.05.004.

[26] **Domingo-Ferrer, J.**, **Blanco-Justicia, A.**, **Manjón, J.A. and Sánchez, D.** (2022). Secure and Privacy-Preserving Federated Learning via Co-Utility, *IEEE Internet Things J.*, *9*(5), 3988–4000, https://doi.org/10.1109/JIOT.2021.3102155.

[27] **Lycklama, H.**, **Burkhalter, L.**, **Viand, A.**, **Küchler, N. and Hithnawi, A.** (2023). RoFL: Robustness of Secure Federated Learning, *44th IEEE Symposium on Security and Privacy, SP 2023, San Francisco, CA, USA, May 21-25, 2023*, IEEE, pp.453–476, https://doi.org/10.1109/SP46215.2023.10179400.

[28] **Chowdhury, A.R.**, **Guo, C.**, **Jha, S. and van der Maaten, L.** (2022). EIFFeL: Ensuring Integrity for Federated Learning, ***H. Yin, A. Stavrou, C. Cremers and E. Shi**, editors, Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security, CCS 2022, Los Angeles, CA, USA, November 7-11, 2022*, ACM, pp.2535–2549, https://doi.org/10.1145/3548606.3560611.

[29] **Gehlhar, T.**, **Marx, F.**, **Schneider, T.**, **Suresh, A.**, **Wehrle, T. and Yalame, H.** (2023). SAFEFL: MPC-friendly Framework for Private and Robust Federated Learning, *IACR Cryptol. ePrint Arch.*, 555, https://eprint.iacr.org/2023/555.

[30] **Ma, Y.**, **Woods, J.**, **Angel, S.**, **Polychroniadou, A. and Rabin, T.** (2023). *Flamingo: Multi-Round Single-Server Secure Aggregation with Applications to Private Federated Learning*, Cryptology ePrint Archive, Paper 2023/486, https://eprint.iacr.org/2023/486, https://eprint.iacr.org/2023/486.

[31] **Park, J. and Lim, H.** (2022). Privacy-Preserving Federated Learning Using Homomorphic Encryption, *Applied Sciences*, *12*(2), https://www.mdpi.com/2076-3417/12/2/734.

[32] **Qin, Z.**, **Li, G.Y. and Ye, H.** (2021). Federated Learning and Wireless Communications, *IEEE Wirel. Commun.*, *28*(5), 134–140, https://doi.org/10.1109/MWC.011.2000501.

[33] **Zhang, T.**, **Gao, L.**, **He, C.**, **Zhang, M.**, **Krishnamachari, B. and Avestimehr, A.S.** (2022). Federated Learning for the Internet of Things: Applications, Challenges, and Opportunities, *IEEE Internet Things Mag.*, *5*(1), 24–29, https://doi.org/10.1109/IOTM.004.2100182.

[34] **Prayitno**, **Shyu, C.R.**, **Putra, K.T.**, **Chen, H.C.**, **Tsai, Y.Y.**, **Hossain, K.S.M.T.**, **Jiang, W. and Shae, Z.Y.** (2021). A Systematic Review of Federated Learning in the Healthcare Area: From the Perspective of Data Properties and Applications, *Applied Sciences*, *11*(23), https://www.mdpi.com/2076-3417/11/23/11191.

[35] **Dhade, P. and Shirke, P.** (2023). Federated Learning for Healthcare: A Comprehensive Review, *Engineering Proceedings*, *59*(1), https://www.mdpi.com/2673-4591/59/1/230.

[36] **Liu, T.**, **Wang, Z.**, **He, H.**, **Shi, W.**, **Lin, L.**, **An, R. and Li, C.** (2023). Efficient and Secure Federated Learning for Financial Applications, *Applied Sciences*, *13*(10), https://www.mdpi.com/2076-3417/13/10/5877.

[37] **Rathee, M.**, **Shen, C.**, **Wagh, S. and Popa, R.A.** (2022). ELSA: Secure Aggregation for Federated Learning with Malicious Actors, *IACR Cryptol. ePrint Arch.*, 1695, https://eprint.iacr.org/2022/1695.

[38] **Masuda, H.**, **Kita, K.**, **Koizumi, Y.**, **Takemasa, J. and Hasegawa, T.** (2023). Byzantine-Resilient Secure Federated Learning on Low-Bandwidth Networks, *IEEE Access*, *11*, 51754–51766.

[39] **Gu, X.**, **Li, M. and Xiong, L.** (2023). DP-BREM: Differentially-Private and Byzantine-Robust Federated Learning with Client Momentum, *CoRR*, *abs/2306.12608*, https://doi.org/10.48550/arXiv.2306.12608, 2306.12608.

[40] **Ghodsi, Z.**, **Javaheripi, M.**, **Sheybani, N.**, **Zhang, X.**, **Huang, K. and Koushanfar, F.** (2023). zPROBE: Zero Peek Robustness Checks for Federated Learning, *IEEE/CVF International Conference on Computer Vision, ICCV 2023, Paris, France, October 1-6, 2023*, IEEE, pp.4837–4847, https://doi.org/10.1109/ICCV51070.2023.00448.

[41] **Corrigan-Gibbs, H. and Boneh, D.** (2017). Prio: Private, Robust, and Scalable Computation of Aggregate Statistics, *CoRR*, *abs/1703.06255*, http://arxiv.org/abs/1703.06255, 1703.06255.

CURRICULUM VITAE

Şükrü ERDAL

EDUCATION:

- **B.Sc.:** 2007, Marmara University, Faculty of Technical Education
- **M.Sc.:** 2024, Istanbul Technical University, Cybersecurity Engineering and Cryptography, Department of Applied Informatics

PROFESSIONAL EXPERIENCE AND REWARDS:

- **KoçSistem:** November 2020 - , Senior Cyber Security Specialist
- **IBM:** November 2017 - October 2020 , Network Support Specialist
- **ProXGate:** February 2016 - October 2017, Founder & Network and Security Consultant
- **Telcoset:** March 2015 - February 2016, Network and Security Specialist
- **Burgan Bank:** June 2014 - February 2015, Network and Communication Manager
- **BDH - Bilişim Destek Hizmetleri:** April 2011 - June 2014, Network Security Specialist
- **Turcom Teknoloji:** September 2008 - April 2011, Network Engineer

PUBLICATIONS, PRESENTATIONS AND PATENTS ON THE THESIS:

- **Şükrü Erdal**, Ferhat Karakoç, Enver Özdemir, "A survey on privacy and security aspects and solutions for federated learning in mobile communication networks", 11th International Congress on Fundamental and Applied Sciences, 9-11 July 2024, Istanbul, Turkiye