논문 22-47-01-21 The Journal of Korean Institute of Communications and Information Sciences '22-01 Vol.47 No.01
https://doi.org/10.7840/kics.2022.47.1.198

대규모 연합학습을 위한 효율적 분산 클러스터링 기법

김 형 빈◆, 김 용 호*, 유 철 우**, 박 현 희°

Efficient Distributed Clustering Algorithm for Large-Scale Federated Learning

Hyungbin Kim◆, Yongho Kim*, Cheolwoo You**, Hyunhee Park°

요 약

인공지능의 활용성이 커짐에 따라, 학습을 위한 데이터 활용 과정에서 개인정보 보호 이슈가 발생하고 있다. 이러한 이슈를 해결하기 위해 제안된 연합학습은, 스마트폰과 같은 분산 장치에서 학습이 이루어지며 분산 장치와 서버 간 원본 데이터의 교환 없이 학습이 진행된다. 연합학습은 참여하는 분산 장치의 데이터가 독립적이고 동일한 확률 분포를 갖는다고 가정되지만, 실제 연합학습에 참여하는 분산 장치의 데이터 분포는 불균일하기 때문에 이에 맞는 통계적 이질성이 고려되어야 한다. 본 논문에서는 대규모 연합학습 환경에서 각각의 분산 장치가 갖는 데이터의 불균일 분포 문제 개선을 목표로 한다. 분산 장치의 학습 결과로 도출된 가중치를 활용한 새로운 기법을 제안하며, 기존 연합학습과의 정확도 및 손실 성능 비교를 통해 시뮬레이션 결과를 보인다.

Key Words : Federated learning, Deep learning, Distributed optimization, Collaborative work, Cluster algorithm

ABSTRACT

As the use of artificial intelligence increases, privacy issues arise in the process of using data for learning. In the federated learning proposed to solve these issues, learning is performed on a distributed device such as a smartphone, and learning proceeds without exchanging original data between the distributed device and the server. In the federated learning, it is assumed that the data of the participating distributed devices are independent and have the same probability distribution, but since the data distribution of the distributed devices participating in the actual federated learning is non-independent and non-identically, the statistical heterogeneity should be considered. In this paper, we aim to improve the problem of non-independent and non-identically distribution of data in each distributed device in a large-scale federated learning environment. The proposed method uses the weights derived from the learning results of the distributed device, and the simulation results are shown by comparing the accuracy and loss rate performance with existing federated learning.

※ 본 논문은 2021년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No. 2021-0-00368, 6G 서비스를 위한 인공지능/머신러닝 기반 자율형 MAC 개발)
◆ First Author : Myongji University, Department of Information and Communication Engineering, hbkim@mju.ac.kr, 학생회원
° Corresponding Author : Myongji University, Department of Information and Communication Engineering, hhpark@mju.ac.kr, 종신회원
* Myongji University, Department of Information and Communication Engineering, yhkim@mju.ac.kr, 학생회원
** Myongji University, Department of Information and Communication Engineering, cwyou@mju.ac.kr, 종신회원
논문번호 : 202110-272-B-RE, Received September 30, 2021; Revised November 1, 2021; Accepted November 9, 2021

Ⅰ. 서 론

하드웨어 장치 및 소프트웨어 기술의 발전에 따라 최근 인공지능 기술이 크게 발전되면서 사용자 맞춤 추천 시스템, 이미지 인식 및 자연어 처리와 같은 다양한 분야에서 활용되고 있다[1]. 인공지능은 인간과 같은 판단을 내리기 위해 학습 과정을 거쳐야 한다. 다양한 데이터를 반복 학습하여 판단 알고리즘을 정하기 때문에 학습은 필수적이다. 인공지능의 학습을 진행하기 위해서는 목적에 맞는 방대한 양의 데이터 수집 또한 필수적이다. 그러나 데이터를 수집하는 과정에서 개인 정보 보호 문제가 발생한다[2].

이를 해결하기 위해 구글에서 연합학습(Federated learning)을 제안하였다[3]. 연합학습은 스마트폰과 같은 다수의 분산 장치와 하나의 서버로 이루어지며, 분산 장치에서 학습이 이루어지는 구조를 갖는다. 분산 장치에서의 학습 이후 도출된 가중치가 서버로 전송되고, 서버는 수신된 가중치들을 취합하는 기능을 하기 때문에 분산 장치가 학습에 사용한 데이터를 서버에 전송하지 않아도 학습이 이루어진다. 연합학습은 학습 시에 원본 데이터가 서버로 전송되지 않음으로써, 기존 머신러닝 기법들과 달리 데이터의 익명성이 보장되는 강점을 갖는다. 그러나 연합학습에 참여하는 분산 장치가 많아지면 각 분산 장치가 갖는 데이터의 불균일한 분포로 인하여 시스템 이질성, 통계적 이질성 문제 및 큰 통신 비용이 발생할 수 있다[4].

따라서 이러한 문제를 해결하기 위한 연구가 진행되어왔다. 첫 학습 진행 시에, 각 분산 장치에서 도출된 가중치를 서버에서 취합하고 표준편차를 계산하여, 평균에서 표준편차 범위 내의 값을 갖는 분산 장치들만을 이후 학습에 참여하도록 하는 연구가 진행되었다[5]. 이 연구에서 제안하는 기법을 통해 왜곡된 학습을 유발하는 분산 장치를 제외하고 학습을 진행하여 정확성과 손실 측면에서 기존 연합학습에 비해 향상된 성능을 보인다. 그러나 이 경우 연합학습에 전혀 참여하지 못하는 분산 장치가 발생하게 된다. 기존 연합학습에 참여하는 분산 장치들을 2개의 연합학습으로 분리하여 학습을 진행하는 연구가 진행되었다[6]. 이 연구에서 제안하는 기법을 통해 유사한 가중치를 갖는 분산 장치들을 2개의 연합학습으로 재구성하여 정확성 및 손실 측면에서 기존 연합학습 대비 향상된 성능을 보인다. 그러나 연합학습을 재구성할 때 미리 정한 2개의 연합학습으로 분리가 이루어지기 때문에, 분산 장치의 가중치에 따라 재구성이 이루어질 연합학습 수의 최적값을 찾아내는 연구가 필요하다.

또한, 연합학습을 무선 네트워크에 적용하였을 때 발생할 수 있는 문제를 해결하기 위한 연구가 진행되었다[7,8]. [7]에서 제안하는 기법은 분산 장치의 컴퓨팅 능력과 전송 에너지를 고려하여 무선 네트워크에서 패킷 오류 기반의 연합학습 손실 함수 최소화 문제를 해결하였다. [8]에서 제안하는 기법은 시간 할당, 대역폭 할당, 전력 제어 등의 시스템 에너지를 고려하여 기존 연합학습 대비 59.5% 감소한 에너지 소비량을 보인다. [7]과 [8]에서의 연구는 무선 네트워크 기반 연합학습의 수렴율(convergence rate)을 수식적으로 증명하여 에너지 소비량을 감소시킬 수 있으나, 연합학습에 참여하는 분산 장치의 수가 큰 경우에는 그에 맞는 시스템 에너지가 요구되기 때문에 동일한 시스템 에너지 내에서 대규모 연합학습에 대응할 수 있는 연구가 필요하다.

각각의 분산 장치가 갖는 데이터가 불균일하게 분포된 상태일 때 맨해튼 거리와 유클리드 거리를 활용하여 연합학습을 클러스터링하는 연구가 진행되었다[9]. 이 연구에서는 MNIST 데이터셋[10]을 사용하여 시뮬레이션을 진행하였으며, 시뮬레이션 결과 제안하는 기법이 기존 연합학습에 비해 더 적은 커뮤니케이션 라운드에서 성능이 수렴함을 보였다. 그러나 이 연구에서는 클러스터링 결과로 생성될 클러스터의 개수를 직접 지정해주어야 하는 한계가 있다.

따라서 본 논문에서는 대규모 연합학습이 갖는 데이터 불균일 분포 문제를 개선하기 위해 분산 장치의 가중치에 따라 연합학습을 분리하고, 분리해야 할 연합학습 수의 최적값을 찾는 기법을 제안한다.

Ⅱ. 본 론

2.1 연합학습

기존 머신러닝 기법들은 학습을 위해 모든 데이터를 중앙 서버에 취합하는 과정이 필요하다. 그러나 이 과정에서 데이터와 관련된 개인 정보 보호 문제가 발생할 수 있다. 이러한 문제를 해결하기 위해 연합학습은 중앙 집중형 모델과 달리, 데이터를 소유하고 있는 분산 장치에서 직접 데이터를 처리하여 모델을 학습하는 구조가 제안되었다. 각각의 분산 장치에서 학습을 통해 도출된 가중치들만 중앙 서버로 전송되기 때문에 분산 장치의 데이터를 중앙 서버에 취합하지 않고 학습을 진행할 수 있다.

그림 1은 기존 연합학습의 프로토콜을 나타낸다. 연합학습의 진행 순서는 크게 3단계로 구분할 수 있으며 로컬 모델 학습, 글로벌 모델 갱신, 로컬 모델 갱

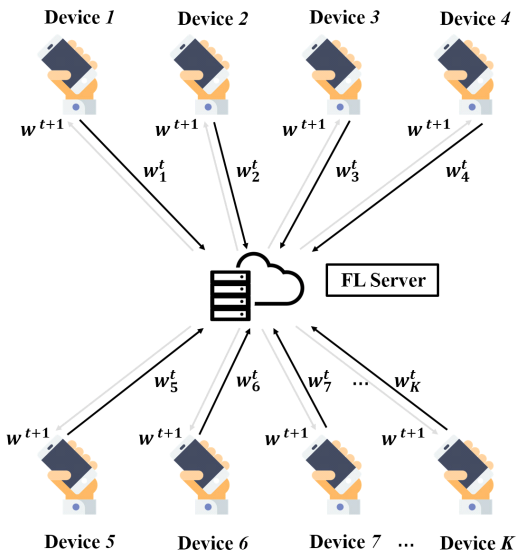


그림 1. 연합학습.
Fig. 1. The federated learning.

신이 반복되며 학습이 이루어진다. 이 3단계가 1회 반복하는 것을 커뮤니케이션 라운드 1회라고 정의한다.

2.1.1 로컬 모델 학습

연합학습의 첫 번째 커뮤니케이션 라운드를 시작할 때, 서버와 통신 가능한 분산 장치를 결정한다. 로컬 모델 학습이란 각 분산 장치가 갖고 있는 로컬 데이터를 처리하여 학습하는 것이다. 로컬 모델 학습 결과로써 분산 장치별 가중치가 도출된다. K개의 분산 장치가 연합학습에 참여하는 경우, K개의 분산 장치로부터 학습된 가중치를 중앙 서버로 전송한다.

2.1.2 글로벌 모델 갱신

글로벌 모델 갱신이란 중앙 서버에서 K개의 분산 장치로부터 전송받은 가중치를 취합하는 것이다. 가중치를 취합하기 위한 다양한 방법들이 연구되었으며, FedAvg와 FedSD와 같은 방법이 있다[3][5].

FedAvg는 가장 보편적으로 사용되는 방법이다. FedAvg는 글로벌 모델을 갱신하기 위해 모든 가중치를 평균화함으로써 글로벌 모델을 갱신한다. 연합학습에 참여하는 분산 장치의 수를 K, 연합학습의 커뮤니케이션 라운드 수를 t라고 할 때, FedAvg를 통한 글로벌 모델 갱신은 식 (1)로 표현된다.

$$w^{t+1} = \frac{\sum_{k=1}^{K} w_t^k}{K} \tag{1}$$

FedSD는 FedAvg에서 사용된 평균화 방법에 더해 가중치 표준편차를 활용한다. 모든 가중치에 대한 평균이 $A(w)$이고, 가중치 표준편차가 $SD(w)$일 때,

$$A(w) - SD(w) \leqq w_k \leqq A(w) + SD(w) \tag{2}$$

식 (2)을 만족하는 가중치만 평균화하여 글로벌 모델을 갱신한다.

2.1.3 로컬 모델 갱신

중앙 서버를 통해 처리된 글로벌 모델의 가중치를 분산 장치로 전송한다. 처리된 가중치는 다음 커뮤니케이션 라운드에서 분산 장치가 학습할 때 초기 가중치로 사용된다.

2.2 제안하는 기법

그림 2를 통해 본 논문에서 제안하는 Efficient Distributed Clustering 알고리즘(EDC 알고리즘)을 간략하게 나타낸다. EDC 알고리즘을 통해 기존 연합학습을 Ω개의 연합학습으로 재구성하며, 분산 장치의 가중치에 따라 Ω의 최적값을 찾을 수 있다. 그림 2는 Ω가 3인 상황이며, 분산 장치 1~3과 분산 장치 4~5,

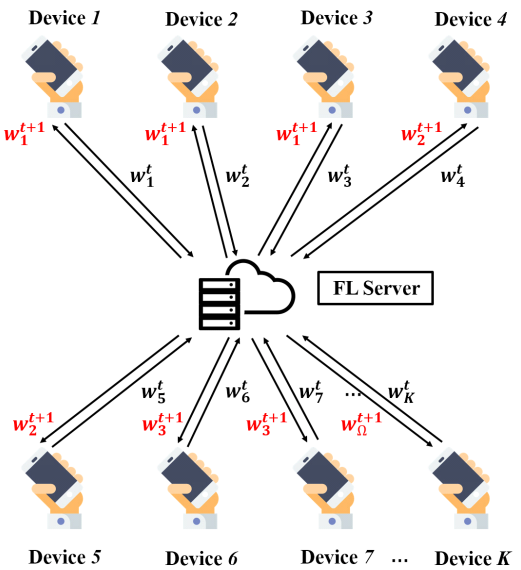


그림 2. 제안하는 기법.
Fig. 2. The proposed method.

분산 장치 6~7이 다른 연합학습으로 재구성되어 글로벌 모델에 의해 각각 w^{t+1_1}, w^{t+1_2}, w^{t+1_3}의 초기 가중치를 갖는다.

EDC 알고리즘은 기존 연합학습과 동일한 방식으로 로컬 모델 학습이 진행되고, 기존 연합학습의 글로벌 모델 갱신 단계 및 로컬 모델 갱신 단계에서 차이점을 가진다. 소단원에서 EDC 알고리즘을 소개하고자 한다.

2.2.1 글로벌 모델 갱신

분산 장치에서 학습이 이루어진 이후, 분산 장치의 레이어별 가중치가 생성된다. 본 논문에서는 분산 장치의 레이어별 가중치를 활용하여 Ω의 최적값을 찾는 EDC 알고리즘을 제안한다. 그림 3은 EDC 알고리즘의 순서를 나타내며, 그림 3 좌측의 행렬은 분산 장치의 레이어별 가중치를 나타낸다.

EDC 알고리즘과 기존의 연합학습의 차이는 글로벌 모델 갱신 및 로컬 모델 갱신에서 발생한다.

중앙 서버에서 분산 장치의 레이어별 가중치 평균화와 레이어별 전체 분산 장치의 가중치 평균화를 진행한다. 연합학습에 참여하는 분산 장치의 수는 K, 분산 장치가 갖는 레이어의 수는 N, 가중치를 갖는 행렬의 크기는 $I\times J$이다. $k=1,\cdots,K$이고 $n=1,\cdots,N$일 때, $k_{i,j}^{n}$는 k번째 분산 장치가 갖는 n번째 레이어의 (i,j) 위치에 있는 가중치이다. k번째 분산 장치가 갖는 레이어별 가중치 평균인 avg_k^n는 식 (3)으로 구할 수 있다.

$$avg_k^n = \frac{\sum_{i=1}^{I}\sum_{j=1}^{J} k_{i,j}^{n}}{I\times J}, (n=1,\cdots,N) \tag{3}$$

avg_{server}^n는 레이어별 모든 분산 장치의 가중치 평균이며, 식 (4)로 구할 수 있다.

$$avg_{server}^n = \frac{\sum_{k=1}^{K}\sum_{i=1}^{I}\sum_{j=1}^{J} k_{i,j}^{n}}{K\times I\times J}, (n=1,\cdots,N) \tag{4}$$

Ω의 최적값을 구하기 위해 avg_k^n와 avg_{server}^n의 관계를 분석한다. 그림 3처럼 레이어가 2개인 경우, avg_k^n와 avg_{server}^n 각각 2개의 값을 갖는다. 분산 장치가 갖는 모든 레이어에 한해서, 같은 레이어에 대응하는 avg_k^n와 avg_{server}^n를 비교했을 때 avg_k^n가 큰 레이어의 수를 γ(compared value)로 설정한다.

x축은 분산 장치의 번호, y축은 γ의 개수로 설정하여 그림 3 우측과 같은 그래프를 만들 수 있다. 그림 3 우측의 그래프는 γ에 따라 분산 장치의 번호를 오름차순으로 정렬하여 나타낸 것이다. 그림 3 우측의 그래프와 같이 오름차순 또는 내림차순으로 정렬된 형태에서 γ에 따른 분산 장치의 상호 유사도를 측정한다. 상호 유사도 측정 방법으로는 유클리드 거리 측정 방법을 사용한다[11].

유클리드 거리는 두 데이터 간의 거리를 속성별 거리 제곱의 합으로 측정하는데, 이는 다차원 공간에서

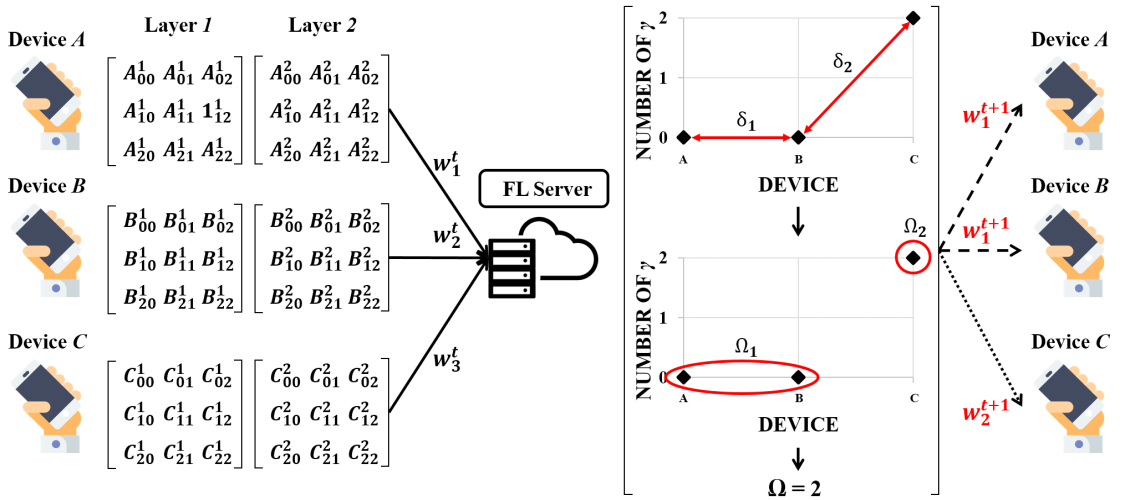


그림 3. EDC 알고리즘의 순서.
Fig. 3. The order of EDC algorithm.

두 점 사이의 직선거리를 의미한다. δ_e가 k번째와 $k+1$번째 분산 장치의 거리이고, δ_e에서 e는 $1,\cdots,K-1$이다.

점 $p=(p_1,p_2,\cdots,p_M)$와 점 $q=(q_1,q_2,\cdots,q_M)$이 있을 때 δ_e를 식으로 표현하면 식 (5)와 같다.

$$\delta_e=\sqrt{\sum_{m=1}^{M}(p_m-q_m)^2},(e=1,\cdots,K-1) \tag{5}$$

전체 유클리드 거리의 평균이 δ_{avg}이고 인접 좌표의 유클리드 거리가 $\delta_e(e=1,\cdots,K-1)$이면, δ_{avg}은 식 (6)와 같다.

$$\delta_{avg}=\frac{\sum_{e=1}^{K-1}\delta_e}{K-1} \tag{6}$$

인접 좌표의 유클리드 거리가 전체 유클리드 거리의 평균보다 클 때, 해당 분산 장치를 새로운 연합학습으로 분류한다. 인접 좌표의 유클리드 거리와 전체 유클리드 거리의 평균을 비교하는 의사코드를 알고리즘 1에 작성하였다.

그림 3 우측 그래프의 경우, 식 (5)를 통해 전체 유클리드 거리의 평균과 인접 좌표의 유클리드 거리를 계산한다. A 분산 장치와 B 분산 장치의 유클리드 거리가 δ_1이고 B 분산 장치와 C 분산 장치의 유클리드 거리가 δ_2이면 전체 유클리드 거리의 평균 δ_{avg}는 $\frac{\delta_1+\delta_2}{2}$이다. 좌표에 해당하는 x와 y를 대입하여 계산하면 다음과 같다.

알고리즘 1 δ_e와 δ_{avg}비교

Input: $\delta_e\,(e=1,\cdots,K-1)$
$\leftarrow$ *the euclidean distance between* e and $e+1$
Input: $\delta_{avg} \leftarrow$ *the average of all* δ_e
Input: $n \leftarrow 0$
1: **for** $e \leftarrow 1$ **to** $K-1$
2: $C_e \leftarrow e^{th}$ *device*
3: **if** $(\delta_e \le \delta_{avg})$ **then**
4: $C_{e+1} \in \Omega_n$
5: **else**
6: $C_{e+1} \in \Omega_{n+1}$
7: **end for**

$$\delta_1=1 \tag{7}$$

$$\delta_2=\sqrt{2}\,(\approx 1.414) \tag{8}$$

$$\delta_{avg}=\frac{1+\sqrt{2}}{2}(\approx 1.207) \tag{9}$$

알고리즘 1에 δ_1과 δ_2, δ_{avg}를 적용하면 C 분산 장치가 이전 분산 장치와는 다른 연합학습으로 분류되는 것을 확인할 수 있고, Ω의 최적값을 확인할 수 있다.

2.2.2 로컬 모델 갱신

알고리즘 1을 통해 정해진, 분산 장치가 속한 연합학습에 맞는 가중치를 분산 장치에 전송한다. 분산 장치로 전송된 가중치는 다음 커뮤니케이션 라운드에서 분산 장치가 학습할 때 초기 가중치로 사용된다.

본 논문에서 제안하는 EDC 알고리즘을 통해 분산 장치의 레이어별 가중치를 분석하여 기존 연합학습을 Ω개의 연합학습으로 재구성함으로써, 분산 장치의 수가 클 때 각 분산 장치가 갖는 데이터의 불균일한 분포 문제를 해결한다.

Ⅲ. 시뮬레이션

시뮬레이션을 위한 데이터로는 MNIST 데이터셋을 사용하였다. 연합학습에 참여하는 분산 장치는 100개로 설정하였으며 커뮤니케이션 라운드는 100회로 설정하였다. 표 1은 시뮬레이션 환경을 나타낸다. 표 2는 EDC 알고리즘의 시뮬레이션을 위해 사용한 모델 구조를 나타낸다. 표 2의 학습 파라미터가 있는 레이어에 제안하는 기법을 적용하여 Ω의 최적값을 찾아낸다. 클러스터

EDC 알고리즘을 통해 Ω개의 연합학습으로 재구성되기 때문에 Ω개의 정확도(accuracy)와 손실(loss)이 얻어진다. 성능 검증의 기준을 얻고자 본 시뮬레이션과 동일한 데이터 분포 조건으로 기존 연합학습을 진행하였다. 기존 연합학습의 정확도 및 손실을 Ω개

표 1. 시뮬레이션 환경
Table. 1. The simulation environment

Simulation environment	
Local epoch	1
Batch size	100
Optimizer	SGD
Train/test ratio	4:1

표 2. 모델의 레이어별 파라미터
Table. 2. The parameters of each layer of the model

Layer	Layer type	Feature Maps	Input size	Trainable parameters	Activation
Input	image	-	32x32	-	-
C1	Conv	6	28x28	156	Tanh
S2	Pool	6	14x14	0	-
C3	Conv	16	10x10	2416	Tanh
S4	Pool	16	5x5	0	-
F5	Dense	-	400	48120	Tanh
F6	Dense	-	120	10164	Tanh
F7	Dense	-	84	850	Tanh
Output	Dense	-	10	-	Softmax

정확도의 평균과 Ω개 손실의 평균과 비교함으로써 성능 검증을 진행하였다. 본 시뮬레이션 환경에서 $\Omega=6$을 얻었고, 표 3은 기존 연합학습과 재구성된 6개의 연합학습에 대한 정확도 및 손실을 나타낸다. 재구성된 6개의 연합학습을 각각 클러스터 1부터 클러스터 6으로 정의한다. 그림 4과 그림 5는 각각 정확도와 손실에 대한 기존 연합학습과 클러스터 6개의 성능을 나타낸다.

정확도와 손실에 대한 기존 연합학습과 클러스터 6개의 성능을 각각 비교한 결과, 클러스터 1, 클러스터 3, 클러스터 4의 경우 기존 연합학습에 비해 낮은 성능을 보인다. 그러나 클러스터 2, 클러스터 5, 클러스터 6은 정확도와 손실 측면에서 기존 연합학습에 비해 높은 성능을 보인다. 이를 통해, 클러스터에 참여하는 데이터 수와 같은 클러스터의 환경이 성능에 영향을 끼치는 것을 확인할 수 있다.

정확도와 손실 측면에서 기존 연합학습에 비해 낮은 성능을 보이는 클러스터 1, 클러스터 3, 클러스터 4의 성능 중에서도 클러스터 1의 성능이 크게 낮은 것

표 3. 커뮤니케이션 라운드 100회 수행 시 시뮬레이션 성능
Table. 3. The Simulation performance at 100th communication round

	정확도	손실
기존 연합학습	0.9714	0.1010
클러스터 1	0.9096	0.3093
클러스터 2	0.9787	0.0697
클러스터 3	0.9673	0.1176
클러스터 4	0.9698	0.1083
클러스터 5	0.9741	0.0898
클러스터 6	0.9764	0.0792

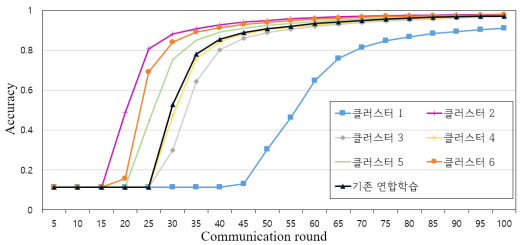


그림 4. 시뮬레이션 결과 정확도.
Fig. 4. The accuracy of the simulation.

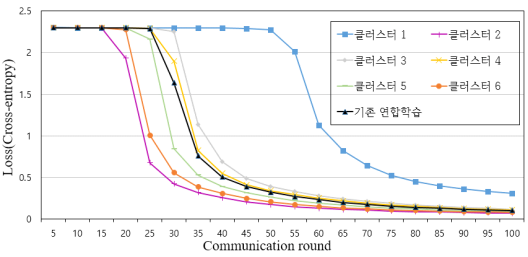


그림 5. 시뮬레이션 결과 손실.
Fig. 5. The loss(cross-entropy) of the simulation.

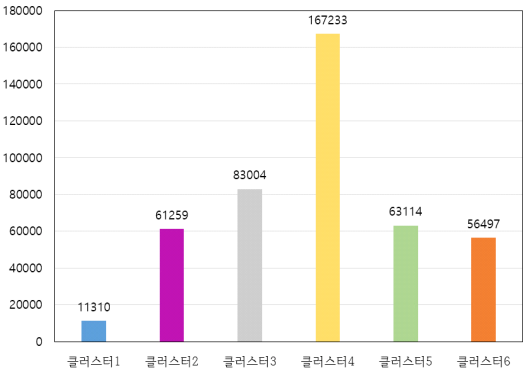


그림 6. 클러스터별 데이터 수.
Fig. 6. The number of data per cluster.

을 보인다. 이는 클러스터 1에 분류된 분산 장치들이 갖는 데이터의 수가 다른 재구성 연합학습에 비해 적었기 때문이다. 그림 6에서 각 재구성 연합학습이 갖는 데이터의 수를 나타낸다. 이를 통해, 클러스터에 분류된 분산 장치들이 갖는 데이터의 수가 작다면 기존 연합학습에 비해 낮은 성능을 보임을 확인할 수 있다.

Ⅳ. 결 론

본 논문에서는 기존 연합학습에 참여하는 분산 장치가 많을 때 데이터의 불균일 분포 문제 개선을 목표로 연구를 진행하였다. 본 논문에서 제안하는 EDC 알고리즘을 사용하여 분산 장치의 가중치를 분석하였고, 분산 장치의 가중치에 따라 Ω의 최적값을 찾아 연합학습을 Ω개로 재구성하여 성능을 확인하였다. 시뮬레이션 결과, 클러스터에 참여하는 데이터 수와 같은 클러스터의 환경이 정확도와 손실에 대한 클러스터 성능에 영향을 끼치는 것을 확인할 수 있었다. 또한, 학습에 사용될 데이터의 수와 같은 클러스터의 환경이 갖추어진 경우, 동일한 커뮤니케이션 라운드일 때 기존의 연합학습과 비교하여 성능이 향상됨을 확인하였으며 별도의 시스템 에너지가 추가되지 않아도 모델 자체적인 클러스터링을 통해 대규모 연합학습을 진행할 수 있음을 확인하였다.

References

[1] E. Brynjolfsson and A. Mcafee, "Artificial intelligence, for real," *Harvard Business Rev.,* 2017.

[2] J. H. Ziegeldorf, O. G. Morchon, and K. Wehrle, "Privacy in the Internet of Things: threats and challenges," *Secur. and Commun. Netw.*, vol. 7, no. 12, pp. 2728-2742, 2014.

[3] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Acras, "Communication-efficient learning of deep networks from decentralized data," in *Proc. AISTATS*, vol. 54, pp. 1273-1282, Feb. 2017.

[4] P. Kairouz, et al., "Advances and open problems in federated learning," *CoRR*, vol. abs/1912.04977, 2019.

[5] H. Kim, Y. Kim, G. Woo, J. Kim, and H. Park, "FedSD : Federated learning algorithm with standard deviation of weights for each user," in *Proc. Symp. KICS*, pp. 662-663, Jun. 2021.

[6] H. Kim, Y. Kim, and H. Park. "Reducing model cost based on the weights of each layer for federated learning clustering," *2021 ICUFN IEEE*, pp. 405-408, 2021.

[7] M. Chen, Z. Yang, W. Saad, C. Yin, H. V. Poor and S. Cui, "A Joint Learning and Communications Framework for Federated Learning Over Wireless Networks," in *IEEE Transactions on Wireless Communications*, vol. 20, no. 1, pp. 269-283, Jan. 2021

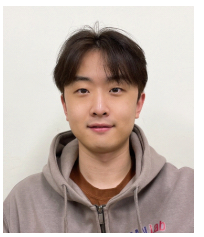
[8] Z. Yang, M. Chen, W. Saad, C. S. Hong and M. Shikh-Bahaei, "Energy Efficient Federated Learning Over Wireless Communication Networks," in *IEEE Transactions on Wireless Communications*, vol. 20, no. 3, pp. 1935-1949, March 2021

[9] C. Briggs, Z. Fan, and P. Andras, "Federated learning with hierarchical clustering of local updates to improve training on non-IID data," *CoRR*, abs/2004.11791, 2020.

[10] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," in *Proc. IEEE*, vol. 86, no. 11, pp. 2278-2324, 1998.

[11] A. K. Jain and R. C. Dubes, *Algorithms for clustering data*. Prentice Hall, Englewood Cliff, New Jersey, 1988.

김 형 빈 (Hyungbin Kim)



2021년 : 명지대학교 정보통신공학과 졸업

2021년~현재 : 명지대학교 정보통신공학과 석사과정

<관심분야> AI/ML 모델링, 연합학습, 컴퓨터 비전, 데이터 분석 및 알고리즘

[ORCID:0000-0002-0190-5018]

김 용 호 (Yongho Kim)

2021년 : 한국성서대학교 컴퓨터소프트웨어학과 졸업
2021년~현재 : 명지대학교 정보통신공학과 석박사통합과정
<관심분야> AI/ML 모델링, 지능형 시스템 개발, 컴퓨터 비전, 연합학습
[ORCID:0000-0001-7099-4336]

유 철 우 (Cheolwoo You)

1993년 : 연세대학교 전자공학과 학사 졸업
1995년 : 연세대학교 전자공학과 석사 졸업
1999년 : 연세대학교 전자공학과 박사 졸업
1999년~2003년 : LG전자 책임연구원
2003년~2004년 : EoNex 책임 연구원
2004년~2006년 : 삼성전자 책임 연구원
2006년~현재 : 명지대학교 정보통신공학과 교수
<관심분야> 5G mobile communications system, IoT, M2M, new multiple access schemes, multiple antenna transmission, advanced FEC
[ORCID: 0000-0003-3519-3490]

박 현 희 (Hyunhee Park)

2011년 : 고려대학교 전자컴퓨터공학과 공학박사
2011년~2012년 : 고려대학교 정보기술사업단 연구교수
2012년~2014년 : 프랑스 INRIA Research Center Postdoctoral researcher
2014년~2017년 : LG전자 차세대표준연구소 선임연구원
2017년~2020년 : 한국성서대학교 컴퓨터소프트웨어학과 조교수
2020년~현재 : 명지대학교 정보통신공학과 부교수
<관심분야> 무선통신 표준화, 통신시스템, 데이터 분석 및 알고리즘
[ORCID:0000-0003-3810-7367]