

CID 기반 참조 및 검증 구조를 활용한 연합학습 모델 데이터 관리 메커니즘 제안

김진수*, 박남제**

A Data Management Mechanism for Federated Learning Models using a CID-based Reference and Verification Structure

Jinsu Kim*, Namje Park**

이 논문은 2025년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2025S1A55B5A160080)

요 약

본 논문에서는 스마트 온실 환경에서의 연합학습 기반 데이터 관리 구조를 대상으로 데이터 무결성 강화를 위한 메커니즘을 제안한다. 제안 메커니즘은 모델 가중치를 직접 전달하지 않고 CID를 기반으로 참조하는 구조를 적용하며, IPFS를 통해 모델 가중치를 저장하고 블록체인을 통해 CID와 검증 정보를 관리하는 방식으로 구성된다. 이를 통해 분산 환경에서 생성되는 모델 데이터의 저장과 참조를 분리하고, 데이터 관리 흐름을 재구성할 수 있다. 또한 제안 방식은 모델 데이터의 위치와 상태를 식별할 수 있도록 하며, 분산된 환경에서 생성되는 모델 정보의 무결성을 유지하기 위한 구조를 제공한다. 더불어 데이터 공유 없이도 다양한 환경에서 생성된 정보를 반영할 수 있는 학습 환경을 고려한다. 제안 방식은 저장 비용을 일정하게 유지하는 반면 조회 지연은 증가하는 특성을 보이고, 동시 요청 증가 상황에서도 일관된 성능 변화를 유지하는 것으로 나타났다.

Abstract

This paper proposes a mechanism for enhancing data integrity in federated learning-based data management for smart greenhouse environments. The proposed mechanism adopts a Content Identifier (CID)-based reference structure, where model weights are not directly transmitted but stored in IPFS and managed through blockchain with associated CID and verification information. This approach separates model storage and reference, enabling reconstruction of the data management process in distributed environments. In addition, the proposed method provides a structure to identify the location and state of model data while maintaining data integrity of model information generated across distributed nodes. Furthermore, it supports learning across diverse environments without direct data sharing. The proposed method maintains a constant storage cost while exhibiting increased retrieval latency, and demonstrates consistent performance behavior even under increasing concurrent requests.

Keywords

federated learning, blockchain, IPFS, CID, smart greenhouse

* 제주대학교 사이버보안인재교육원 연구원

- ORCID: <https://orcid.org/0000-0003-1009-3928>

** 제주대학교 초등컴퓨터교육학과 교수(교신저자)

- ORCID: <https://orcid.org/0000-0003-4434-8933>

• Received: Mar. 27, 2026, Revised: Aug. 05, 2026, Accepted: Aug. 08, 2026

• Corresponding Author: Namje Park

Dept. of Computer Education, Teachers College, Jeju National University, 61 Iljudong-ro, Jeju-si, Jeju Special Self-Governing Province, 63294, Korea

Tel.: +82-64-754-4914, Email: namjepark@jejunu.ac.kr

I. 서 론

스마트 온실 환경에서는 온도, 습도, 일사량과 같은 다양한 환경 데이터를 기반으로 난방 에너지 소비를 효율적으로 관리하는 것이 중요하다[1]. 특히 난방 에너지는 운영 비용과 직결되기 때문에 이를 사전에 예측하고 제어하는 기술에 대한 요구가 증가하고 있으며, 이를 위해 머신러닝 기반의 예측 모델이 활용되고 있다. 그러나 개별 온실에서 수집되는 데이터는 특정 환경 조건에 국한되는 한계를 가지며, 지역별 기후와 시설 구조에 따라 데이터 분포가 달라 단일 온실에서 학습된 모델은 다양한 환경을 충분히 반영하기 어렵다[2][3].

이러한 한계를 보완하기 위해 연합학습 방식이 고려될 수 있다. 연합학습은 온실에서 생성된 데이터를 직접 공유하지 않고 모델을 공동으로 학습할 수 있으며, 글로벌 모델을 통해 다양한 환경 조건을 반영할 수 있다[4]-[6]. 그러나 모델 가중치를 직접 전달하는 구조는 데이터 전송 부담을 증가시키며, 다수의 클라이언트가 참여하는 환경에서는 통신 부하와 중앙 서버 의존성이 증가하는 문제가 존재한다[7]-[9]. 또한 분산 환경에서는 각 노드에서 데이터의 신뢰성을 특정 주체에 의존하지 않고 검증할 필요가 있으며, 중앙 서버 구조는 단일 장애 지점과 신뢰 비대칭 문제를 야기할 수 있다. 이에 따라 데이터 기록과 검증을 분산된 방식으로 수행할 수 있는 구조와 모델 가중치의 변경 여부를 검증할 수 있는 관리 방식이 요구된다[10]. 이러한 요구사항은 특정 주체에 의존하지 않고 데이터 기록과 검증을 수행할 수 있는 구조를 필요로 하며, 이를 위해 블록체인을 활용한 분산 기록 방식이 고려될 수 있다.

본 논문에서는 모델 가중치를 직접 전달하지 않고 CID 기반 참조 구조를 적용한 데이터 관리 메커니즘을 제안한다. 제안 메커니즘은 IPFS를 이용하여 모델 가중치를 분산 저장하고, 블록체인을 통해 CID와 검증 정보를 함께 관리한다. 이를 통해 데이터 저장과 참조를 분리하고 모델 가중치의 변경 여부를 검증할 수 있으며, 특정 주체에 의존하지 않는 데이터 관리와 연합학습 환경에서의 일관된 학습 과정을 동시에 고려할 수 있다.

II. 관련 연구 분석

2.1 연합학습 기반 에너지 예측 연구 분석

에너지 소비 예측은 스마트 온실 및 스마트 그리드 환경에서 효율적인 자원 관리와 밀접하게 연관된 문제로, 다양한 머신러닝 기반 접근 방식이 제안되고 있다. 특히 데이터의 분산성과 프라이버시 보호 요구가 증가함에 따라, 데이터를 직접 공유하지 않고도 모델을 학습할 수 있는 연합학습 기반 접근 방식이 점차 주목받고 있다[11][12].

Sarmiento et al.[13]의 연구에서는 CNN과 LSTM을 결합한 연합학습 구조를 적용하여 분산된 환경에서 에너지 소비를 예측하는 방법을 제안하였으며, 다양한 에너지 데이터셋을 활용한 실험을 통해 중앙집중형 모델과 비교 가능한 성능을 보일 수 있음을 확인하였다. 또한 Mendes et al.[14]의 연구에서는 건물 단위의 에너지 수요 예측을 위해 연합학습 기반 프레임워크를 제안하였으며, 데이터 공유 없이도 협력적인 예측이 가능함을 보였다. 한편 Wang et al.[15]의 연구에서는 건물 에너지 데이터를 대상으로 개인화 연합학습 구조를 적용하여 데이터 이질성 문제를 완화하고 예측 정확도를 향상시키는 방법을 제시하였다.

이와 같은 연구들은 연합학습이 에너지 예측 문제에 효과적으로 적용될 수 있음을 보여주고 있으나, 학습 과정에서 생성되는 모델 가중치의 관리 및 신뢰성 확보에 대한 고려는 상대적으로 제한적이다. 특히 분산 환경에서 생성된 모델 정보의 무결성을 보장하고, 이를 효율적으로 공유하기 위한 구조에 대한 추가적인 연구가 필요하다.

2.2 연합학습 모델 무결성 연구 분석

연합학습은 데이터를 직접 공유하지 않고 모델을 학습할 수 있다는 장점을 가지지만, 학습 과정에서 생성되는 모델 가중치의 신뢰성과 무결성에 대한 문제가 지속적으로 제기되고 있다[16]-[18]. 특히 분산된 환경에서는 각 클라이언트가 생성한 모델 업데이트의 정당성을 검증하기 어렵기 때문에, 악의적

인 업데이트나 비정상적인 데이터가 전체 모델에 영향을 미칠 수 있다.

Bonawitz et al.[19]의 연구에서는 클라이언트의 모델 업데이트를 보호하기 위한 secure aggregation 기법을 제안하여, 개별 업데이트를 보호하면서도 집계 가능하도록 하였다. 또한 Kairouz et al.[20]은 연합학습 환경에서 발생할 수 있는 다양한 공격 시나리오와 보안 문제를 체계적으로 분석하고, 데이터 유출 및 모델 오염에 대한 취약성을 지적하였다. 한편 Li et al.[21]은 비정상적인 모델 업데이트를 탐지하기 위한 방법을 제안하여, 학습 과정에서의 신뢰성 확보를 위한 접근 방식을 제시하였다.

이와 같은 연구들은 연합학습 환경에서의 보안 문제를 완화하기 위한 다양한 방법을 제안하고 있으나, 모델 가중치의 저장 위치와 참조 구조를 포함한 데이터 관리 방식에 대한 고려는 상대적으로 제한적이다. 특히 모델 가중치의 무결성을 검증하면서도, 이를 효율적으로 공유할 수 있는 구조에 대한 연구는 부족한 실정이다.

2.3 블록체인/IPFS 기반 분산 데이터 관리 분석

블록체인과 분산 저장 기술은 데이터의 무결성과 신뢰성을 확보하기 위한 수단으로 다양한 분야에서 활용되고 있다[22][23]. 특히 연합학습 환경에서는 중앙 서버에 대한 의존성을 줄이고, 데이터의 변경 이력을 관리하기 위한 방법으로 블록체인 기반 접근 방식이 제안되고 있다. 또한 IPFS와 같은 분산 파일 시스템은 대용량 데이터를 효율적으로 저장하고 공유할 수 있는 구조를 제공한다.

Kim et al.[24]의 연구에서는 블록체인을 활용하여 학습 환경에서 학습 데이터를 안전하게 관리하는 구조를 제안하였으며, 데이터의 위변조를 방지하기 위한 분산형 학습 환경을 구성하였다. 또한 Kumar et al.[25]은 블록체인과 연합학습을 결합하여 분산 환경에서 데이터의 신뢰성과 협력 학습 구조를 강화하는 방안을 제안하였으며, 글로벌 모델이 블록체인을 통해 공유되는 구조를 제시하였다. 한편 Benet[26]은 IPFS를 제안하여 분산 환경에서 데이터의 저장 및 공유를 효율적으로 수행할 수 있는 구조를 정의하였으며, 콘텐츠 기반 주소 체계를 통해

데이터 무결성 검증이 가능함을 보였다. 기존 연구는 블록체인과 IPFS를 활용한 데이터 관리에 초점을 두고 있으나, 저장과 참조를 분리한 구조적 접근은 제한적이다. 표 1은 관련 연구와 제안 메커니즘 간의 구조적 차이를 요약하여 나타낸다.

표 1. 관련 연구와 제안 메커니즘의 비교 (O: 지원, X: 미지원)

Table. 1. Comparison of related works and proposed mechanism (O: Supported, X: Not supported)

Feature	[24]	[25]	[26]	Proposed
Blockchain-based management	O	O	X	O
IPFS-based storage	X	X	O	O
CID-based reference	X	X	O	O
Separation of storage and reference	X	X	△	O
Integrated data management	X	X	X	O

III. 연합학습 기반 분산 네트워크 환경에서의 온실 에너지 데이터 관리 메커니즘 제안

본 논문에서는 스마트 온실에서 요구되는 온실 데이터를 분산 저장하고 데이터의 변경 여부를 검증할 수 있는 관리 구조를 제공하기 위한 메커니즘을 제안한다. 제안 메커니즘은 크게 클라이언트, IPFS, 블록체인, 집계 서버의 4가지 구성요소로 구성된다. 본 논문에서 다루는 무결성은 모델 데이터의 변경 여부를 검증하는 수준의 데이터 동일성 확인에 초점을 둔다. 즉, 저장된 모델 가중치가 전송 및 저장 과정에서 변경되지 않았는지를 검증하는 것을 의미하며, 해당 데이터가 정상적인 학습 결과인지 또는 품질적으로 유효한지를 판단하는 수준까지는 포함하지 않는다. 이러한 범위는 CID 기반 참조 구조와 해시 검증을 통해 구현되며, 데이터의 생성 과정 자체에 대한 신뢰성 검증과는 구분된다.

클라이언트는 종단에 위치하는 온실을 의미하며, 온실에서 측정되는 데이터를 기반으로 학습을 수행하며, 집계 서버에 의해 생성된 글로벌 모델을 반영하는 과정을 수행한다. 제안 메커니즘은 모델 가중치를 직접 전달하거나 단순히 CID를 기록하는 방식과 달리, 데이터의 저장과 참조, 그리고 검증 정보를 결합하여 관리하는 구조를 가진다. IPFS는 학습

가중치의 특성상 트랜잭션 내에 포함되기 어려운 데이터를 외부에서 관리하기 위한 분산 저장소로 사용되며, 단순 저장 기능을 넘어 모델 데이터와 연계된 관리 구조의 일부로 활용된다. 클라이언트와 집계 서버에서 생성된 학습 모델 가중치는 IPFS에 기록되며, 해당 데이터에 대한 CID는 블록체인에 기록된다. 이때 CID는 단순한 데이터 식별 정보로 사용되는 것을 넘어, 검증 해시 및 기록 정보와 함께 관리되어 모델 데이터의 상태와 생성 이력을 함께 추적할 수 있도록 구성된다. 블록체인은 학습에 요구되는 가중치의 CID와 검증에 요구되는 해시, 기록자의 식별정보 등의 데이터를 기록하며, 데이터 위치를 저장하는 역할뿐만 아니라 모델 생성 및 전달 과정에 대한 검증 정보를 함께 관리하는 구조를 형성한다.

기록된 정보는 클라이언트와 집계 서버가 학습 가중치의 정보를 얻기 위한 정보로 활용된다. 집계 서버는 블록체인에 기록되어 있는 CID를 기반으로 IPFS에서 실제 학습 모델 가중치 정보를 획득하고, 획득한 가중치 정보를 집계하여 글로벌 모델을 생성한다. 생성된 글로벌 모델은 IPFS에 저장하며, 클라이언트와 동일하게 식별정보, CID, 검증 해시를

포함하여 블록체인에 기록한다. 이러한 구조는 단순한 데이터 참조를 넘어, 모델 데이터의 생성, 저장, 검증 및 활용까지의 전체 흐름을 통합적으로 관리할 수 있도록 한다.

또한 제안 구조는 데이터 자체가 아닌 CID 기반 참조를 중심으로 구성되므로 데이터 저장과 조회가 특정 노드에 종속되지 않고 분산 환경에서 독립적으로 수행될 수 있으며, IPFS 기반 분산 저장을 통해 노드 의존성을 줄이고 데이터 접근의 유연성을 확보할 수 있다. 더불어 블록체인에 기록되는 CID 및 검증 정보는 모든 참여 노드가 동일하게 참조할 수 있는 기준으로 작용하여 분산 환경에서도 일관된 데이터 관리가 가능하도록 한다.

그림 1은 제안 메커니즘의 전반적인 개요를 나타낸다. 클라이언트에서는 블록체인에 저장되어 있는 집계 서버 가중치의 CID를 기반으로 IPFS에서 글로벌 모델을 받아 반영하고 학습을 수행한다. 이후 생성된 로컬 가중치는 IPFS에 저장하며, 접근을 위한 CID를 블록체인에 식별정보와 검증해시를 포함하여 기록한다. 집계 서버는 블록체인 내의 CID를 기반으로 로컬 가중치를 획득하여 집계를 통해 글로벌 모델을 생성한다.

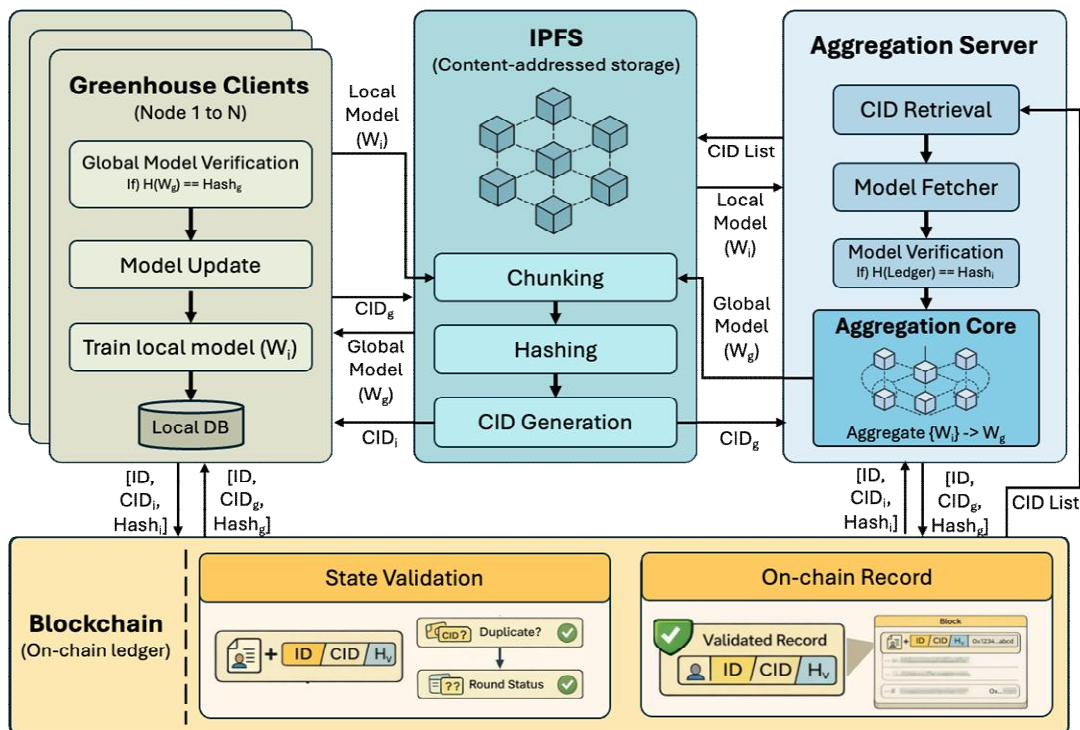


그림 1. 블록체인 및 IPFS 기반 연합학습 모델 무결성 검증 개념도

Fig. 1. Conceptual framework of blockchain and IPFS-based federated learning model integrity verification

본 논문에서 고려하는 위협은 모델 가중치의 변조 및 위·변조된 데이터의 전달과 같은 데이터 무결성 위협에 초점을 둔다. 특히 전송 과정에서 모델 가중치가 변경되거나, 악의적인 노드가 변조된 데이터를 공유하는 상황을 가정한다. 제안 구조는 IPFS에 저장된 데이터와 블록체인에 나저 기록된 CID 및 검증 해시를 비교함으로써 데이터의 변경 여부를 확인할 수 있도록 하며, 이를 통해 전송 및 저장 과정에서 발생할 수 있는 데이터 변조를 탐지할 수 있다. 다만 본 논문은 데이터의 품질이나 학습 결과의 적합성에 대한 검증까지는 포함하지 않는다.

3.1 약어 정리

표 2는 제안 메커니즘에서 사용되는 기호와 약어를 정리한 것이다.

표 2. 기호/약어 정리

Table. 2. Notation and Abbreviations

Abbrev.	Definition
$W_i^{(t)}$	Local model weights generated by client at round t
$W_g(t)$	Global model weights generated at round t
$CID_i(t)$	Content identifier corresponding to the local model
$CID_g(t)$	Content identifier corresponding to the global model
ID_i	Unique identifier assigned to each client
Round _t	Identifier representing the training round
$C(\cdot)$	Operation that converts model weights into a content identifier

3.2 학습 모델의 데이터 흐름 및 생성 과정

그림 2는 제안 메커니즘에서 구성요소 간의 데이터 전개 과정을 나타낸다.

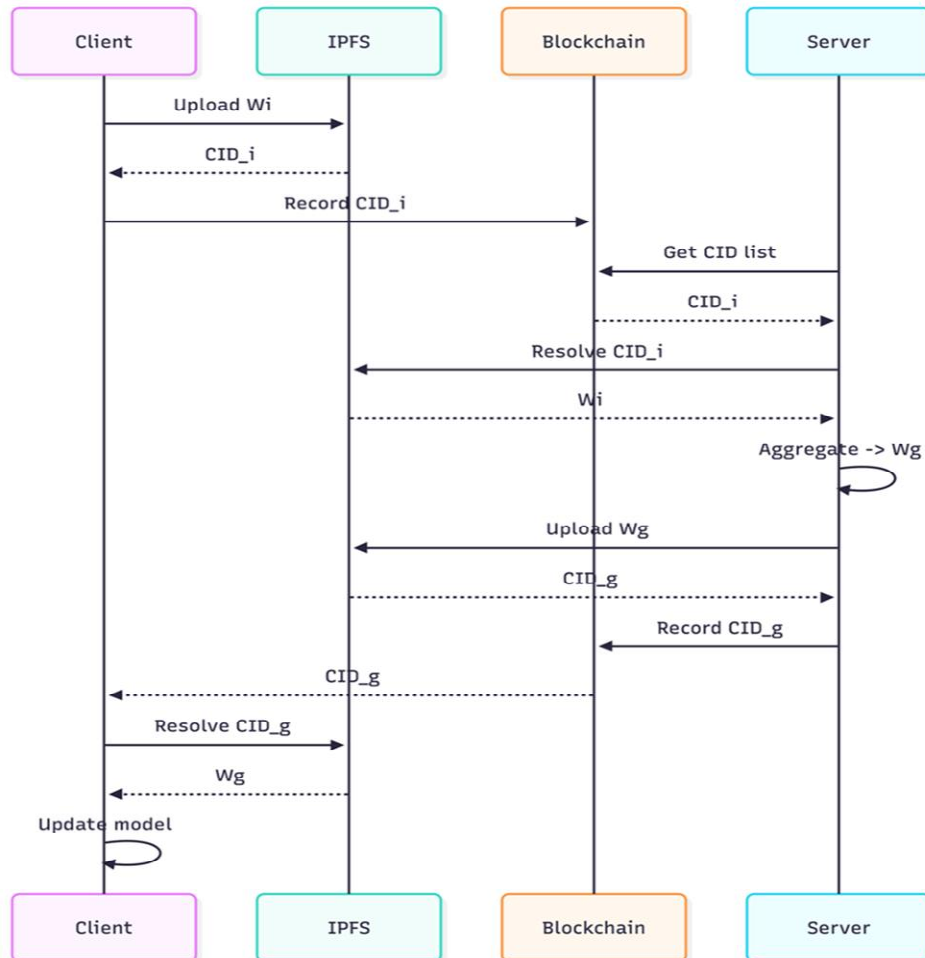


그림 2. 블록체인 및 IPFS 기반 연합학습 모델 처리 흐름도

Fig. 2. Process Flow of Blockchain and IPFS-based Federated Learning Model Handling

클라이언트는 로컬 모델 가중치를 생성하며, 생성된 가중치는 IPFS에 기록한다. IPFS는 클라이언트로부터 받은 로컬 모델 가중치에 대한 CID를 제공한다. 클라이언트는 CID를 포함하여 학습 가중치를 기록하기 위한 트랜잭션을 생성하며, 트랜잭션의 구조는 식 (1)과 같이 전개된다.

$$\begin{aligned} Transaction_i &= [ID_i, t, CID_i^{(t)}, Hash(W_i^{(t)})] \\ CID_i^{(t)} &= C(W_i^{(t)}) \end{aligned} \quad (1)$$

집계 서버에서는 글로벌 모델을 생성하기 위해 블록체인 내의 ID_i 를 사용하여 동일 라운드(t) 내의 CID_i 를 획득한다. 집계 서버에서는 수집된 CID를 통해 식 (2)와 같이 글로벌 모델을 생성한다.

$$W_g^{(t)} = aggregate(W_1^{(t)}, W_2^{(t)}, \dots, W_n^{(t)}) \quad (2)$$

글로벌 모델을 생성한 집계 서버는 IPFS에 글로벌 모델을 기록하여 생성된 CID와 자신의 식별정보, 검증 정보, 라운드 정보를 포함하여 블록체인에 식 (3)과 같이 기록한다.

$$\begin{aligned} Transaction_g &= [ID_g, t, CID_g^{(t)}, Hash(W_g^{(t)})] \\ CID_g^{(t)} &= C(W_g^{(t)}) \end{aligned} \quad (3)$$

클라이언트에서는 블록체인에 기록된 집계 서버 트랜잭션을 통해 식별정보, 라운드, CID, 검증 해시를 획득하여 검증하고, 로컬 모델에 적용한다.

3.3 블록체인 기반 모델 관리 및 검증 구조

제안 메커니즘에서 블록체인은 모델 가중치 자체를 저장하지 않고, CID와 검증 정보를 기반으로 모델을 관리하는 역할을 수행한다. 블록체인에 기록되는 정보는 모델의 저장 위치를 식별하기 위한 참조정보로 활용되며, 학습 과정에서 생성되는 데이터의 일관성을 유지하기 위한 기준으로 사용된다.

블록체인에 기록되는 데이터는 식 (4)와 같은 구조를 가진다.

$$[Enc(ID), \partial, CID^{(t)}, Hash(W^{(t)})] \quad (4)$$

여기서 $Enc(ID)$ 는 참여 노드의 식별정보를 암호화한 값을 의미한다.

식별 정보의 보호를 위해 클라이언트와 집계 서버는 서로 다른 방식으로 암호화를 수행한다. 클라이언트는 자신의 식별 정보를 집계 서버의 공개키를 이용하여 암호화한 후 블록체인에 기록하며, 이를 통해 외부에서는 식별 정보를 확인할 수 없도록 한다. 반면 집계 서버는 자신의 식별 정보를 비밀키를 기반으로 생성하며, 클라이언트는 집계 서버의 공개키를 이용하여 해당 정보를 검증할 수 있다.

블록체인은 입력되는 CID에 대해 상태 기반 검증을 수행한다. 동일 학습 라운드 내에서 중복된 CID가 기록되는 것을 방지하며, 각 CID가 현재 라운드에 해당하는지 여부를 확인한다. 또한 기록되는 데이터가 정의된 구조를 만족하는지를 검증하여, 비정상적인 정보가 기록되는 것을 방지한다.

이와 같은 검증 과정은 모델 가중치 자체를 직접 검증하지 않더라도, 기록되는 정보의 신뢰성을 유지할 수 있도록 한다. 이후 집계 서버와 클라이언트는 블록체인에 기록된 CID를 기준으로 모델을 참조하게 되며, 이를 통해 전체 학습 과정이 수행된다.

IV. 제안 메커니즘의 구조적 성능 분석

본 장에서는 제안 메커니즘의 특성을 분석하기 위해 저장 비용, 조회 지연, 그리고 확장성 측면에서 실험을 수행하였다. 본 실험의 목적은 절대적인 성능을 평가하는 것이 아니라, 데이터 관리 구조에 따른 성능 변화의 경향을 분석하는 데 있다. 직접 저장 방식은 데이터를 직접 저장하고 조회하는 구조를 가지는 반면, 제안 방식은 데이터 대신 CID를 통해 참조하는 구조를 가진다. 이러한 차이를 기반으로 저장 비용, 조회 지연, 그리고 동시 요청 증가에 따른 성능 변화를 비교하였다. 본 실험은 단일 환경에서 수행되었으며, 실제 분산 네트워크의 특성을 완전히 반영하지는 않는다. 그러나 동일 조건에서의 비교를 통해 제안 메커니즘의 상대적 특성과 동작 경향을 분석할 수 있다.

4.1 저장 비용 분석

그림 3은 모델 크기에 따른 저장 비용을 직접 저장 방식과 제안 방식에서 비교한 결과를 나타낸다. 직접 저장 방식은 데이터를 블록체인에 직접 기록하는 구조를 가지므로, 모델 크기가 증가함에 따라 저장 비용이 선형적으로 증가하는 경향을 보인다. 반면 제안 방식은 데이터 대신 CID와 메타데이터만을 기록하는 구조를 가지므로, 모델 크기와 관계없이 저장 비용이 거의 일정하게 유지되는 것을 확인할 수 있다. 이러한 차이는 로그 스케일에서도 명확하게 나타나며, 두 방식 간 저장 구조의 근본적인 차이를 보여준다. 이러한 차이는 데이터 저장과 참조를 분리한 구조적 특성에 기인하며, 분산 환경에서도 온체인에 기록되는 데이터 크기를 일정하게 유지할 수 있다는 점에서 의미를 가진다.

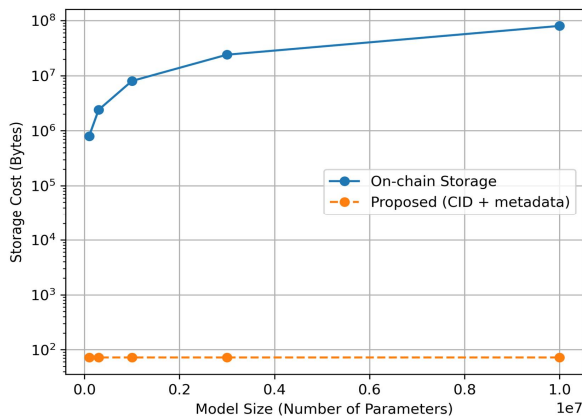


그림 3. 데이터 크기에 따른 저장 비용 비교
Fig. 3. Storage cost comparison according to data size

4.2 조회 지연 분석

그림 4는 모델 크기에 따른 조회 지연을 직접 저장 방식과 제안 방식에서 비교한 결과를 나타낸다. 직접 저장 방식은 로컬에 저장된 데이터를 즉시 읽어오는 구조를 가지므로 전반적으로 매우 낮은 지연을 보인다. 반면 제안 방식은 CID를 기반으로 외부 저장소에서 데이터를 조회하는 과정을 포함하므로 추가적인 탐색 및 전송 시간이 발생한다. 이로 인해 모든 구간에서 직접 저장 방식에 비해 높은 지연이 나타나며, 특히 모델 크기가 증가함에 따라

지연 증가 폭이 커지는 경향을 확인할 수 있다. 또한 제안 방식은 반복 측정 결과에서 일정 수준의 변동성을 보이며, 이는 데이터 조회 과정에서의 처리 지연 및 시스템 자원 상태의 영향을 받기 때문으로 해석된다. 이와 같은 지연 특성은 분산 저장 구조에서의 데이터 접근 방식과 밀접하게 연관되며, 외부 저장소를 참조하는 구조에서 발생하는 자연스러운 결과로 해석될 수 있다. 즉, 저장 효율성과 조회 지연 간의 상충 관계를 반영하는 구조적 특성으로 볼 수 있다.

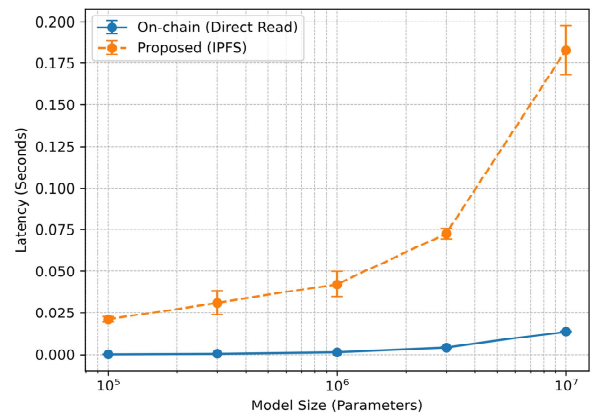


그림 4. 데이터 조회 방식에 따른 지연 시간 비교
Fig. 4. Latency comparison for data retrieval methods

4.3 확장성 분석

그림 5는 동시 요청 수 증가에 따른 조회 지연 변화를 나타낸다. 동시 요청 수가 증가함에 따라 전체 지연 시간은 점진적으로 증가하는 경향을 보인다. 이는 다수의 요청이 동시에 처리되면서 시스템 자원의 경쟁이 발생하기 때문으로 해석된다. 특히 일정 수준 이상의 요청이 발생하는 경우 지연 증가 폭이 커지는 구간이 관찰된다. 그러나 전체적인 증가 양상은 일관된 패턴을 유지하며, 요청 수 증가에 따라 지연이 급격히 증가하거나 시스템이 불안정해지는 현상은 나타나지 않았다. 이는 제안 방식이 부하가 증가하는 상황에서도 안정적으로 동작할 수 있음을 보여준다. 이와 함께 제안 방식은 블록체인을 활용하여 CID 및 검증 정보를 관리하는 구조를 가지므로, 트랜잭션 생성 및 검증 과정에서 일정 수준의 처리 지연이 발생할 수 있다. 다만 모델 가중

치 자체를 기록하지 않고 참조 정보를 중심으로 관리함으로써, 온체인 데이터 처리 부담을 제한하는 구조를 가진다. 이러한 특성은 분산 환경에서의 데이터 처리 과정에 영향을 미칠 수 있으며, 구조적 관점에서의 확장성을 유지하는 데 기여한다.

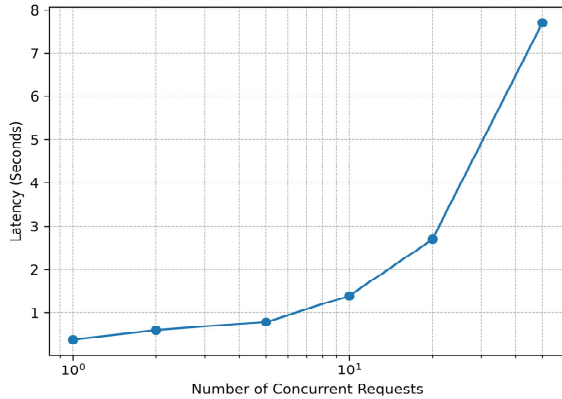


그림 5. 동시 요청 수 증가에 따른 지연 시간 변화
Fig. 5. Latency variation with increasing concurrent requests

V. 결 론

본 논문에서는 스마트 온실 환경에서의 에너지 데이터를 분산 저장하고, 연합학습 과정에서 데이터의 무결성을 강화하는 메커니즘을 제안하였다. 제안 메커니즘은 모델 가중치를 직접 전달하지 않고 CID를 기반으로 참조하는 방식을 적용하여, 분산 환경에서의 데이터 관리 흐름을 재구성하였다. 또한 IPFS를 통해 모델 가중치를 저장하고, 블록체인을 통해 CID와 검증 정보를 기록함으로써 모델 정보의 저장과 참조가 분리된 형태를 정의하였다.

제안 방식은 모델 가중치의 직접적인 전송 없이도 학습이 가능하도록 하며, CID 기반 참조를 통해 데이터의 위치와 상태를 명확하게 식별할 수 있는 특징을 가진다. 또한 블록체인을 활용한 기록 방식을 통해 분산 환경에서의 데이터 흐름을 일관되게 관리할 수 있으며, 모델 정보의 변경 여부를 검증할 수 있는 기반을 제공한다. 이러한 특성은 분산된 환경에서의 모델 관리와 학습 과정의 흐름을 보다 명확하게 구성할 수 있도록 한다.

또한 제안 메커니즘의 특성을 분석하기 위해 저장 비용, 조회 지연 및 확장성 측면에서 실험을 수

행하였으며, 데이터 크기 증가와 관계없이 저장 비용을 일정하게 유지하는 반면 조회 과정에서는 추가적인 지연이 발생하는 특성을 확인하였다. 또한 동시 요청 수 증가에 따라 지연은 점진적으로 증가하지만 일관된 성능 변화를 유지하는 것을 확인하였다.

한편, 제안 방식은 실제 환경에서의 적용을 고려할 때 추가적인 검토가 필요한 부분이 존재한다. 제안 메커니즘은 스마트 온실과 같이 다수의 독립적인 운영 주체가 존재하는 환경에서, 데이터 공유 없이도 학습 정보를 반영하면서 모델 가중치의 변경 여부를 검증할 수 있는 구조로 활용될 수 있다. 이를 통해 분산 환경에서의 데이터 관리 일관성을 확보하고, 특정 주체에 의존하지 않는 학습 구조를 구성할 수 있는 가능성을 가진다.

또한 본 논문에서 제안한 방식은 모델 가중치의 변경 여부를 검증하는 수준에 초점을 두고 있어, 모델의 품질이나 학습 결과의 적합성을 보장하는 수준까지는 포함하지 않는다. 더불어 실제 운용 환경에서의 통신 비용, 블록체인 처리 오버헤드 및 다양한 환경 조건에 따른 성능 변화에 대한 정량적 분석은 추가적인 연구가 필요하다.

References

- [1] A. Soussi, E. Zero, A. Ouammi, D. Zejli, S. Zahmoun, and R. Sacile, "Smart greenhouse farming: a review towards near zero energy consumption", *Discover Cities*, Vol. 2, Art. no. 55, May 2025. <https://doi.org/10.1007/s44327-025-00096-w>.
- [2] D. Katzin, E. J. van Henten, and S. van Mourik, "Process-based greenhouse climate models: Genealogy, current status, and future directions", *Agricultural Systems*, Vol. 198, Art. no. 103388, Feb. 2022. <https://doi.org/10.1016/j.agsy.2022.103388>.
- [3] J. Kim and N. Park, "De-Identification Mechanism of User Data in Video Systems According to Risk Level for Preventing Leakage of Personal Healthcare Information", *Sensors*, Vol. 22, No. 7, Art. no. 2589, Mar. 2022. <https://doi.org/10.3390/s22072589>

- [4] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. Arcas, "Communication-efficient learning of deep networks from decentralized data", Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS), Fort Lauderdale, Florida, USA, pp. 1273-1282, May 2017. <https://doi.org/10.48550/arXiv.1602.05629>.
- [5] S.G. Lee and S.K. Park, "Threat Detection and Gateway Security Design for BLE-based IoT Environments", Journal of Convergence Science, Technology, and Society, Vol. 5, No. 1, pp. 63-71, Jun. 2026. <https://doi.org/10.56366/jcsts.2026.5.1.63>.
- [6] P. Kim and N. Park, "Development of Information Security Unplugged for DDoS Core Principles for Elementary School Students", Journal of Convergence Science, Technology, and Society, Vol. 3, No. 1, pp. 25-30, Jun. 2024. <https://doi.org/10.56366/jcsts.2024.3.1.25>.
- [7] J. Konečný, H. B. McMahan, D. Ramage, and P. Richtárik, "Federated optimization: Distributed machine learning for on-device intelligence", arXiv preprint, arXiv:1610.02527, Oct. 2016. <https://doi.org/10.48550/arXiv.1610.02527>.
- [8] J. Kim, and N. Park, "Role-based Access Control Video Surveillance Mechanism Modeling in Smart Contract Environment", Transactions on Emerging Telecommunications Technologies, Vol. 33, No. 4, e4227, Feb. 2022. <https://doi.org/10.1002/ett.4227>
- [9] J. Konečný, H. B. McMahan, F. X. Yu, P. Richtárik, A. T. Suresh, and D. Bacon, "Federated learning: Strategies for improving communication efficiency", arXiv preprint, arXiv:1610.02527, Oct. 2016. <https://doi.org/10.48550/arXiv.1610.05492>.
- [10] N. H. Tran, W. Bao, A. Zomaya, M. N. Nguyen, and C. S. Hong, "Federated learning over wireless networks: Optimization model design and analysis", IEEE INFOCOM 2019 - IEEE Conference on Computer Communications, Paris, France, pp. 1387-1395, Jun. 2019. <https://doi.org/10.1109/INFOCOM.2019.8737464>.
- [11] Q. Yang, Y. Liu, T. Chen, and Y. Tong, "Federated machine learning: Concept and applications", ACM Transactions on Intelligent Systems and Technology, Vol. 10, No. 2, pp. 1-19, Jan. 2019. <https://doi.org/10.1145/3298981>.
- [12] J. Kim, N. Park, G. Kim, and S. Jin, "CCTV Video Processing Metadata Security Scheme Using Character Order Preserving-Transformation in the Emerging Multimedia", Electronics, Vol. 8, No. 4, Apr. 2019. <https://doi.org/10.3390/electronics8040412>.
- [13] E. M. de M. Sarmiento, I. F. Ribeiro, P. R. N. Marciano, Y. G. Neris, H. R. de O. Rocha, V. F. S. Mota, and R. da S. Villaça, "Forecasting energy power consumption using federated learning in edge computing devices", Internet of Things, Vol. 25, pp. 1-15, Apr. 2024. <https://doi.org/10.1016/j.iot.2023.101050>.
- [14] N. Mendes, J. Mendes, J. Mohammadi, and P. Moura, "Federated learning framework for prediction of net energy demand in transactive energy communities", Sustainable Energy, Grids and Networks, Vol. 40, pp. 1-13, Dec. 2024. <https://doi.org/10.1016/j.segan.2024.101522>.
- [15] R. Wang, L. Bai, R. Rayhana, and Z. Liu, "Personalized federated learning for buildings energy consumption forecasting", Energy and Buildings, Vol. 323, pp. 1-14, Nov. 2024. <https://doi.org/10.1016/j.enbuild.2024.114762>.
- [16] E. Bagdasaryan, A. Veit, Y. Hua, D. Estrin, and V. Shmatikov, "How to backdoor federated learning", Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics, Online, Vol. 108, pp. 2938-2948, Aug. 2020.
- [17] W. Jung, "Study on Automated Testing Methods Using AI MCP: Focused on Implementation and Application", Journal of Convergence Science,

- Technology, and Society, Vol. 4, No. 1, pp. 63-69, Jun. 2025. <https://doi.org/10.56366/jcsts.2025.4.1.63>.
- [18] J. Kim, E. Choi, and N. Park, "A Proposal for a Mobility-Control Data Transfer Mechanism Based on a Block Network Utilizing End-to-End Authentication Data", Mathematics, Vol. 12, No. 13, Art. no. 2073, Jul. 2024. <https://doi.org/10.3390/math12132073>.
- [19] K. Bonawitz, V. Ivanov, B. Kreuter, A. Marcedone, H. B. McMahan, S. Patel, D. Ramage, A. Segal, and K. Seth, "Practical Secure Aggregation for Privacy-Preserving Machine Learning", Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security, Dallas, Texas, USA, pp. 1175-1191, Oct. 2017. <https://doi.org/10.1145/3133956.3133982>.
- [20] P. Kairouz, et al., "Advances and Open Problems in Federated Learning", Found. Trends Mach. Learn., Vol. 14, No. 1-2, pp. 1-210, Jun. 2021. <https://doi.org/10.1561/22000000083>.
- [21] T. Li, A. K. Sahu, A. Talwalkar, and V. Smith, "Federated Learning: Challenges, Methods, and Future Directions", IEEE Signal Processing Magazine, Vol. 37, No. 3, pp. 50-60, May 2020. <https://doi.org/10.1109/MSP.2020.2975749>.
- [22] J. Kim and N. Park, "Suggestions for a Blockchain-based Smart Factory Data Verification Mechanism to Enhance Cloud Data Reliability", Journal of Convergence Science, Technology, and Society, Vol. 2, No. 2, pp. 45-49, Jan. 2023.
- [23] J. Kim, and N. Park, "Blockchain-based Smart Grid De-Identification MetadataTransmission Mechanism", Journal of KIIT, Vol. 22, No. 7, pp. 169-177, Jul. 2024. <https://doi.org/10.14801/jkiit.2024.22.7.169>.
- [24] J. Kim and N. Park, "Blockchain-Based Data-Preserving AI Learning Environment Model for AI Cybersecurity Systems in IoT Service Environments", Applied Sciences, Vol. 10, No. 14, Art. no. 4718, Jul. 2020. <https://doi.org/10.3390/app10144718>.
- [25] R. Kumar, A. A. Khan, J. Kumar, Zakria, N. A. Golilarz, S. Zhang, Y. Ting, C. Zheng, and W. Wang, "Blockchain-Federated-Learning and Deep Learning Models for COVID-19 Detection Using CT Imaging", IEEE sensors journal, Vol. 21, No. 14, pp. 16301-16314, Jul. 2021. <https://doi.org/10.1109/JSEN.2021.3076767>.
- [26] J. Benet, "IPFS - Content Addressed, Versioned, P2P File System", arXiv preprint, arXiv.1407.3561, Jul. 2014. <https://doi.org/10.48550/arXiv.1407.3561>.

저자소개

김 진 수 (Jinsu Kim)



2024년 8월 : 제주대학교
융합정보보안학과(공학박사)
2018년 9월 ~ 현재 : 제주대학교
사이버보안인재교육원 연구원
관심분야 : 클라우드, 지능형
영상감시 시스템, IoT

박 남 제 (Namje Park)



2008년 2월 : 성균관대학교
컴퓨터공학과(공학박사)
2003년 4월 ~ 2008년 12월 : ETRI
정보보호연구단 선임연구원
2009년 1월~ 2010년 8월 : UCLA
Post-Doc., ASU Research
Scientist
2010년 9월 ~ 현재 : 제주대학교 교육대학
초등컴퓨터교육전공 교수,
대학원 융합정보보안협동과정 교수
관심분야 : 융합기술보안, 컴퓨터교육, 스마트그리드, IoT,
해사클라우드