The Transactions of the Korea Information Processing Society
TKIPS Vol.13, No.2, pp.76~90

eISSN: 3022-7011
https://doi.org/10.3745/TKIPS.2024.13.2.76

Systematic Research on Privacy-Preserving Distributed Machine Learning

Min Seob Lee† · Young Ah Shin†† · Ji Young Chun†††

ABSTRACT

Although artificial intelligence (AI) can be utilized in various domains such as smart city, healthcare, it is limited due to concerns about the exposure of personal and sensitive information. In response, the concept of distributed machine learning has emerged, wherein learning occurs locally before training a global model, mitigating the concentration of data on a central server. However, overall learning phase in a collaborative way among multiple participants poses threats to data privacy. In this paper, we systematically analyzes recent trends in privacy protection within the realm of distributed machine learning, considering factors such as the presence of a central server, distribution environment of the training datasets, and performance variations among participants. In particular, we focus on key distributed machine learning techniques, including horizontal federated learning, vertical federated learning, and swarm learning. We examine privacy protection mechanisms within these techniques and explores potential directions for future research.

Keywords : Distributed Machine Learning, Privacy-Preserving Technologies, Federated Learning, Swarm Learning

프라이버시를 보호하는 분산 기계 학습 연구 동향

이 민 섭† · 신 영 아†† · 천 지 영†††

요 약

인공지능 기술은 스마트 시티, 자율 주행, 의료 분야 등 다양한 분야에서 활용 가능성을 높이 평가받고 있으나, 정보주체의 개인정보 및 민감정보의 노출 문제로 모델 활용이 제한되고 있다. 이에 따라 데이터를 중앙 서버에 모아서 학습하지 않고, 보유 데이터셋을 바탕으로 일차적으로 학습을 진행한 후 글로벌 모델을 최종적으로 학습하는 분산 기계 학습의 개념이 등장하였다. 그러나, 분산 기계 학습은 여전히 협력하여 학습을 진행하는 과정에서 데이터 프라이버시 위협이 발생한다. 본 연구는 분산 기계 학습 연구 분야에서 프라이버시를 보호하기 위한 연구를 서버의 존재 유무, 학습 데이터셋의 분포 환경, 참여자의 성능 차이 등 현재까지 제안된 분류 기준들을 바탕으로 유기적으로 분석하여 최신 연구 동향을 파악한다. 특히, 대표적인 분산 기계 학습 기법인 수평적 연합학습, 수직적 연합학습, 스웜 학습에 집중하여 활용된 프라이버시 보호 기법을 살펴본 후 향후 진행되어야 할 연구 방향을 모색한다.

키워드 : 분산 기계 학습, 프라이버시 보호, 연합 학습, 스웜 학습

1. 서 론

데이터 기술의 급격한 발전과 데이터 중심의 사회로의 전환은 정보보호의 중요성을 더 증가시켰다. 특히, 대규모 데이터를 활용하는 기계 학습과 데이터 과학 분야에서는 개인 정보 보호와 데이터 보안 문제가 핵심적인 관심사로 떠오르고 있다. 이러한 문제를 해결하기 위한 한 방안으로, 분산 기계 학습이 주목받고 있다. 이러한 방식은 중앙 서버에 데이터를 집중시키지 않고, 개별 장치에서 데이터를 처리하여 학습하는 것을 의미한다. 이는 데이터의 소유자가 자신의 데이터를 직접 통제하며 프라이버시를 유지할 수 있도록 한다[73,74].

본 논문에서는 분산 기계 학습 환경에서 프라이버시를 보호하는 기술에 집중하여, 이 분야의 최신 연구 동향과 기술의 기법들을 살펴본다. 먼저, 데이터를 로컬로 학습을 수행하면서도 발생할 수 있는 위험을 보호하는 프라이버시 보호 기술들이 있다. 간략하게, 중요한 정보를 여러 조각으로 나누어 안전하게 공유하는 Secret Sharing 기법, 참여자들이 안전하게 암호화 키를 공유하고 합의하는 방식인 Key Agreement 기법, 데이터를 암호화한 상태에서도 연산을 수행가능하게 하는 Homomorphic Encryption 기법, 특정 함수만을 계산할 수 있는 특별한 키를 생성하여 유연성을 제공하는 Functional

※ 이 논문은 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 2021R1F1A1063992).
† 준 회 원 : 고려대학교 정보보호대학원 Ph.D., 석·박사통합과정
†† 비 회 원 : 고려대학교 정보보호대학원 Ph.D., 석·박사통합과정
††† 정 회 원 : 서울사이버대학교 빅데이터·정보보호학과 조교수
Manuscript Received : December 5, 2023
Accepted : December 21, 2023
* Corresponding Author : Ji Young Chun(jychun@iscu.ac.kr)

Encryption 기법, 데이터 집합에 노이즈를 추가하여 프라이버시를 보호하는 Differential Privacy 기법 등이 존재한다. 이러한, 프라이버시 보호 기술을 분산 기계 학습에 적용한 연구들을 살펴본다.

두 번째로, 학습을 시작하기 전 데이터를 전처리하는 과정에 대한 연구를 분석한다. 이 과정에서는 Data Cleansing, Privacy-Preserving Record Linkage, Private Set Interaction 과 같은 기법들이 중요한 역할을 한다. Data Cleansing은 불완전하거나 오류가 있는 데이터를 정정하거나 제거하는 과정으로, 데이터의 품질을 향상시킨다. Privacy-Preserving Record Linkage는 서로 다른 데이터 소스에서 동일한 개체에 대한 기록을 식별하면서 개인정보를 보호하는 기법이다. 마지막으로, Private Set Interaction은 두 개 이상의 데이터셋 간의 공통 요소를 찾아내면서 각 데이터셋의 개인정보를 보호하는 방법이다. 이러한 기법들은 데이터셋을 정화하고, 통일된 형식에 맞게 정렬한 후 학습에 사용할 공통 데이터셋을 선별하는 중요한 작업으로, 최근 분산 기계 학습에 어떻게 적용되는지 동향을 분석하는 데 중요하다.

마지막으로, 분산 기계 학습은 다양한 접근 방식으로 분류될 수 있으며, 이는 각각의 특성과 사용되는 암호학적 프리미티브에 따라 다른 연구 동향을 보여준다. 첫 번째로, 분산 기계 학습은 서버 기반(Server-based)과 서버리스(Server-less) 접근 방식으로 나뉜다. 서버 기반 접근 방식은 중앙 서버가 데이터 처리와 학습 관리를 담당하는 반면, 서버 리스 접근 방식은 서버에만 의존하지 않고 클라우드 기반의 자원을 활용하여 데이터를 처리한다. 두 번째로, 분산 기계 학습은 데이터의 구조에 따라 수직적(Vertical) 또는 수평적(Horizontal) 분산 기계 학습 기법으로 분류된다. 수직적 분산 학습은 다른 특성을 가진 데이터셋을 결합하는 데 초점을 맞추고, 수평적 분산 학습은 동일한 특성을 가진 대규모 데이터셋을 처리하는 데 적합하다. 마지막으로, Cross-Device와 Cross-Silo 접근 방식은 분산 학습의 환경에 따라 나뉩니다. Cross-Device는 다수의 소형 장치가 참여하는 환경을, Cross-Silo는 제한된 수의 대형 조직이나 데이터 센터가 참여하는 환경을 의미한다. 이러한 각각의 접근 방식은 분산 기계 학습의 다양한 적용 사례와 연구 방향을 제시한다.

전체적인 연구의 목적은 분산 기계 학습 환경에서 프라이버시를 보호하는 기술의 최신 동향을 분석하고, 이러한 기술들이 어떻게 데이터 보호와 학습 효율성 사이의 균형을 달성할 수 있는지 확인한다. 이 연구는 기존 연구에서 다루지 않았던 새로운 관점으로 보며, 향후 연구 방향에 대한 중요한 통찰력을 제공할 것이다.

논문의 구성은 다음과 같다. 2장에서는 프라이버시 보호 기술에 대해 설명하고, 3장에서는 데이터 전처리 기법에 대해 다룬다. 4장에서는 분산 기계 학습 기법의 다양한 적용 사례를 분석한다. 마지막으로, 결론에서는 향후 연구 방향에 대해 논의한다.

2. 프라이버시 보호 기술

분산 기계 학습은 데이터를 중앙집중식 서버에 모아서 처리하지 않고, 각각의 기기에서 로컬로 학습을 수행하여, 개인정보를 보호하면서도 기계 학습 모델을 효과적으로 학습시킬 수 있는 기술이다. 이러한 분산 기계 학습 방식은 데이터의 소유자가 데이터를 제어하고, 데이터의 프라이버시를 보호할 수 있게 한다. 그러나, 분산 기계 학습 환경에서도 데이터의 프라이버시 보호는 여전히 중요한 과제로 여겨지고 있어, 이를 해결하기 위한 다양한 기술적 접근 방식이 제시되고 있다[75].

프라이버시 보호 기술은 분산 기계 학습 환경에서 데이터의 프라이버시를 보호하고, 민감 정보의 노출을 방지하기 위해 중요한 역할을 한다. 이러한 기술들은 데이터의 무결성과 기밀성을 보장하며, 학습과 추론 과정에서 발생할 수 있는 위험을 최소화한다. 따라서, 분산 기계 학습을 다룬 연구에서 사용되고 있는 Secret Sharing, Key Agreement, Homomorphic Encryption, Functional Encryption, Differential Privacy 기법에 대해 설명하고, 이러한 기술들이 어떠한 방식으로 활용되고 있는지를 다룬다.

2.1 Secret Sharing

Secret Sharing은 중요한 정보를 안전하게 공유하는데 사용되는 암호학 기법으로, Shamir와 Blakley가 처음으로 제안하였다[1]. 이 기법은 하나의 비밀 정보 S는 그림자라고 불리는 n개의 여러 조각으로 나누어진다. 각 조각만으로는 원래의 비밀 정보를 알아낼 수 없도록 설계되어 있고, 미리 정해진 k개($k \leq n$) 이상의 조각을 갖고 있을 경우에만 숨겨진 S를 알 수 있게 된다.

Shamir가 제안한 Secret Sharing 방식을 살펴보면, 먼저 비밀 정보를 분배하는 키 생성 기관(KGC, Key Generation Center)가 존재한다. 키 생성 기관은 먼저 파수가 (k-1)인 다항식을 랜덤하게 선택해서 아래의 식과 같이 만든다.

$$f(x) = a_0 + a_1 x + \cdots + a_{k-1} x^{k-1}$$

이 경우, 상수 값을 숨기고자 하는 비밀 정보는 $S = a_0 = f(0)$ 이고, 모든 계수인 $a_0, a_1, \cdots, a_{k-1}$ 은 소수 p개의 원소로 구성된 유한체(finite field) F_p상의 원소로 한다. $f(x)$를 사용해서 n개의 조각을 $f(i) \bmod p$ 를 통해 생성한다. (단, $0 \leq i < n$) 그 후 키 생성 기관은 각 조각을 n명의 사용자들에게 안전하게 분배한다.

비밀 정보를 다시 복구하기 위해서는 n개의 조각 중에서 임의의 k개 이상의 조각을 모아야 한다. 만약 임의의 k개의 조각을 $(s_{i_1}, s_{i_2}, \cdots, s_{i_k})$라 하면 아래의 식을 통해서 비밀 정보를 다시 복구할 수 있다.

$$S = f(0) = \sum_{i \in (i_1, i_2, \cdots, i_k)} s_i \beta i$$

$$= \sum_{i \in (i_1, i_2, \cdots, i_k)} s_i \left(\prod_{j \in (i_1, i_2, \cdots, i_k) - i} \frac{x_j}{x_j - x_i} \right) \mathrm{mod} p$$

(β_i는 Lagrange 계수)

위의 Shamir 기법과 같이, Secret Sharing은 데이터의 소유자와 기계 학습 모델의 참여자들 사이에서 민감한 정보를 안전하게 공유하고, 동시에 데이터의 프라이버시와 기밀성을 달성할 수 있다. 이러한 프라이버시 보호 기술은 분산 기계 학습 환경에서 중요하게 사용되며, 여러 참여자들이 각자의 데이터를 공유하지 않고 공동의 머신러닝 모델을 훈련할 수 있도록 하였다[47]. Xu 등의 [53]에서는 Secret Sharing 기법을 사용하여 사용자의 로컬 가중치의 기밀성을 보장하고, 서버의 집계 결과를 더 정확하게 검증할 수 있도록 하였다.

2.2 Key Agreement

모든 원고는 국문 또는 영문으로 작성하여야 한다. 국문 논문의 경우, 서지 정보(제목, 저자, 소속, 교신저자의 주소와 이메일), 표, 그림, 감사의 글, 참고문헌 등은 모두 영문으로 기술하여야 한다. 심사를 위한 초기 투고 원고에는 저자 정보를 포함시키지 말아야 한다. 하지만, 논문 수락 판정을 받은 후 제출하는 최종본에는 저자 정보를 포함시켜야 한다[4, 5].

Key Agreement 기법은 두 개 이상의 통신 참여자들이 네트워크를 통해 안전하게 암호화 키를 공유하고 합의하는 암호학적 기법으로 Diffie-Hellman이 처음 제안하였다[2]. 이 기법은 참여자들이 공개 채널을 통해 서로 안전하게 비밀 키를 생성하고 교환할 수 있도록 하여 이렇게 생성된 키는 후속 통신의 암호화와 복호화에 사용된다. 구체적으로는 ($KA.param, KA. \geq n, KA.agree$)라는 세 가지 알고리즘으로 구성된다.

$KA.param(k) \rightarrow pp$: 공개 매개변수라는 pp를 생성

$KA.gen(pp) \rightarrow (s_u^{SK}, s_u^{PK})$: 어떠한 사용자 u가 자신의 비밀키, 공개키 쌍을 생성. 여기서 s_u^{SK}는 비밀키(Private key), s_u^{PK}는 공개키(Public key)를 의미한다.

$KA.agree(s_u^{SK}, s_u^{PK}) \rightarrow s_u, v$: 사용자 u가 자신의 비밀키 s_u^{SK}와 다른 사용자 v의 공개키 s_u^{PK}를 결합하여, u와 v 사이의 비밀 공유키 su,v를 생성한다.

Diffie-Hellman의 키 동의 방식에서는 아래와 같은 알고리즘으로 과정이 진행된다.

$KA.param(k) \rightarrow (G', g, q, H)$: 소주 차수 q의 그룹 G', 생성자 g, 그리고 해시 함수 H를 뽑는다.

$KA.gen(G', g, q, H) \rightarrow (x, g^x)$: 랜덤 x로부터 비밀키 s_u^{SK}를 만들고, g^x를 공개키 s_u^{PK}로 사용한다.

$KA.gen(x_u, g^{x_u}) \rightarrow s_u, v$: 마지막 단계에서는 s_u, v를 출력하여, u 와 v 사이에 공유된 비밀키를 생성한다.

키 동의 프로토콜은 Diffie-Hellman이 제안한 이후, 해시 및 대칭 암호화 기반, PKC(Public-Key Cryptography)기반, ID 기반 등으로 연구되어 오고 있다. 특히 분산 기계 학습 환경에서의 참여자들이 서로 신뢰할 수 없거나 중앙 서버와 같은 인증 기관이 없는 경우에도 안전한 통신을 보장하기 위해 사용된다. 참여자들은 민감한 데이터를 안전하게 공유하고, 모델 학습과 추론 과정에서 발생할 수 있는 데이터 유출과 무단 접근을 방지할 수 있다[47, 49].

2.3 Homomorphic Encryption

Homomorphic Encryption (HE)은 암호화된 데이터에 대해 연산을 가능하게 하는 암호학적 기법으로, 이를 통해 데이터를 암호화한 상태에서 덧셈, 뺄셈, 곱셈 등의 연산을 수행할 수 있다. 일반적으로, HE는 가능한 연산의 수에 따라 부분 동형 암호(Partial Hommorphic Encryption), 유한 동형 암호(Somewhat Homomorphic Encryption), 완전 동형 암호(Fully Homomorphic Encryption)와 같이 세 가지로 분류된다. PHE는 덧셈과 곱셈 연산 중 한 가지만 가능하게 하고, FHE는 두 가지 연산을 모두 가능하게 한다. SWHE는 FHE처럼 두 연산을 모두 가능하게 하지만, 연산의 수가 제한적이다. 연합학습을 포함한 분산 기계 학습 환경을 다룬 연구에서 동형 암호가 모델 매개변수를 보호하여 사용자 데이터의 프라이버시를 높였다[39, 50, 51, 54, 55, 56, 57, 63, 66, 67].

분산 기계 학습에 동형 암호를 적용하기 위해서는 계산 비용이 중요하다. 따라서, 곱셈의 높은 계산 비용으로 인해 FHE보다 PHE가 더 많이 사용된다. 분산 기계 학습에 사용되는 PHE를 설명하기 위해 아래에 Paillier의 기법을 설명한다[3].

Pailier의 HE 기법은 아래의 식과 같이 5개의 알고리즘으로 구성되어 있다.

$HE = P(HE.Gen, HE.Enc, HE.Dec, HE.Add, HE.Mul)$

공개키 쌍 $(pk, sk) = HE.Gen(1^\lambda)$에 대해 $c = HE.Enc_{pk}(m)$이고,

$HE.Dec_{sk}(c) = HE.Dec_{sk}(HE.Enc_{pk}(m)) = m$이다.

위의 공개키 쌍 $(pk, sk) = HE.Gen(1^\lambda)$와 상수 k에 대해 $c_1 = HE.Enc_{pk}(m_1)$, $c_2 = HE.Enc_{pk}(m_2)$를 만족하면, 모든 메시지 $m_1, m_2 \in M$ 에 대해 암호문을 복호화하지 않고도 아래의 두 식을 만족한다.

$$HE.Add_{pk}(c_1, c_2) = HE.Enc_{pk}(m_1 + m_2)$$
$$HE.Mul_{pk}(k, c_2) = HE.Enc_{pk}(k \bullet m_2)$$

2.4 Functional Encryption

Functional Encryption (FE) 기법은 특정 함수만을 계산할 수 있는 특별한 키를 생성할 수 있게 한다. 이 기법을 통해, 키의 소유자는 암호화된 데이터에 대해 특정 함수를 계산할 수 있지만, 데이터 자체에 대한 접근은 제한된다. 간단히 말하

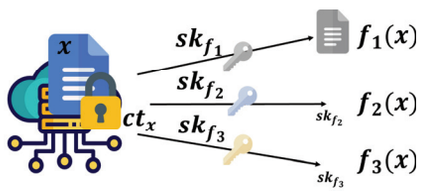

Fig. 1. Fuctional Encryption

면, 함수 키(functional key) sk_f와 x에 대한 암호문 ct_x가 주어지면, 복호화를 통해서 연산의 결과로 $f(x)$가 출력으로 나오고 x에 대한 어떠한 정보도 나오지 않는 것을 말한다.

이는 분산 기계 학습 환경의 참여자들이 민감한 데이터를 공유하지 않고도 특정 계산을 수행하거나 모델을 학습시킬 수 있게 해준다. Functional Encryption은 데이터의 프라이버시를 보호하면서도 필요한 계산을 수행할 수 있는 유연성을 제공하며, 이는 분산 학습 시나리오에서 중요한 역할을 수행한다[40, 49].

2.5 Differential Privacy

Differential Privacy (DP)는 개인 데이터의 프라이버시를 보호하기 위해 데이터 집합에 노이즈를 추가하는 동시에, 유의미한 통계적 정보의 쿼리를 공유할 수 있게 하는 기법이다. 기본적으로, 데이터 집합에 쿼리를 실행할 때, 개별 데이터 포인트의 존재 여부가 결과에 미치는 영향을 제한한다. 이를 수학적으로 표현하면, 두 데이터 집합 D_1과D_2가 하나의 요소만 다를 때, 위의 데이터로 임의의 알고리즘 A를 실행했을 때, $A(D_1)$와 $A(D_2)$의 결과 분포가 거의 유사해야 한다. 이는 아래의 식과 같이 표현된다.

$$\Pr[A(D_1)\in S] \le e^{\epsilon} \times \Pr[A(D_2)\in S]$$

여기서 S는 출력의 임의의 부분 집합이고, ϵ는 '프라이버시 손실'을 나타내는 작은 양의 값이다. 이 ϵ값은 결과의 차등성을 결정하며, 값이 작을수록 프라이버시 보호가 강화된다. DP는 이러한 방식으로 개인의 프라이버시를 보호하면서도 데이터 집합의 전체적인 통계적 특성을 유지할 수 있게 한다. 따라서, 분산 기계 학습 특히 연합학습에 DP를 적용하여 가중치 유출 공격과 같은 위험을 막고자 하는 연구가 필요하며, 실제로 연구되고 있다[4].

DP는 전역적 차분 프라이버시(GDP), 국소적 차분 프라이버시(LDP)로 나뉜다. GDP는 중앙 집계자가 데이터 결합 전체에 노이즈를 추가하는 방식이고, LDP는 각 클라이언트 또는 유저가 자신의 데이터에 노이즈를 추가하는 방식이다. LDP는 GDP에 비해 신뢰성 문제가 발생하지 않는 장점이 있고, GDP는 LDP에 비해 적은 노이즈의 추가에도 전체 데이터 집합에 대한 DP를 보장할 수 있다는 장점이 있다.

최근, DP가 분산 기계 학습과 연합학습 등에 적용되어 사용되고 있다. 이러한 학습 방식에는 여러 기기나 노드가 데이터를 공유하지 않고도 중앙 집중식 모델을 학습할 수 있다. DP를 적용함으로써, 각 참여자의 데이터는 보호되면서도, 전체 모델은 여러 참여자의 데이터로 유의미한 학습을 할 수 있다고 연구되고 있다[48, 65].

3. 프라이버시를 보호하는 데이터 전처리 기법

서버의 주도로 여러 클라이언트와 협력하여 학습을 진행하기 위해서는 각 클라이언트에 분산되어 저장된 데이터셋을 정화하고 통일된 형식에 맞게 정렬한 후, 학습에 사용할 공통 데이터셋을 선별하는 작업이 선수되어야 한다. 최근 동향에 따르면 모든 전처리 과정이 이미 수행되었음을 전제하고, 그 다음 단계인 학습 기법에 중점하여 제안되고 있다[76]. 본 단원은 관련 연구에서 다소 생략되어있는 전처리 과정을 구체적으로 제시하고자 여러 유형의 분산 기계 학습에서 사용되는 프라이버시를 보호하는 전처리 기법을 정리한다.

3.1 Data Cleansing

데이터 정화(Data Cleansing)는 어떠한 유형의 데이터 분석을 수행하더라도 가장 먼저 처리되어야 하는 작업이다. 일반적인 데이터 정화 작업은 데이터에 내재되어 있는 다양한 오류들을 핸들링하기 위해서 사용하는 방법들을 의미한다. 데이터에서 존재하는 오류들은 크게 다섯 가지로 분류될 수 있다 [5]. 첫 번째로, 중복된 데이터셋을 제거하는 것이다. 두 번째로는, 목표와 부합되지 않고 관련성이 낮은 데이터들을 제외하는 과정이다. 세 번째로는 모델 성능에 악영향을 미칠 수 있는 특정 이상값들을 선별하는 과정이다. 네 번째로는 누락된 데이터를 식별하고 제외하거나 보완하는 과정이다. 마지막으로는 구조적으로 통일되지 않은 형식에 따르는 이상 문제를 수정하는 것이다.

일반적인 기계 학습 분야에서 최근까지 제안된 오류 핸들링 과정은 크게 두 가지의 접근 방식으로 분류될 수 있다[6]. 먼저 Outlier 탐지 기반의 데이터 정화 방법[7-10]은 전체 데이터와 비교하였을 때, 다수의 데이터와 멀리 위치하여 있거나, 차이가 크게 나는 데이터 엔트리(entry)를 Outlier로써 식별하여 정화하는 작업이다. 해당 정화 방법을 실현하기 위한 기법으로는 통계적 속성을 활용한 방법[7], 거리 기반 [8], 클러스터링 기반[9], 밀도 기반[10]의 방법 등이 존재한다.

다음으로 기계 학습 기반의 데이터 정화 방법 [11-15]가 있다. 이는 데이터 정화 단계에서 기계 학습을 적용하는 것으로, 어떠한 데이터가 오류인지를 분류하기 위한 모델을 활용하여 학습을 진행한다. 이처럼, 다양한 방식으로 제안된 데이터 정화 작업은 전체 학습의 성능에 직간접적으로 영향을 미칠 수 있는 중요한 작업으로 정확하고 효율적인 방법으로 수행되어야 함을 알 수 있다.

그러나, 연합학습을 포함한 분산 기계 학습에서는 모든 데이터셋이 하나의 데이터베이스에 집결되어 있지 않기 때문에

일반적인 데이터 정화 작업과 비교하여 더 복잡하고 정밀한 과정으로 진행된다. 클라이언트들은 각기 보유 중인 데이터 샘플 및 feature 정보가 서로 다르지만, 프라이버시를 보호하기 위해 각 클라이언트의 데이터 정보들을 노출시키지 않으면서 상호 간에 데이터 정화 작업을 수행해야 한다.

해당 도전 과제에 주목하여 분산 학습 환경에 적용 가능한 데이터 정화 기법이 최근까지 다양하게 제안되고 있다[6, 16-19]. 2011년 Popa 등[17]은 동형 암호를 활용하여 암호화된 데이터를 바탕으로 데이터를 정화할 수 있는 기법을 제안하였다. 2016년 Krishnan 등[16]은 Differential Privacy 기법을 활용하여 데이터 정화 작업을 수행할 수 있는 기법을 제안하였다. 그러나, 위 두 기법은 중앙화된 방식으로 구현되거나, 연산 부하가 높아서 오늘날 모바일 환경(Cross-Device) 및 기관 간 환경(Cross-Silo)에 동시에 적용 가능한 기법이라고 보기 어렵다.

이후 Mohassel 등[18]과 Demmler 등[19]의 연구와 같이 데이터 정화 기법의 실용화를 위해 향상된 이점을 가진 동형 암호 기법 기반의 정화 기법과 안전한 다자간 연산 기법을 통해 제안되기 시작하였다. 가장 최근인 2021년 Ma 등[6]은 데이터 프라이버시를 보존하기 위해 모바일 엣지 환경에 적합한 분산 데이터 정화 기법을 새롭게 제안하였다. 해당 기법은 기존 Outlier인 데이터 엔트리를 식별하는 AVF(Attribute Value Frequency) 알고리즘에 안전한 양자간 연산을 활용하고자 하였다. 각 클라이언트는 두 서버의 주도하에 알고리즘의 연산 결과값(score)을 기준으로 정렬시켜 데이터를 노출하지 않은 채 오류 데이터를 식별하는 기법을 제안하였다.

한편, 현재까지의 연구 동향을 조사한 결과 분산 기계 학습 기법에 직접적으로 적용이 가능한 데이터 정화 기법을 프라이버시를 보호하는 관점에서 제안된 사례는 극히 드물다는 것을 확인하였다. 과거의 연구는 데이터를 한 곳으로 집계하였을 때에만 적용이 가능한 기법이 중점으로 제안되었다. 또한, 대다수의 연구 동향은 학습 과정에서의 프라이버시만 보호하기 위하여 설계되었거나, 프라이버시를 보호하는 데이터 정화 기법만을 집중하여 학습 단계를 고려하지 않았다.

3.2 Privacy-Preserving Record Llinkage

자료 연계(Record Linkage) 기법이란 서로 다른 데이터베이스 간의 결합을 위해서 사용되는 기법[20]으로, 주로 공통된 고유 식별 번호를 사용해 데이터 샘플을 기준으로 연계하는 기법을 의미한다[21]. 전통적인 자료 연계 기법은 개인의 고유 식별자를 생성하기 위해 이름, 성, 나이 등의 개인정보를 활용하였다. 그러나 각 데이터베이스에 저장된 정보가 민감 및 개인정보를 바탕으로 고유 식별자를 생성할 경우 국
내외의 법적 규제에 따라 활용상 제약이 가해질 수 있다. 이러한 문제를 해결하기 위해서 전통적인 자료 연계 기법은 신뢰하는 기관을 중축으로 데이터를 결합하여 데이터 활용의 제약 문제를 해결하고자 하였다. 신뢰 기관이 자료 연계를 수행

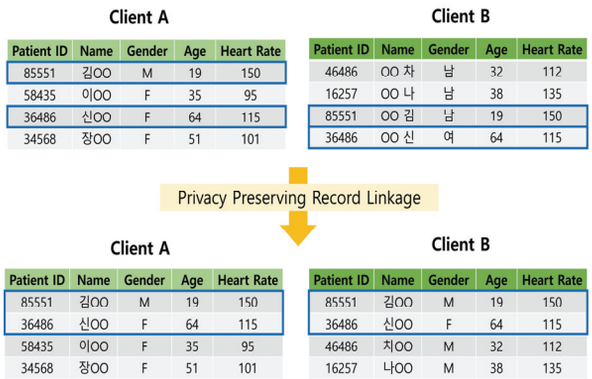

Fig. 2. Privacy Preserving Record Linkage

하여 각 자료에 대한 가명정보를 데이터베이스 관리자에게 전달하는 것이다. 우리나라의 경우 안전하게 서로 다른 개인정보처리자의 보유 정보를 결합하기 위해서 지정된 결합전문기관들이 중앙화된 결합키관리기관의 결합키연계정보를 이용하여 결합을 진행한 후 데이터셋을 생성한다[22]. 그러나, 이러한 방법은 사전에 신뢰하는 중앙 개체를 주축으로 가명정보처리 및 결합이 진행되어야 하는 특징이 있으며, 다량의 데이터를 대상으로 다수의 기관들의 요청이 발생할 경우 시간 및 비용의 측면에서 비효율적일 수 있다. 즉, 현재까지 제안된 자료 연계 기법은 프라이버시 보호 분산 기계 학습 기법에 활용하기에 아직 실용화되어 있지 않다는 한계가 존재한다.

프라이버시를 보호하는 자료 연계(Privacy-Preserving Anonymous Linking, PPRL) 또는 PER(Private Entity Resolution) 기법은 신뢰 기관의 도움을 최소화한 방법으로 데이터베이스 간의 자료 연계를 수행하는 기법으로 프라이버시를 보호하면서 데이터의 활용가치가 높이기 위한 기술들을 의미한다. 일반적으로 PPRL의 기법은 크게 데이터 전처리 단계, Blocking 단계, 매칭 단계 순으로 작동된다[23, 24]. 첫 번째 데이터 전처리 단계에서는 각 클라이언트가 보유하는 데이터베이스에서 통일된 방식으로 연계 시 사용될 데이터의 정화 및 인코딩 과정을 수행한다. 다음으로, 데이터 Blocking 단계에서는 자료들을 최대한 간소화하여 매칭 결과가 모호할 수 있는 데이터들을 사전에 제거함으로써 연계 과정에서의 비교 대상의 수를 줄인다. 해당 과정이 끝나면 매칭 과정에서 사용될 데이터들을 추려내어 해당 자료들을 바탕으로 타 데이터베이스와 더 자세히 비교할 수 있게 된다.

매칭 단계에서는 인코딩 및 Blocking된 데이터들을 대상으로 유사도(Similarity)를 기반으로 그룹 또는 클러스터를 생성하여 자료들을 비교하는 과정을 수행한다. 매칭 단계는 크게 확률론적 매칭 기법과 결정론적 매칭 기법으로 분류될 수 있다 [25, 26, 28]. 확률론적 매칭 기법은 두 개의 자료들이 서로 동일 데이터 샘플에 대한 자료일 가능성을 확률론적 방법으로 계산하는 기법들을 의미한다[27, 28]. 대표적인 기법으로는 1960년대에 제안된 Fellegi-Sunter의 기법[21]으로, 연계 변수

의 값을 매겨 매긴 점수를 바탕으로 매칭과 비매칭의 자료들을 분류하는 기법이다. 이 외에도 효율성을 증대시킬 수 있는 블룸 필터(Bloom Filter) 알고리즘[29]와 이의 발전된 형태인 CLK(Cryptographic Long-term Key) 기법[30] 등이 존재한다. 그러나, 확률론적 기법의 경우 다량의 데이터를 바탕으로 연계하기 수월하다는 장점이 있으나, 정확도가 떨어져 거짓 긍정의 비율이 결정론적 매칭 기법에 비해 높다는 단점이 존재한다.

이와 반대로 결정론적 매칭 기법은 두 자료 간의 연계를 하나 이상의 연계 값들의 일치 및 불일치를 기반으로 연계를 수행하는 기법들을 의미한다 [28]. 대표적인 기법으로는 해시 함수를 사용한 기법[31-33]과 동형 암호 등의 프리미티브를 활용한 암호화 기법[34, 35]가 존재한다. 결정론적 매칭 기법은 거짓 긍정의 비율이 상대적으로 낮으며[36, 37], 유사도 연산 과정이 비교적 단순하여 확장이 가능하다는 장점이 있으나, 대게 신뢰하는 제3의 기관(TTP, Trusted Third Party)을 중심으로 기법이 수행되어 탈중앙성이 떨어진다는 한계가 존재한다.

한편, 최근 2017년 분산 기계 학습 기술이 제안되면서[38] 분산화된 환경에서도 적용 가능한 PPRL 기술의 필요성이 증가되기 시작하였다. 특히, 클라이언트 간에 데이터 샘플이 동일하고, feature가 서로 다른 수직적 연합학습 환경에서 PPRL 기법이 필수적이며, 현재까지 많은 연구[27, 40, 41]이 사전에 전처리 과정이 수행되었음을 전제하고 학습 기법을 설계하고 있다. 2017년 Hardy 등[39]는 두 클라이언트가 수직적으로 분할된 데이터를 보유하고 있을 때 데이터 샘플을 기준으로 PPRL을 실현하기 위하여 부분 동형 암호 및 CLK를 활용하여 자료 연계 기법을 제안하였다. 해당 기법은 전처리 과정을 포함하여 학습 기법을 설계하였다는 의의를 가지나, 서버를 포함하여 3자간 프로토콜로 제안된 기법으로써 확장이 어렵고, 모든 데이터를 암호화해야 한다는 점에서 효율성이 저하된다는 단점이 존재한다.

수직적 연합학습 환경 이외에도 의료 환경과 같이 데이터가 민감 또는 개인정보에 해당될 경우 더욱 안전하게 데이터를 전처리한 후 학습되어야 한다. 이를 실현하기 위해 최근 Stammler 등[42]와 Southwell[43] 등의 연구와 같이 의료 데이터를 바탕으로 안전하게 PPRL을 설계하는 연구가 수행되기 시작하였다. 그러나, 현재까지 제안된 연구에 따르면 특정 산업 분야에 특화된 PPRL 기법이 설계되거나 사용된 기법을 활용하여 통합 학습 기법을 제안하는 연구는 부족한 실정이다.

3.3 Private Set Intersection

PSI(Private Set Intersection)란 두 명 이상의 참여자가 각 보유 데이터들을 노출하지 않으면서 공통 데이터에 대한 교차 정보만을 안전하게 알아내는 기술이다. PSI는 다자간 수행하는 SMC(Secure Multiparty Computation) 기술로, PPRL의 세부 기술이라고 볼 수 있다. PSI 참여자들은 교차 집합에 속하는 공통 데이터 정보 이외 나머지 정보에 대한 프라이버시를

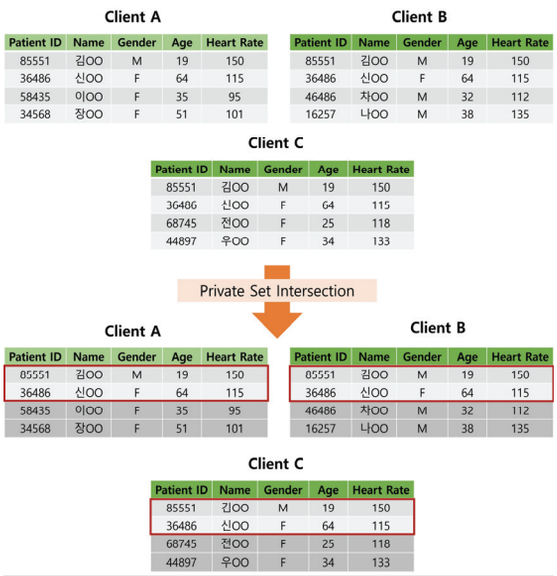

Fig. 3. Private Set Intersection

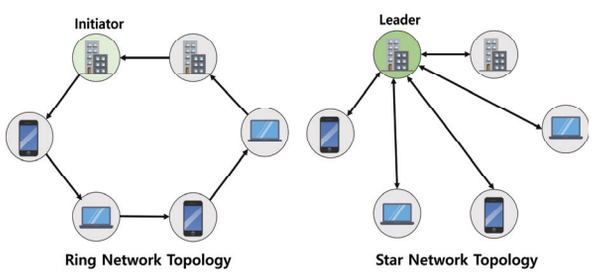

Fig. 4. Type of Network Topology

보호할 수 있어 전체 데이터베이스를 한 곳으로 모으지 않아도 학습 데이터를 정의할 수 있다. 이러한 이유로 분산 기계 학습의 전처리 기술로써 활발히 활용되고 있으며, 특히 데이터 샘플은 동일하나 feature 정보가 다른 수직적 연합학습과, 모델을 분리하여 학습하는 분할 학습의 전처리 과정에서 적용되고 있다. 본 논문에서는 분산 기계 학습 기법에서 활용가능한 PSI 기법의 동향을 분석하고자 다자간 PSI 기법에 집중하여 기술한다.

다자간 PSI 기법에서는 크게 두 종류의 토폴로지 구조로 분류할 수 있다[44]. 첫 번째로는 Fig. 4의 왼쪽 그림과 같이 링 네트워크 토폴로지 구조를 따르는 유형으로, 각 참여자들은 양옆의 참여자들과 연결된 상태에서 첫 번째 참여자(Initiator)부터 순차적으로 통신한다. 링 토폴로지에서 설계된 다자간 PSI 기법들은 참여자 간 비슷한 연산량을 요구로 하는 특징을 가진다.

두 번째로는 스타 네트워크 토폴로지 구조를 따르는 유형으로, 한 참여자(Leader)를 중심으로 나머지 참여자들이 연결된 상태에서 통신을 하며 PSI 기법을 수행한다. 스타 토폴로지 구조는 특정 참여자를 통신 및 리소스 성능이 뛰어난 개체로 가정하고 있다. 따라서, 무거운 연산량을 바탕으로 설계되

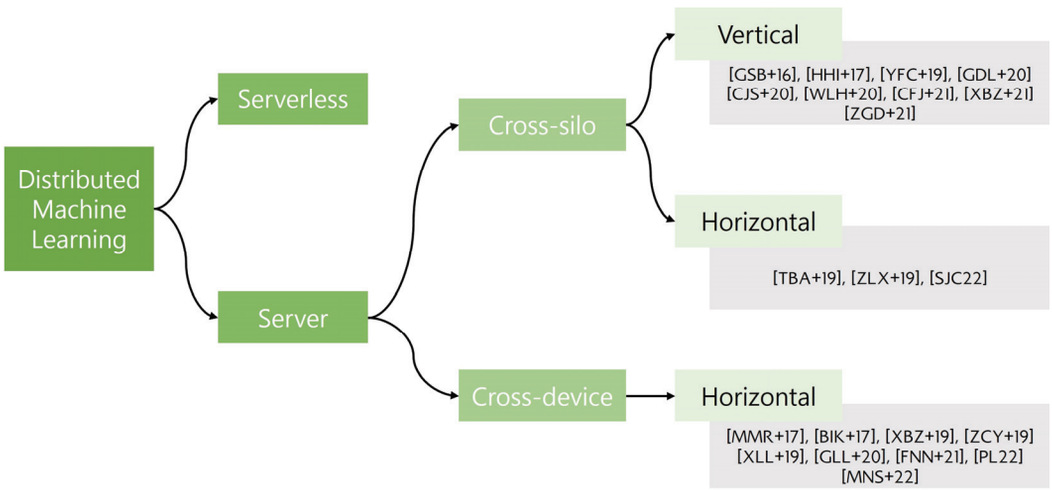

Fig. 5. Classfication of Federated Learning

어도 해당 개체의 도움을 받아 효율적으로 처리할 수 있다는 장점이 존재한다. 그러나, Fig. 4와 같이 참여자 수가 증가할수록 리더 클라이언트의 측면에서 통신 부담이 증가하여 단일 장애점 오류(SPoF, Single Point of Failure)의 가능성이 커지고, 전체 시스템의 신뢰성이 감소될 수 있다. 따라서, 스타 네트워크 토폴로지 기반의 다자간 PSI 기법은 효율성과 안전성의 적절한 Trade-off를 고려하여 설계해야 하는 요구사항이 존재한다.

4. 분산 기계 학습 기법

본 장에서는 클라이언트의 보유 데이터를 노출하지 않으면서 여러 클라이언트와 협력하여 하나의 기계 학습 모델을 학습하는 분산 기계 학습 기법의 최신 연구 동향을 살펴본다. 또한 현재까지 제안된 기법들을 대상으로 서버의 존재 유무, 학습 데이터셋의 특징, 학습 주체(클라이언트)의 기기 성능 등 복합적인 기준에 따라 분류를 진행한다. 이를 통해 다양한 이름으로 제안된 분산 기계 학습 기법의 유사점과 차이점을 발견하고 향후 연구 방향을 분석하고자 한다.

4.1 분산 기계 학습 유형

프라이버시를 보호하는 분산 기계 학습 기법은 세 가지 기준에 따라 분류될 수 있다. 첫 번째로, 중앙 서버의 존재 여부에 따라 Server-based와 Serverless로 분류된다. 중앙 서버가 존재하는 경우 서버의 주관으로 학습을 위한 전처리 과정에서부터 전체 모델을 업데이트하는 과정을 원활히 수행할 수 있다. 그러나, 중앙 서버가 악의적일 경우 학습 과정에서 보안 위협이 발생할 수 있으므로, 최근 연구 동향에서는 중앙 서버가 악의적일 경우에도 프라이버시가 보호되도록 안전하게 설계되거나, 중앙 서버에만 의존하지 않고 학습이 진행 가능한 탈중앙화 분산 기계 학습 기법이 제안되고 있다.

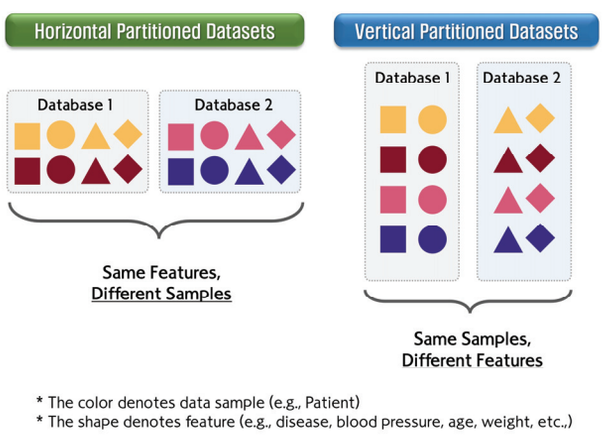

Fig. 6. Classification of Distribution Types in Training Datasets

두 번째로, 학습 데이터셋의 특징에 따라 수평적 분산 기계 학습과 수직적 분산 기계 학습 기법으로 분류된다. Fig. 6은 전체 학습 데이터셋의 모습을 나타낸다. 학습 데이터셋은 크게 레이블 정보와 feature 정보로 구분되며, 각각의 데이터는 데이터 샘플로부터 생성된다. 수평적 분산 기계 학습의 경우 클라이언트 간에 데이터 샘플이 서로 다르되, feature 정보는 동일한 경우를 의미한다. 반면, 수직적 분산 기계 학습의 경우 feature 샘플이 동일하나, 데이터 샘플이 서로 다른 경우를 의미한다. 일례로, 수평적 분산 기계 학습은 하나의 샘플에 대한 feature 정보는 반드시 하나의 클라이언트에만 저장된다는 가정에서 학습된다고 볼 수 있지만, 수직적 분산 기계 학습은 하나의 샘플에 대한 feature 정보가 여러 클라이언트에 분할되어 저장된다는 상황에서 학습을 진행하는 특징을 보인다.

마지막으로, 학습 클라이언트의 기기 성능에 따라 Cross-Device와 Cross-Silo로 구분된다. 클라이언트가 IoT(Internet-of-Things) 기기들과 같이 저사양의 기기들로 전제되는 경우 Cross-Device 환경이라고 불린다. 이와 달리 클라이언트가

Table 1. Horizontal Federated Learning

	Year	Partitioned data	Cryptographic Primitives	Server Existence	Client Types	Model
[46]	2017	HFL	X	O	Cross-Device	FedAvg
[47]	2017	HFL	SS, KA, AE	O	Cross-Device	X
[48]	2019	HFL	AHE, DP, SMC	O	Cross-Silo	DT, CNN
[49]	2019	HFL	SMC, FE	O	Cross-Device	CNN
[50]	2019	HFL	AHE	O	Cross-Device	CNN
[51]	2019	HFL	AHE	O	Cross- Silo	LSTM
[52]	2019	HFL	SS, HH, KA	O	Cross-Device	CNN
[53]	2020	HFL	SS, KA, HH, COM	O	Cross-Device	X
[54]	2021	HFL	SMC,FHE,DP,SS	O	Cross-Device	NLP, IC, NIDS
[55]	2022	HFL	AHE, SMC	O	Cross-Device	high-quality tree boosting
[56]	2022	HFL	AHE	O	Cross-Silo	X or FedAvg
[57]	2022	HFL	FHE	O	Cross-Device	CNN

다량의 데이터셋을 보유하면서 통신 및 연산 성능이 높은 고사양 기기들일 경우에는 Cross-Silo 환경이라고 한다. 각 연구에서 전제하는 클라이언트 성능에 따라 활용되는 프라이버시 보호 기술 및 적용 목적이 달리질 수 있다.

본 논문에서는 위 분류 기준들을 융합하여 현재까지 제안된 연구 내용들을 유기적으로 분석하고자 한다. 이를 통해 다양한 방식으로 제안된 연구 내용들의 특징들을 새롭게 발견하여 향후 연구 방향을 파악하는 것을 목표로 한다. Fig. 5는 최근 연구 내용들을 대상으로 분류한 결과를 보여준다. 각 세부 장에서는 대표적인 프라이버시를 보호하는 분산 기계 학습 기법을 중심으로 현 연구 동향에 대한 분석을 기술한다.

4.2 수평적 연합학습

가장 먼저 제안된 분산 기계 학습 기법 유형인 수평적 연합학습은 McMahan 등[46]에 의해 처음 제안되었다. McMahan 등 [46]은 탈중앙화된 분산 기계 학습 기법을 처음으로 연합학습(Federated Learning)으로 정의하였으며, k명의 클라이언트가 존재할 때, 글로벌 모델을 업데이트하는 알고리즘인 FederatedAveraging 알고리즘을 Table 2와 같이 제안하였다. FederatedAveraging 알고리즘은 확률적 경사하강법(SGD, Stochastic Gradient Descent)에 바탕을 둔다. 가장 대중적인 딥러닝 환경에서 많이 사용되는 확률적 경사하강법은 전체 데이터셋을 정해진 크기를 가진 배치(batch)를 단위로 반복 주기마다 랜덤 샘플을 선택하여 기울기를 계산하는 최적화 과정이다. 매 라운드마다 학습에 참여하도록 선택된 클라이언트는 보유 데이터셋을 바탕으로 로컬 모델을 업데이트한 후 가중치 w를 계산하여 중앙 서버에 전송한다. 중앙 서버는 가중치 값들을 집계(aggregate)하여 평균값을 산출한 후 다음 라운드의 가중치 값으로 업데이트를 진행한다.

2017년 Bonawitz 등 [47]은 해당 개념을 발전하여 프라이버시를 보호하는 분산 기계 학습 기법을 새롭게 제안하였다.

Table 2. FederatedAveraging Algorithm

Central Server:

Initialization:
- Initialize W

For each round t**:**
- $M = \max(C \cdot K, 1)$
- S = select random set of M clients from K
- **For each client** $k \in S$ **in parallel:**
  - $W_k = ClientUpdate_k(W)$
- $W = \sum_{k \in S} W_k$

ClientUpdate(W**):**

Data Splitting:
- Split D into batches of size B

For each local epoch e **from 1 to** E**:**
- **For each batch** b **in** D**:**
  - $W = W - \eta \cdot \nabla L(W; b)$
- **Return** W **to server**

제안된 Secure Aggregation 기법은 Secret Sharing, Key Agreement, 공개키 암호화 및 서명 기법 등을 다양하게 활용하여 서버로 하여금 각 클라이언트의 로컬 모델 업데이트(가중치 w_t) 정보를 알아내기 어렵도록 설계하였다. 각 참여자는 다른 클라이언트의 비밀 조각(secret share) 정보들을 생성한다. 그리고, 각 클라이언트와 양자 간 Pairwise Mask 값을 생성하여 계산된 가중치 값에 Mask 값들을 추가하여 서버에게 전송한다. 서버는 각 참여자로부터 받은 모든 값을 바탕으로 집계하면, 동일 Mask의 Pairwise 쌍들이 더해지고 0으로 최종적으로 계산되어 가중치들의 집계 값만 알게 된다.

만약, 서버가 해당 주기에 참여하는 모든 참여자들로부터 값을 전송받지 못할 경우 Pairwise Mask 값이 0으로 상쇄되지 않을 수 있다. 이러할 경우 서버는 참여한 다른 클라이언트

로부터 못 받은 클라이언트에 해당되는 비밀 조각 정보들을 받아 mask 값을 복원하여 상쇄한다. 해당 연구는 프라이버시를 보호하는 연합학습의 시초라 할 수 있다. 이후, 최근까지 연합학습에서는 클라이언트의 로컬 모델 업데이트 값을 보호하기 위한 많은 연구들이 다양한 프리미티브를 활용하여 제안되었다. 더불어, 클라이언트 이외에 중앙 서버가 악의적이거나 프로토콜은 정직하게 따르지만 중요 정보를 알아내고자 하는 Honest-but-Curious한 개체일 경우에 안전한 기법이 제시되었다.

최근까지 제안된 프라이버시를 보호하는 수평적 연합학습 연구 분석 결과는 Table 1과 같다. 수평적 연합학습은 각 클라이언트가 보유하는 데이터셋의 모습이 수평적으로 분할된 데이터셋일 경우를 뜻한다. 즉, 한 데이터 샘플에 대한 feature 정보들은 반드시 하나의 기기에 저장된다는 가정에 따르는 모델을 의미한다. 이러한 분포 유형은 주로 Cross-Device인 환경에서 자주 정의될 수 있다. 대게 한 IoT(Internet-of-Things) 디바이스는 소유자로부터 발생한 데이터를 수집하기 때문이다. 예를 들어, 구글 키보드인 Gboard 경우 [Gboard] 하나의 모바일 디바이스 내에 저장된 사용자의 데이터를 바탕으로 입력 데이터를 예측하여 수평적 연합학습에 해당된다. Cross-Device 환경에서 설계된 학습 기법들은 클라이언트를 연산량 및 저장 공간의 리소스가 상대적으로 적은 기기들로 가정한다. 이러한 이유로 프라이버시를 보호하기 위해서 활용되는 암호학적 기법들은 높은 연산량을 요구하지 않도록 효율적으로 설계되어야 하는 요구사항이 존재한다. 실제로 높은 연산량을 요구로 하는 동형 암호 기법들을 바탕으로 설계된 연합학습 연구[48, 51, 56]은 클라이언트 유형을 Cross-Silo로 가정하고 있다.

2019년 Xu 등 [52]의 연구에서는 사용자의 로컬 가중치의 기밀성을 보장하기 위해 이중 마스킹 프로토콜을 사용하여, 클라우드 서버가 집계 결과의 정확성에 대한 증명을 제공하도록 요구하여 연합학습 환경에서 보안과 검증 가능성을 향상시킴을 보였다. 그 후, 2020년 Guo 등[53]의 연구에서는 고차원 가중치를 갖는 클라이언트에게 유리하며 통신 효율성이 높고 빠르게 검증가능한 프로토콜을 제안하였다. Guo 등 연구의 목표와 같이 높은 통신, 계산 비용을 해결하기 위해, 2021년 Fereidooni 등[54]의 연구에서는 각 학습 단계에서 2라운드의 통신만 필요로 하며, 높은 계산 비용을 발생시키는 암호화 기법을 사용하지 않고, 안전한 양자간 계산을 통해 더 효율적인 프로토콜을 제안하였다. Park 등[55]의 연구에서는 2022년 중앙 서버가 개인정보를 추론하는 것을 방지하기 위해, 동형 암호를 사용하여 암호화된 로컬 모델 파라미터를 복호화 없이 집계하는 알고리즘을 제안함으로써, 프라이버시를 강화하였다. 이와 같은 동형 암호를 사용한 또 다른 연구로는, Ma 등[57]의 연구에서는 모델 업데이트가 집계된 공개 키를 통해 암호화되어 서버와 공유되기 전에 암호화를 진행하였다. 이러한 모델에서는 참여하는 모든 클라이언트의 협력이 필요하다. 따라서 퍼블릭하게 공유된 모델 업데이트로부터의 프라이버시를 더 강화하였다. 이어, Shin 등[56]의 연구에서는 동형 암호화를 기반으로 클라이언트 간에 안전한 계산을 통해 데이터

Table 3. Vertical Federated Learning

	Year	Type	Crypto.	Novelty of PM	Utilized PM	Referring PM	Server Existence	Client Types	Label Owing Party	Model
[62]	2016	VFL	SMC	X	X	X	O	Cross-Silo	One party	Linear Regression
[39]	2017	VFL	AHE	△	Entity Resolution,CLK	[C12]	O	Cross-Silo	One party	Linear Regression
[63]	2019	VFL	AHE	X	X	[SHL+17]	O	Cross-Silo	Active/ One party (Guest)	Logistic Regression
[64]	2020	VFL	X	X	X	X	O	Cross-Silo	Active parties	Nonlinear (Classification)
[65]	2020	VFL	DP	X	X	X	O	Cross-Silo	Active parties	Logistic Regression, Deep Learning(CNN)
[66]	2020	VFL	DP, AHE, SMC	X	Entity Resolution	[SHL+17]	O	Cross-Silo	Active Parties, Passive Parties	Classification
[67]	2021	VFL	AHE	X	PSI	[LC04]	O	Cross-Silo	Active parties	Regression
[40]	2021	VFL	FE	X	Entity Resolution	[C19]	O	Cross-Silo	One party	Logistic Regression
[68]	2021	VFL	X	X	X	X	O	Cross-Silo	Active parties	Logistic Regression

Crypto. : Cryptographic Primitives, PM : Preprecessing methods, VFL : Vertical Federated Learning, SMC : Secure Multiparty Computation, AHE : Additive Homomorphic Encryption, DP : Differential Privacy, FE : Functional Encryption

셋 크기와 집계된 로컬 업데이트 파라미터를 보호함으로써, 데이터셋이 클라이언트 간에 균일하게 분포되지 않았거나 일부 클라이언트가 각 라운드에서 탈락될 때 데이터셋 크기를 보호하는 장점을 갖는 것을 보였다.

4.3 수직적 연합학습

전체 데이터셋이 수직적으로 분할되어 각 클라이언트가 나누어 가지고 있음을 가정하는 수직적 연합학습은 보다 실제 환경에 근사하도록 설계되어 기존 수평적 연합학습의 한계를 해결하고자 하는 학습 방법론이다. 앞서 기술된 수평적 연합학습은 학습 서비스 제공 기업의 이익 충돌 등의 이유로 실용화의 한계를 가진다[58]. 각 클라이언트가 서비스 제공 기업일 경우 보유 데이터 샘플과 feature 정보를 공유하는 과정에서 이익 충돌이 발생할 수 있기 때문이다. 그러나, 수직적 연합학습은 애플리케이션 환경을 학습 기법에 맞게 적절히 정의하고, 전처리 과정을 초기에 잘 구현한다면 각 클라이언트는 동일 데이터 샘플에 대한 feature 정보를 공유할 수 있다. 그 결과, 제공 기업의 이익 충돌 발생 가능성을 낮출 수 있어 여러 기관이 협력하여 모델을 생성 가능하다.

수직적 연합학습은 하나의 데이터 샘플에 대한 여러 feature들이 다수의 클라이언트에 분할되어 저장됨을 가정한다. 예를 들면 의료 환경에서의 수직적 연합학습을 수행할 시 한 환자는 안과, 피부과 등 다양한 병원에 동시에 방문할 수 있음을 전제하는 것이다. 이 가정은 프라이버시 보호를 실현해야 하는 의료 환경과 같이 민감 정보를 활용하고자 하는 특수 분야에 매우 합리적일 수 있는 가정이며, 이러한 이유로 Table 3과 같이 최근에 많은 연구가 활발히 진행되고 있다.

수직적 연합학습은 대게 Cross-Silo 환경을 가정하고 있다. 그 이유는 데이터 샘플에 대한 정보가 여러 클라이언트에 나뉘어 저장될 가능성은 Cross-Device 환경보다 Cross-Silo 환경에서 높이 존재하기 때문이다. 또한, 수평적 연합학습과 달리 학습 모델을 분할(split)하여 서버와 클라이언트별로 달리 가진다. 주로 함께 적용되는 분할 학습(Split Learning)은 연합학습 이후에 제시된 분산 학습 개념으로. 서버는 상위 부분의 모델(Top Model)을. 각 클라이언트는 하위 부분의 모델(Sub-Model)을 바탕으로 학습을 진행한다. 분할 학습과 함께 학습되는 수직적 연합학습은 클라이언트가 모델의 얕은(Shallow) 부분을 로컬에서 먼저 학습을 진행하고, 연산 결과값(Intermediate Result)을 서버에게 전송한다. 이후, 서버는 각 클라이언트로부터 받은 연산 결과값을 입력값으로 하여 깊은(Deep) 레이어에 해당하는 모델을 마저 학습한다. 분할 학습과 융합된 수직적 연합학습은 단편적으로 진행하는 분할 학습과 달리 클라이언트 측의 학습 과정을 병렬적으로 처리할 수 있어[59] 효율적이며, 전달되는 값이 학습 모델의 업데이트 정보가 아닌 중간 결과값으로 중간자 공격 등으로부터 더 안전해질 수 있는 특징을 가진다.

한편, 수직적 연합학습은 서로 다른 참여자가 보유하고 있는 데이터셋을 공유하지 않은 채 동일 데이터 샘플에 대한 feature 정보를 나열해야 하므로, 학습을 진행하기 전 프라이버시를 보호하는 전처리 과정이 필요하다. 전처리 과정에서는 보유 데이터셋을 노출하지 않아야 하며, 이질적인 포맷 환경에서 학습에 사용될 데이터셋을 선택해야 하므로 이는 해결하기 어려운 동시에 반드시 진행되어야 하는 절차에 해당한다.

그러나, 최신 연구 동향을 분석한 결과 대부분의 연구는 전처리에 활용 가능한 암호 프리미티브 논문을 단순히 참조하고, 그 이후 단계인 학습 단계에 초점을 두어 프라이버시를 보호하는 학습 기법을 제안하였다. 또한, 참조하는 Privacy-Preserving Record Linkage, PSI, Entity Resolution 기법들은 대부분 과거에 이론적으로 제안된 연구일 뿐 제안하는 학습 기법에 특화되어 제안된 전처리 기법이 아니므로 학습 알고리즘에 직접적으로 적용하는 것이 어려울 수 있다는 한계점이 발견되었다[39, 40, 63, 66-68].

한편, 학습 단계에서는 대다수의 연구가 프라이버시를 보호하는 기법을 바탕으로 안전한 수직적 연합학습 알고리즘들이 제안되었다. 일례로, 2016년 Gascón 등[62]은 처음으로 수직적으로 분할된 데이터를 대상으로 학습 가능한 프로토콜을 Garbled Circuit 기반 다자간 연산(SMC, Secure Multi-party Computation) 기법을 바탕으로 설계하였다. 이후 동형 암호[39, 63, 66, 67]과 차등 정보보호[65, 66] 기법을 활용하는 다수의 수직적 연합학습 알고리즘도 제안되었다. 또한, 2021년 Xu 등[40]은 학습 데이터를 클라이언트 간 보호하도록 함수 암호(Functional Encryption)를 활용한 기법을 제안하였다. 함수 암호는 하나 이상의 암호문과 개인키를 입력하면, 연암호화된 평문들을 정해진 함수를 통해 계산한 연산 결과값을 평문으로 출력하는 기법이다. 함수 암호는 연산 결과값이 평문으로 출력된다는 점에서 동형 암호 기법과 차별된다. 또한, 학습 모델도 동형 암호처럼 선형 모델에 국한되지 않아 비선형 기계 학습 모델도 학습할 수 있다. 더불어, 근사 오차에 대한 추가적인 연산을 필요로 하지 않아 정확도 및 효율성도 향상시킬 수 있다. 이러한 장점으로 최근 함수 암호를 활용한 학습 기법 설계에도 많은 주목을 받고 있다.

수직적 연합학습은 최근 제안된 학습 유형이므로 공통적인 시스템 모델을 따르지 않으며, 정의하는 시스템 모델에 따라 서로 다른 학습 기법이 설계된다. 일례로, 레이블 정보를 소유하는 주체를 서버로 정의하는 연구[39, 62, 65]와 클라이언트 중 하나로 정의하는 연구[40, 63,] 다수의 클라이언트로 정의하는 연구[64, 66-68]로 분류된다. 레이블 정보의 소유 주체를 서로 달리 정의하면 전체 모델을 업데이트하는 과정이 동일 암호학적 프리미티브를 사용된다고 할지라도 프로세스가 달라질 수 있다. 또한, 최근에 제안된 연구를 제외하고, 대다수의 연구는 레이블 소유 주체를 단일 개체로 가정하였음을 발견하였다. 상대적으로 신뢰하는 개체인 중앙 서버 혹은 리더 클라이언트가 전체 보유 데이터셋에 대한 레이블 정보를 초기부터 가지고 있음을 가정한다는 것이다. 그러나, 이러한 가정

은 실 환경에 근사하도록 설계하고자 하는 수직적 연합학습의 추구 목표와 충돌되는 가정이다. 실제 환경에서는 다수의 개체가 보유 데이터셋의 부분화된 레이블 정보를 각각 나눠가질 수 있는 확률이 존재한다. 하지만, 레이블 정보가 분할되어 저장될 경우 프라이버시를 보호하는 학습 기법을 설계하기가 도전적인 이유로, 현재까지는 해당 모델을 바탕으로 설계된 연구가 미비하다. 레이블 프라이버시에 대한 보안 요구사항을 추가로 고려해야 하며, 최종적으로 모델을 집계(Aggregation)하는 프로토콜이 복잡해질 수 있기 때문이다.

이처럼, 정의하는 환경에 적합하도록 안전하게 학습 기법을 설계하는 것도 의의가 있으나, 분산 기계 학습의 활용 가능성을 높이고, 더 넓은 적용 분야로의 확장을 고려하기 위해서는 가장 대중적인 실제 환경과 유사하도록 시스템 모델을 표준화하여 정의하는 과정이 향후 연구에 필요하다. 이를 실현하기 위해선 수직적 연합학습 연구는 이절적인 데이터베이스를 보유하고 있는 다중 클라이언트 간에 전처리 단계를 안전하게 수행하고, 이후 학습 단계에서 최종 모델을 업데이트하는 완전한 단계에서의 프라이버시 보호 기법에 관한 연구가 앞으로 필요해질 것으로 보인다.

4.4 스웜 학습

스웜 학습(Swarm Learning)은 분산 기계 학습의 새로운 패러다임으로, 수직적 연합학습과 유사하게 다수의 클라이언트 간 협력을 통해 학습을 진행하는 방식이다. 이는 2021년 Warnat 등[69]의 연구에서 제시된 바와 같이, 각 클라이언트가 독립적으로 Covid-19, 결핵, 백혈병, 폐 질환 등의 데이터를 처리하고 학습 모델을 업데이트하며, 이러한 업데이트를 네트워크를 통해 공유하는 방식으로 진행되어 효과적으로 질병을 예측할 수 있음을 보이며. 중앙 서버의 필요성을 최소화하며, 데이터의 프라이버시와 보안을 강화하는 데 중점을 두었다.

스웜 학습의 핵심은 각 클라이언트가 로컬 데이터를 기반으로 독립적으로 학습을 진행하고, 학습된 모델의 파라미터만을 네트워크를 통해 공유하는 것이다. 2022년 Saldanha 등[70]의 연구에서는 이러한 접근이 의료 분야에서 특히 유용함을 보였다. 구체적으로는, 스대규모 다기관 데이터셋에서 대장암의 분자 변이를 예측하는 인공지능 모델을 개발함으로써, 스웜 학습이 프라이버시를 보장함과 동시에 복잡한 히스토파톨로지 이미지를 분석하는데 효과적임을 보였다.

스웜 학습은 의료, 금융, IoT(Internet of Things)등의 분야에서 유용하게 적용될 수 있다. 2022년 Basak 등[71]의 연구에서는 인간 활동 인식을 위한 스웜 학습 기반 딥러닝 접근법을 제시하며 스웜 학습과 딥러닝을 결합 DSwarm-Net을 제안하였다. 이는 다양한 병원과 연구 기관이 각자의 환자 데이터를 기반으로 독립적으로 모델을 학습하고, 이를 통해 얻은 지식을 공유함으로써 보다 정확하게 인간의 행동을 인식하여 포괄적인 진단 모델을 개발할 수 있음을 보였다.

스웜 학습의 구현에 있어서는 다양한 암호학적 기법이 적용될 수 있다. Wang 등[72]의 연구에서는 대규모 최적화 문제를 해결하기 위한 스웜 학습 기반 알고리즘을 제안하여, 클라이언트 간에 공유되는 정보가 외부에 노출되더라도, 실제 데이터의 내용을 보호할 수 있게 해주고 동시에 효율성을 증진시켰다.

스웜 학습은 아직 초기 단계의 연구 분야이며, 이에 대한 구체적인 연구와 적용 사례가 점차 증가하고 있다. 특히, 스웜 학습이 적용될 수 있는 다양한 실제 환경과 시나리오에 대한 연구가 필요하며, 이를 통해 분산 기계 학습의 새로운 가능성을 탐색할 수 있을 것이다. 또한, 스웜 학습의 효율성과 안정성을 높이기 위한 기술적인 개선과 함께, 데이터 프라이버시와 보안을 강화하는 방법에 대한 연구도 중요한 과제로 남아있다.

5. 결 론

프라이버시를 보장하는 분산 기계 학습 연구는 다수의 참여자가 보유 데이터셋을 노출하지 않으면서 협력을 통해 하나의 학습 모델을 안전하게 생성할 수 있도록 보장한다. 일반적인 분산 기계 학습은 대부분 학습 성능의 개선을 위해 데이터 이질성 및 Non-IID 문제 해결 등에 다수 집중되고 있다. 본 연구는 이와 달리 사용자의 데이터셋을 보호하기 위해 프라이버시를 강화하는 분산 기계 학습 분야에 중점을 두어 동향을 탐색하였다. 특히, 현재까지 활발히 연구되고 있는 분산 기계 학습 연구를 서버의 존재 유무, 데이터셋의 분포 환경, 참여자의 리소스 등 다양한 분류기준을 바탕으로 유기적으로 분류 및 분석하였다. 그중 대표적인 분산 기계학습 연구인 수평적 연합학습, 수직적 연합학습, 스웜 러닝에 관련된 연구 동향을 살펴보았다.

연구 결과 제안된 연구들은 학습 모델, 데이터 분포 환경 등에 특화하여 세분화된 환경에서 안전하고 효율적인 학습 기법을 설계하고 있음을 공통적으로 발견하였다. 이로 인해 학습 가능한 모델들이 제한적일 수 있으며, 활용하는 암호 프리미티브에 따라 실용화의 한계를 가질 수 있음을 보였다. 이러한 문제점은 특히 수직적 연합학습 연구에서 두드러지게 나타났으며, 레이블 정보 등 학습 데이터셋의 일부를 보유하고 있는 주체를 누구로 가정하느냐에 따라 전체 학습 기법 및 시스템 모델의 가정 사항이 달라졌다. 향후 분산 기계 학습의 실용화를 위해서는 실제적인 환경에 근사하여 범용적으로 사용될 수 있는 안전한 프레임워크 개발 연구가 필요할 것으로 보인다.

또한, 본 연구는 탈중앙화된 분산 기계 학습 연구가 대두되면서 주목을 받고 있는 스웜 학습에 대해서도 분석하였다. 스웜 학습에서는 수평적 연합학습 및 수직적 연합학습과 달리 아직까지는 프라이버시를 보호하기 위한 연구들이 많이 제시되지 않았다. 이에 학습 토폴로지와 프라이버시 보호 대상에 대한 정의가 이루어져야 할 것이며, 이를 바탕으로 확장 연구를 진행하는 것이 유망한 연구 분야일 것으로 전망된다.

References

[1] A. Shamir, "How to share a secret," *Communications of the ACM*, Vol.22, No.11, pp.612-613, 1979.

[2] W. Diffie and M. E. Hellman, "New directions in cryptography," *Democratizing Cryptography: The Work of Whitfield Diffie and Martin Hellman*, pp.365-390, 2022.

[3] P. Paillier, ''Public-key cryptosystems based on composite degree residuosity classes,'' in *International Conference on the Theory and Applications of Cryptographic Techniques*, pp.223-238, 1999.

[4] Q. Li, Z. Wen, Z. Wu, S. Hu, N. Wang, and Y. Li, "A survey on federated learning systems: Vision, hype and reality for data privacy and protection," *IEEE Transactions on Knowledge and Data Engineering,* 2021.

[5] "What is Data Cleansing?" [Internet], https://aws.amazon.com/ko/what-is/data-cleansing/

[6] L. Ma, Q. Pei, L. Zhou, H. Zhu, L. Wang, and Y. Ji, "Federated Data Cleaning: Collaborative and Privacy-Preserving Data Cleaning for Edge Intelligence," in *IEEE Internet of Things Journal*, Vol.8, No.8, pp.6757-6770, 2021. doi: 10.1109/JIOT.2020.3027980.

[7] A. Koufakou, E. G. Ortiz, M. Georgiopoulos, G. C. Anagnostopoulos, and K. M. Reynolds, "A scalable and efficient outlier detection strategy for categorical data," in *19th IEEE International Conference on Tools with Artificial Intelligence* (ICTAI), Vol.2, pp.210-217, 2007.

[8] S. D. Bay and M. Schwabacher, "Mining distance-based outliers in near linear time with randomization and a simple pruning rule," in *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp.29-38, 2003.

[9] F. Jiang, G. Liu, J. Du, and Y. Sui, "Initialization of K-modes clustering using outlier detection techniques," *Information Sciences*, Vol.332, pp.167-183, 2016.

[10] M. M. Breunig, H.-P. Kriegel, R. T. Ng, and J. Sander, "LOF: Identifying density-based local outliers," *Proceedings of the 2000 ACM SIGMOD international conference on Management of data*, Vol.29, No.2, pp.93-104, 2000.

[11] A. Arasu, M. Götz, and R. Kaushik, "On active learning of record matching packages," in *Proceedings of the 2010 ACM SIGMOD International Conference on Management of data*, pp.783-794, 2010.

[12] S. Mudgal et al., "Deep learning for entity matching: A design space exploration," in *Proceedings of the 2018 International Conference on Management of Data*, pp.19-34, 2018.

[13] T. Rekatsinas, X. Chu, I. F. Ilyas, and C. Ré, "Holoclean: Holistic data repairs with probabilistic inference," *Proceeding VLDB Endowment*, Vol.10, No.11, pp.1190-1201, 2017.

[14] M. Yakout, L. Berti-Équille, and A. K. Elmagarmid, "Don't be SCAREd: Use SCalable automatic REpairing with maximal likelihood and bounded changes," in *Proceedings of the 2013 ACM SIGMOD International Conference on Management of Data*, pp.553-564, 2013.

[15] M. Yakout, A. K. Elmagarmid, J. Neville, M. Ouzzani, and I. F. Ilyas, "Guided data repair," *Proceeding VLDB Endowment*, Vol.4, No.5, pp.279-289, 2011.

[16] S. Krishnan, J. Wang, M. J. Franklin, K. Goldberg, and T. Kraska, "PrivateClean: Data cleaning and differential privacy," in *Proceedings of the 2016 International Conference on Management of Data*, pp.937-951, 2016.

[17] R. A. Popa, C. Redfield, N. Zeldovich, and H. Balakrishnan, "CryptDB: Protecting confidentiality with encrypted query processing," in *Proceedings of the twenty-third ACM symposium on operating systems principles*, pp.85-100, 2011.

[18] P. Mohassel and Y. Zhang, "SecureML: A system for scalable privacypreserving machine learning," in *IEEE Symposium on Security and Privacy* (SP), San Jose, CA, USA, pp.19-38, 2017.

[19] D. Demmler, T. Schneider, and M. Zohner, "Aby-a framework for efficient mixed-protocol secure two-party computation," in *Network and Distributed System Security (NDSS)*, pp.59, 2015.

[20] H. L. Dunn, "Record linkage," *American Journal of Public Health Nations Health*, Vol.36, No.12, pp.1412-1416, 1946.

[21] I. P. Fellegi and A. B. Sunter, "A Theory for Record Linkage", *Journal of the American Statistical Association*, Vol.64, No.328, pp.1183-1210, 1969.

[22] 가명정보결합종합지원시스템 [Internet], https://link.privacy.go.kr/nadac/organ/introData.do

[23] D. Vatsalan, Z. Sehili, P. Christen, and E. Rahm, "Privacy-Preserving Record Linkage for Big Data: Current Approaches and Research Challenges," In: Zomaya, A., Sakr, S. (eds) Handbook of Big Data Technologies. Springer, Cham. 2017. https://doi.org/10.1007/978-3-319-49340-4_25

[24] A. Gkoulalas-Divanis, D. Vatsalan, D. Karapiperis, and M. Kantarcioglu, "Modern privacy-preserving record linkage techniques: An overview," in *IEEE Transactions on Information Forensics and Security*, Vol.16, pp.4966-4987, 2021. doi: 10.1109/TIFS.2021.3114026

[25] S. Gomatam, R. Carter, M. Ariet, and G. Mitchell, "An empirical comparison of record linkage procedures," *Statistics in Medicine*, Vol.21, No.10, pp.1485-1496, 2002.

[26] Peter Christen, "Data matching: concepts and techniques for record linkage, entity resolution, and duplicate detection," *Springer Science & Business Media*, 2012.

[27] A. P. Brown, C. Borgs, S. M. Randall, and R. Schnell, "Evaluating privacy-preserving record linkage using cryptographic long-term keys and multibit trees on large medical datasets," *BMC Medical Informatics and Decision Making*, Vol.17, pp.1-7, 2017. https://doi.org/10.1186/s12911-017-0478-5

[28] I. Lazrig, T. C. Ong, I. Ray, I. Ray, X. Jiang, and J. Vaidya, "Privacy preserving probabilistic record linkage without trusted third party," in *2018 16th Annual Conference on Privacy, Security and Trust* (PST), pp.1-10, 2018.

[29] B. H. Bloom, "Space/time trade-offs in hash coding with allowable errors," *Communications of the ACM*, Vol.13, No.7, pp.422-426, 1970.

[30] R. Schnell, T. Bachteler, and J. Reiher, "A novel error-tolerant anonymous linking code," *Social Science Research Network*, WP-GRLC-2011-02, 2011.

[31] Christine M. O'Keefe, Ming Yung, Lifang Gu, and Rohan Baxter. 2004. "Privacy-preserving data linkage protocols," In *Proceedings of the 2004 ACM Workshop on Privacy in the Electronic Society* (WPES '04). Association for Computing Machinery,NY,USA,94-102. https://doi.org/10.1145/1029179.1029203

[32] S. B. Dusetzina, S. Tyree, A.-M. Meyer, A. Meyer, L. Green, and W. R. Carpenter, "An Overview of Record Linkage Methods," 2014.

[33] S. B. Johnson, G. Whitney, M. McAuliffe, H. Wang, E. McCreedy, L. Rozenblit, and C. C. Evans, "Using global unique identifiers to link autism collections," *Journal of the American Medical Informatics Association*, Vol.17, No.6, pp.689-695, 2010.

[34] A. Inan, M. Kantarcioglu, G. Ghinita, and E. Bertino, "Private record matching using differential privacy," in *Proceeding EDBT*, pp.123-134, 2010.

[35] M. Kuzu, M. Kantarcioglu, A. Inan, E. Bertino, E. Durham, and B. Malin, "Efficient privacy-aware record integration," in *Proceeding EDBT*, Genoa, Italy, pp.167-178, 2013.

[36] A. L. Potosky, G. F. Riley, J. D. Lubitz, R. M. Mentnech, and L. G. Kessler, "Potential for cancer related health services research using a linked Medicare-tumor registry database," *Medical Care*, Vol.31, No.8, pp.732-748, 1993.

[37] S. J. Grannis, J. M. Overhage, and C. J. McDonald, "Analysis of identifier performance using a deterministic linkage algorithm," *Proceeding of AMIA Symposium*, pp.305-309, 2002.

[38] B. McMahan, E. Moore, D. Ramage, S. Hampson, and y Arcas, "Communication-efficient learning of deep networks from decentralized data," *Artificial Intelligence and Statistics*, Vol.54, 2017.

[39] S. Hardy, W. Henecka, H. Ivey-Law, R. Nock, G. Patrini, G. Smith, and B. Thorne, "Private federated learning on vertically partitioned data via entity resolution and additively homomorphic encryption," *arXiv preprint arXiv: 1711.10677*, 2017.

[40] R. Xu, N. Baracaldo, Y. Zhou, A. Anwar, J. Joshi, and H. Ludwig, "Fedv: Privacy-preserving federated learning over vertically partitioned data," *Proceedings of the 14th ACM Workshop on Artificial Intelligence and Security*, 2021.

[41] D. Romanini, A. J. Hall, P. Papadopoulos et al., "Pyvertical: A vertical federated learning framework for multi-headed splitnn," *arXiv:2104.00489*, 2021.

[42] S. Stammler et al., "Mainzelliste SecureEpiLinker (MainSEL): Privacypreserving record linkage using secure multi-party computation," *Bioinformatics*, Vol.2020, pp.1-12, 2020.

[43] A, Southwell et al., "Validating a novel deterministic privacy-preserving record linkage between administrative & clinical data: applications in stroke research," *International Journal of Population Data Science,* Vol.7, No.4, pp.1755, 2022. doi: 10.23889/ijpds.v7i4.1755. PMID: 37152407; PMCID: PMC10161965.

[44] D. Morales, I. Agudo, and J. Lopez, "Private set intersection: A systematic literature review," *Computer Science Review*, Vol.49, pp.100567, 2023, https://doi.org/10.1016/j.cosrev.2023.100567.

[45] A. Adir, E. Aharoni, N. Drucker, E. Kushnir, R. Masalha, M. Mirkin and O. Soceanu, "Privacy-preserving record linkage using local sensitive hash and private set intersection," *ArXiv:2203.14284v1*, 2022.

[46] B. McMahan, E. Moore, D. Ramage, S. Hampson, and y Arcas, "Communication-efficient learning of deep networks from decentralized data," *Artificial Intelligence and Statistics*, PMLR, 2017.

[47] K. Bonawitz et al., "Practical secure aggregation for privacy-preserving machine learning," *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, 2017.

[48] S. Truex, "A hybrid approach to privacy-preserving federated learning," *Proceedings of the 12th ACM Workshop on Artificial Intelligence and Security*, 2019.

[49] R. Xu, N. Baracaldo, Y. Zhou, A. Anwar and H. Ludwig, "Hybridalpha: An efficient approach for privacy-preserving federated learning," *Proceedings of the 12th ACM Workshop on Artificial Intelligence and Security*, 2019.

[50] J. Zhang, B. Chen, S. Yu, and H. Deng, "PEFL: A privacy-enhanced federated learning scheme for big data analytics," *2019 IEEE Global Communications Conference (GLOBECOM)*, IEEE, 2019.

[51] C. Zhang, S. Li, J. Xia, and W. Wang, "{BatchCrypt}: Efficient homomorphic encryption for {Cross-Silo} federated learning," *2020 USENIX Annual Technical Conference (USENIX ATC 20)*, 2020.

[52] G. Xu, H. Li, S. Liu, K. Yang and X. Lin, "Verifynet: Secure and verifiable federated learning," *IEEE Transactions on Information Forensics and Security*, Vol.15, pp.911-926, 2019.

[53] X. Guo et al., "VeriFL: Communication-Efficient and Fast Verifiable Aggregation for Federated Learning," *IEEE Transactions on Information Forensics and Security*, Vol.16, pp.1736-1751, 2020.

[54] H. Fereidooni et al., "SAFELearn: Secure aggregation for private federated learning," *2021 IEEE Security and Privacy Workshops (SPW)*, IEEE, 2021.

[55] J. Park and H. Lim, "Privacy-preserving federated learning using homomorphic encryption," *Applied Sciences*, Vol.12, No.2, pp.734, 2022.

[56] Y. A. Shin, G. Noh, I. R. Jeong, and J. Y. Chun, "Securing a local training dataset size in federated learning," *IEEE Access*, Vol.10, pp.104135-104143, 2022.

[57] J. Ma, SA. Naas, S. Sigg, and X. Lyu, "Privacy-preserving federated learning based on multi-key homomorphic encryption," *International Journal of Intelligent Systems*, Vol.37, No.9, pp.5880-5901, 2022.

[58] Y. Cheng, Y. Liu, T. Chen, and Q. Yang, "Federated learning for privacy-preserving AI," *Communications of the ACM*, Vol.63, No.12, pp.33-36, 2020.

[59] M. G. Poirot, P. Vepakomma, K. Chang, J. K.Cramer, R. Gupta, and R. Raskar, "Split Learning for collaborative deep learning in healthcare," *NeurIPS*, 2019.

[60] B. McMahan and D. Ramage, Google Research, Apr. 2017, [Online] Available: https://blog.research.google/2017/04/federated-learning-collaborative.html

[61] A. Hard et al., "Federated learning for mobile keyboard prediction," *arXiv preprint arXiv:1811.03604*, 2018.

[62] A. Gascón, P. Schoppmann, B. Balle, M. Raykova, J. Doemer, S. Zahur and D. Evans, "Secure linear regression on vertically partitioned datasets," *International Association for Cryptologic Research Cryptology ePrint Archive*, 892, 2016.

[63] K. Yang, T. Fan, T. Chen, Y. Shi, and Q. Yang, "A quasi-newton method based vertical federated learning framework for logistic regression," *arXiv preprint arXiv:1912.00513*, 2019.

[64] B. Gu, Z. Dang, X. Li, and H. Huang, "Federated doubly stochastic kernel learning for vertically partitioned data," *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2020.

[65] T. Chen, X. Jin, Y. Sun, and W. Yin, "Vafl: a method of vertical asynchronous federated learning," *arXiv preprint arXiv:2007.06081*, 2020.

[66] C. Wang, J. Liang, M. Huang, B. Bai, K. Bai, and H. Li, "Hybrid differentially private federated learning on vertically partitioned data," *arXiv preprint arXiv:2009.02763*, 2020.

[67] K. Cheng et al., "Secureboost: A lossless federated learning framework," *IEEE Intelligent Systems*, Vol.36, No.6, pp.87-98, 2021.

[68] Q. Zhang, B. Gu, C. Deng, and H. Huang, "Secure bilevel asynchronous vertical federated learning with backward updating," *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol.35, No.12, 2021.

[69] S. Warnat-Herresthal et al., "Swarm Learning for decentralized and confidential clinical machine learning," *Nature*, Vol.594, pp.265-270, 2021.

[70] O. L. Saldanha et al., "Swarm learning for decentralized artificial intelligence in cancer histopathology," *Nature Medicine*, Vol.28, No.6, pp.1232-1239, 2022.

[71] H. Basak, R. Kundu, PK. Singh, MF. Ijaz, M. Woźniak, and R. Sarkar, "A union of deep learning and swarm-based optimization for 3D human action recognition," *Scientific Reports*, Vol.12, No.1, pp.5494, 2022.

[72] F. Wang, X. Wang, and S. Sun, "A reinforcement learning level-based particle swarm optimization algorithm for large-scale optimization," *Information Sciences*, Vol.602, pp.298-312, 2022.

[73] M. Al-Rubaie and J. M. Chang, "Privacy-preserving machine learning: Threats and solutions," *IEEE Security & Privacy*, Vol.17, No.2, pp.49-58, 2019.

[74] R. Xu, N. Baracaldo, and J. Joshi. "Privacy-preserving machine learning: Methods, challenges and directions," *arXiv preprint arXiv:2108.04417*, 2021.

[75] G. A. Kaissis, Kaissis, M. R. Makowski, D. Ruckert, and R. F. Braren, "Secure, privacy-preserving and federated machine learning in medical imaging," *Nature Machine Intelligence*, Vol.2, No.6, pp.305-311, 2020.
[76] A. Lau, and J. Passerat-Palmbach. "Statistical privacy guarantees of machine learning preprocessing techniques," *arXiv preprint arXiv:2109.02496*, 2021.

이 민 섭
https://orcid.org/0009-0000-5228-9545
e-mail : ms_lee@korea.ac.kr
2020년 동국대학교 컴퓨터공학과(학사)
2020년 ~ 현 재 고려대학교 정보보호대학원 Ph.D., 석·박사통합과정
관심분야 : Privacy-Enhancing Technology & Blockchain & Federated Learning

신 영 아
https://orcid.org/0000-0001-7969-7143
e-mail : yashin95@korea.ac.kr
2020년 성신여자대학교 융합보안학과(학사)
2020년 ~ 현 재 고려대학교 정보보호대학원 Ph.D., 석·박사통합과정
관심분야 : Privacy-Enhancing Technology & Blockchain & Federated Learning

천 지 영
https://orcid.org/0000-0002-5329-8918
e-mail : jychun@iscu.ac.kr
2011년 고려대학교 정보경영공학과(박사)
2021년 ~ 현 재 서울사이버대학교 빅데이터·정보보호학과 조교수
2022년 ~ 현 재 서울사이버대학교 빅데이터·AI센터 부센터장
2023년 ~ 현 재 서울사이버대학교 AI융합대학 학장
관심분야 : Data Privacy & Artificial Intelligence & Federated Learning & Privacy-Enhancing Technology