이질적 연합학습을 위한 도커 기반 테스트베드 구축 및 BWA·ADM 기법 성능 검증

임동건§, 방인규■†, 김태훈■§

§국립한밭대학교 컴퓨터공학과, †국립한밭대학교 지능미디어공학과

20211893@edu.hanbat.ac.kr, {ikbang, thkim}@hanbat.ac.kr

Implementation of Docker-based Testbed for Heterogeneous Federated Learning and Performance Evaluation of BWA and ADM Algorithms

Donggeon Im§, Inkyu Bang■†, Taehoon Kim■§

§Department of Computer Engineering, Hanbat National University

†Department of Intelligence Media Engineering, Hanbat National University

요 약

연합학습(Federated Learning)은 프라이버시를 보호하며 분산된 데이터를 학습할 수 있는 기술이나, 참여 클라이언트 간의 시스템 성능 차이(System Heterogeneity)로 인해 전체 학습 속도가 저하되는 문제(Straggler Problem)가 발생한다. 본 연구에서는 이를 해결하기 위해 제안된 BWA(Batch size and Weighted Aggregation)와 ADM(Adaptive Dataset Management) 기법을 시뮬레이션이 아닌 실제 시스템 환경에서 검증하고자 한다. 이를 위해 Docker 컨테이너 가상화 기술을 활용하여 이질적인 자원 제약 환경을 모사한 테스트베드를 구축하였다. 실험 결과, 저사양 클라이언트 환경에서 ADM 기법 적용 시 종래의 대표 기법인 FedAvg 대비 약 21.3%의 학습 시간을 단축하였으며, BWA 기법 또한 2.7%의 정확도 향상과 함께 8.1%의 시간 단축 효과를 보임을 확인하였다.

Ⅰ. 서 론

모바일 기기 및 IoT 센서의 확산으로 방대한 양의 데이터가 엣지(Edge) 단에서 생성되고 있다. 이러한 데이터를 중앙 서버로 전송하지 않고 로컬에서 학습하는 연합학습(Federated Learning, FL)은 데이터 프라이버시 보호 측면에서 큰 주목을 받고 있다 [1].

그러나 실제 연합학습 환경은 고성능 서버부터 저전력 IoT 기기까지 다양한 컴퓨팅 성능을 가진 디바이스들이 참여하므로, 시스템 이질성(System Heterogeneity) 문제가 필연적으로 발생한다 [2]. 특히 연산 성능이나 통신 성능이 낮은 기기(Straggler)로 인해 글로벌 모델 업데이트가 지연되는 병목 현상은 연합학습 시스템 전반의 효율성을 저해하는 핵심적인 문제로 지적된다.

이를 해결하기 위해 선행연구 [3]에서는 BWA 및 ADM 알고리즘을 제안하고, 쿠버네티스(Kubernetes) 기반 환경에서 그 유효성을 검증하였다. 본 연구는 이를 확장하여, Docker 컨테이너 기반의 경량 테스트베드를 구축하고 개별 클라이언트의 CPU 자원을 명시적으로 제어함으로써(--cpus 옵션 활용), 이질적인 연산 성능을 보다 정밀하게 모사하였다. 이를 통해 저사양 기기가 존재하는 이질적 학습 환경을 현실적으로 모사하고, 이러한 환경에서도 선행연구 [3]의 알고리즘들이 학습 지연을 효과적으로 완화하면서 모델 정확도를 유지할 수 있는지를 재검증하고자 한다.

Ⅱ. 도커 기반 테스트베드 설계

본 연구에서는 실제 디바이스의 성능 차이를 모사하기 위해 Docker Compose를 활용하여 다중 컨테이너 환경을 구축하였다. 총 3개의 클라이언트 컨테이너를 생성하였으며, Docker의 deploy.resources.limits 옵션을 사용하여 각 컨테이너의 CPU 및 메모리 자원을 차등 할당하였다. 표 1은 구축된 테스트베드의 자원 할당 내역이다. Client 0은 고성능 디바이스, Client 2는 저사양 디바이스로 설정하였으며, 특히 Client 2의 CPU를 0.5 코어로 제한하여 의도적으로 연산 지연을 유발하고 병목 상황을 구현하였다.

표 1 이질적 클라이언트 환경 구성을 위한 자원 할당 설정

ID	기기 유형	CPU	메모리	비고
0	High-End (고성능)	4.0	8GB	성능 최대
1	Mid-Range (일반)	1.5	1.5GB	일반
2	Low-End (저사양)	0.5	1GB	병목 유도

Ⅲ. 최적화 알고리즘 구현

본 연구에서는 선행 연구[3]를 참고하여 이질성 해결을 위한 두 가지 핵심 알고리즘을 구현하였다.

1. BWA (Batch size and Weighted Aggregation)

데이터 분포의 불균형(Non-IID)으로 인한 모델 성능 저하를 방지하고 학습 효율을 높이기 위한 기법이다. 강화학습(DRL)을 통해 클라이언트의 상태에 따라 최적의 배치 크기(Batch Size)와 집계 가중치를 동적으로 조절함으로써, 글로벌 모델의 정확도를 극대화하고 수렴 속도를 가속화한다.

2. ADM (Adaptive Dataset Management)

시스템 자원이 부족한 클라이언트의 연산 부하를 줄이기 위한 기법이다. 통신 라운드가 진행됨에 따라 학습 기여도가 낮은 데이

■ Corresponding Authors: Inkyu Bang (ikbang@hanbat.ac.kr), Taehoon Kim (thkim@hanbat.ac.kr)

터를 선별하거나, 데이터셋의 크기(v_n) 자체를 적응적으로 축소하여 저사양 기기의 학습 시간을 단축시킨다.

IV. 실험 및 결과

실험은 CIFAR-10 데이터셋을 사용하였으며, 데이터 분포의 이질성을 반영하기 위해 Non-IID 환경(Bias 0.5)으로 설정하였다. 성능 비교를 위해 기본적인 연합학습 알고리즘인 FedAvg를 Baseline으로 설정하였다.

1. ADM 기법 성능 분석

그림 1은 ADM 적용 시 클라이언트별 데이터 사용 비율(v_n)의 변화를 나타낸다. 고성능인 Client 0은 거의 모든 데이터를 활용하는 반면, 자원이 제한된 Client 1과 2는 라운드가 진행됨에 따라 데이터 사용량을 약 74~86% 수준으로 자동 조절함을 확인할 수 있다.

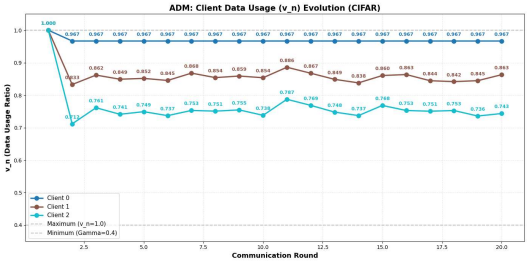

그림 1 ADM 적용 시 데이터 사용량 변화

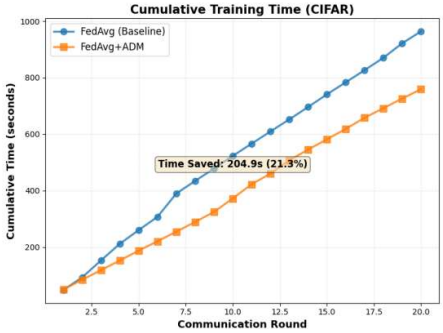

그림 2 ADM 적용 시 누적 학습 시간 비교

그 결과, 그림 2와 같이 ADM은 FedAvg 대비 총 학습 시간을 204.9초 단축하여 약 21.3%의 성능 향상을 달성하였다. 이는 저사양 기기의 연산 부하 감소가 전체 시스템의 대기 시간 단축으로 이어짐을 입증한다. 이때 최종 정확도는 54.0%로 FedAvg(52.8%)와 동등하거나 소폭 우수한 수준을 유지하였다.

2. BWA 기법 성능 분석

BWA 기법 적용 시 배치 크기는 학습 초기 64에서 시작하여 네트워크 상황에 따라 16~32로 동적으로 변화하는 패턴을 보였다.

그림 3에서 확인되는 바와 같이, BWA 기법은 FedAvg 대비 78.1초(8.1%)의 학습 시간 단축 효과를 보였다. 이는 데이터셋 축소를 동반하는 ADM 기법에 비해 시간 단축 폭은 상대적으로 작으나, 원본 데이터셋을 온전히 보존하면서 배치 크기 최적화만으로도 연산 효율을 유의미하게 개선할 수 있음을 시사한다.

또한 정확도 측면에서도 그림 4와 같이 학습 초기부터 안정적인 수렴 곡선을 보였으며, 최종적으로 FedAvg(52.8%) 대비 2.7% 향상된 55.5%의 정확도를 달성하여 속도와 성능을 동시에 개선함을 확인하였다.

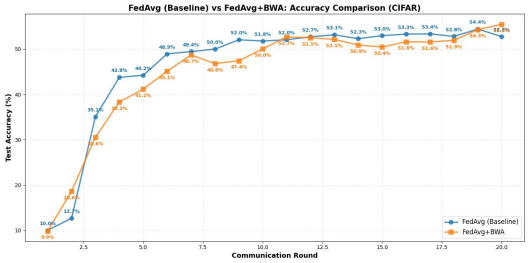

그림 3 BWA 적용 시 누적 학습 시간 비교

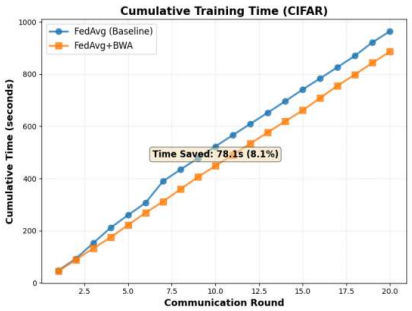

그림 4 BWA 적용 시 정확도 비교

V. 결론

본 연구에서는 연합학습의 이질성 문제를 해결하기 위해 Docker 컨테이너 기술을 활용한 경량화된 테스트베드를 구축하고, 선행 연구[3]에서 제안된 BWA 및 ADM 기법을 구현하여 그 성능을 재검증(Re-verification)하였다. 실험 결과, 구축된 Docker 환경에서 두 기법은 저사양 클라이언트의 병목 현상을 효과적으로 완화하였다. 구체적으로 ADM은 21.3%의 학습 시간 단축을, BWA는 8.1%의 시간 단축과 더불어 2.7%의 정확도 향상을 달성하였다. 본 연구는 선행 연구[3]의 알고리즘이 다양한 가상화 환경(Docker)에서도 여전히 유효함을 실증적으로 확인하고, 자원 제약과 학습 효율 간의 상관관계를 분석했다는 점에서 의의가 있다.

ACKNOWLEDGMENT

본 연구는 2025년도 교육부 및 대전광역시의 재원으로 대전RISE센터의 지원을 받아 수행된 지역혁신중심 대학지원체계(RISE)의 결과(2025-RISE-06-002) 및 대학ICT연구센터사업의 연구결과로 수행되었음(IITP-2024-RS-2024-00437886)

참 고 문 헌

[1] B. McMahan, *et al.*, "Communication-Efficient Learning of Deep Networks from Decentralized Data," in Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS), 2017.

[2] T. Li, A. K. Sahu, M. Zaheer, M. Sanjabi, A. Talwalkar, and V. Smith, "Federated Optimization in Heterogeneous Networks," in Proceedings of Machine Learning and Systems (MLSys), vol. 2, pp. 429-450, 2020.

[3] 방준석, "연합학습 환경에서의 클라이언트 이질성에 관한 연구", 가천대학교 대학원 석사학위논문, 2024.