

# 불안정한 네트워크에서 데이터 부분 손실에 따른 연합 학습 성능 분석

권정민, 박형곤

이화여자대학교

jungmin.kwon@ewha.ac.kr, hyunggon.park@ewha.ac.kr

## Performance Analysis of Federated Learning with Partial Data Loss

Jungmin Kwon, Hyunggon Park

Ewha Womans University

### 요약

패킷 손실이 발생하는 불안정한 네트워크 환경에서 신뢰성을 보장하는 데이터 전달 방법은 네트워크의 지연뿐만 아니라 시스템 성능의 저하를 유발할 수 있는 한계점이 있다. 이러한 문제를 해결하기 위해 UDP와 같은 효율성에 집중한 데이터 전달 방법을 활용한 데이터 전달 방법에 대한 연구가 활발히 이루어지고 있다. 본 논문에서는 불안정한 네트워크 환경에서의 데이터 부분 손실이 연합 학습의 성능에 미치는 영향을 분석하였다. 연합 학습 과정에서 모델 파라미터의 부분 손실이 모델의 성능을 현저히 감소시킬 수 있음을 실험을 통해 확인하였다.

### I. 서론

인공지능(Artificial Intelligence, AI) 기술의 상용화가 확대됨에 따라 AI 모델 학습의 효율성과 보안성을 동시에 향상시킬 수 있는 방법으로 연합 학습(Federated Learning, FL)이 주목받고 있다. FL은 개별 모델이 사용하는 데이터의 직접 공유하는 대신 모델 파라미터를 공유함으로써 모델의 성능을 향상을 가능하게 한다. 일반적인 FL 연구에서는 모델 파라미터를 단일 패키지에 포함시켜 전달한다고 가정하지만, 실제로 공유되는 모델 파라미터의 크기가 메가 바이트(MByte)에서 기가바이트(GByte)에 이르기 때문에 이 가정은 현실적으로 적합하지 않을 수 있다 [1]-[3]. 따라서, 모델 파라미터의 크기가 큰 경우 이를 여러 패키지로 나누어 전송하는 방법을 기본으로 하여 현실적인 네트워크 환경에서 FL의 성능을 분석할 필요가 있다.

패킷 손실이 발생할 수 있는 불안정한 네트워크 환경에서 데이터 전송의 신뢰성을 위해 일반적으로 TCP(Transmission Control Protocol)를 이용한다. 그러나 TCP를 이용한 데이터 전송 방법은 네트워크의 지연을 초래할 수 있으며, 이는 여러 모델의 학습을 동시에 진행해야 하는 FL에서는 치명적인 문제가 될 수 있다. 이에 따라, 지연을 최소화하기 위한 대안적 전송 방법으로 UDP(User Datagram Protocol)를 기반으로 하는 FL 방법이 제안되었다 [4]-[6]. 그 결과, UDP를 기반으로 한 FL은 시간에 민감한 학습에서 지연을 크게 줄여 학습 속도를 향상시켰다. 그러나 불안정한 네트워크에서의 패킷 손실로 인해 FL의 성능 저하 문제가 발생하게 된다. 본 논문에서는 UDP 기반 FL에서 발생하는 패킷 손실이 성능 저하에 미치는 영향을 분석하고자 한다.

### II. 불안정한 네트워크에서의 연합 학습

본 논문에서는 그림 1과 같이  $N$ 개의 로컬 디바이스와 하나의 글로벌 서버로 구성된 UDP 기반의 중앙 집중형 FL 구조를 고려한다. FL 과정은 1) 로컬 모델 학습, 2) 로컬 모델 전송, 3) 모델 통합, 4) 글로벌 모델 전송, 5) 로컬 모델 업데이트의 다섯 단계로 크게 구분되며 이 일련의 과정을 반복하게 된다.

로컬 모델 학습 단계에서는 각 참여하는 로컬 디바이스들이 자체적으로 모델

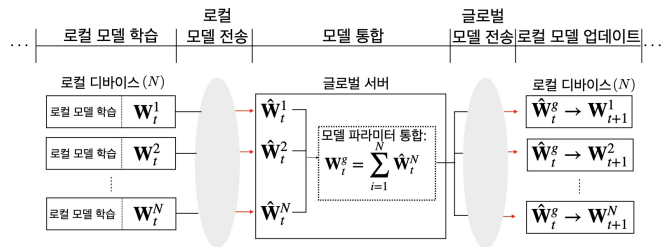


그림 1. 중앙 집중형 FL 구조 및 학습 과정

학습을 진행하고, 로컬 모델 전송 단계에서는 학습된 모델의 파라미터를 글로벌 서버로 전송한다. 즉, 시간  $t$ 에서  $k$  번째 로컬 디바이스에서는 학습된 모델 파라미터  $W_t^k$ 를 글로벌 서버로 전달한다. 이 단계에서 전달되는 모델 파라미터는 패킷 손실이 발생할 수 있는 불안정한 네트워크를 거치게 되고 글로벌 서버에서는 부분적으로 데이터를 손실한 모델 파라미터  $\hat{W}_t^k$  ( $k = 1, \dots, N$ )를 수신하게 된다. 이후 글로벌 서버에서는 수신된 모델 파라미터를 이용하여 모델 통합 단계를 진행하여 새로운 글로벌 모델 파라미터  $W_t^g$ 를 생성한다. 글로벌 모델 전송 단계에서는 새롭게 생성된 글로벌 모델 파라미터를 다시 각 로컬 디바이스로 전송하게 되며, 마지막으로 각 기기에서는 불안정한 네트워크를 거쳐 수신된 글로벌 파라미터  $\hat{W}_t^g$ 를 이용하여 로컬 모델 업데이트를 진행한다.

본 논문에서는 패킷 손실이 발생할 수 있는 네트워크 환경을 고려하며, 이때의 패킷 손실률  $p$ 는 업링크(uplink)와 다운링크(downlink)에서 동일한 값을 나타내고 각 링크마다 독립적으로 손실이 발생한다고 가정한다. 이때 전달하는 모델 파라미터를  $W$ 라고 하고 하나의 모델을 전달하는 패킷의 수가  $M$ 개로 가정할 때, 패킷의 수  $X$ 는 이항분포  $X \sim B(M, p)$ 를 따른다고 가정한다. 수신 단계에서는 이러한 환경에서 전달된 패킷을 기반으로 수신된 모델 파라미터  $\hat{W}$ 를 복구하게 되며, 이때 각 손실된 부분에 해당하는 모델 파라미터는 0

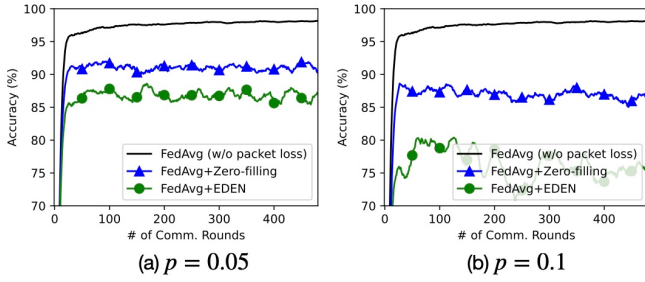


그림 2. MNIST 데이터 셋을 이용한 FL 정확도 성능

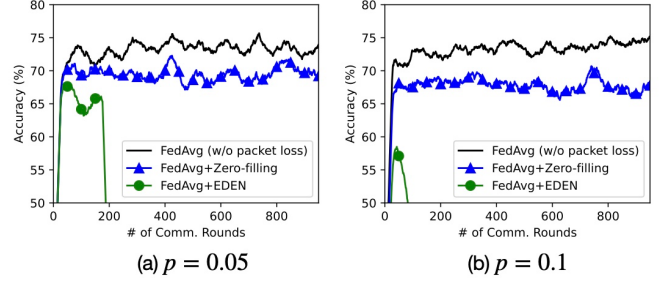


그림 3. CIFAR-10 데이터 셋을 이용한 FL 정확도 성능

값으로 대체하여 복구한다. 이렇게 수신된 모델 파라미터에 대하여 모델 통합 또는 로컬 모델 업데이트가 진행된다. 다음 장에서는 이러한 환경에서 확인되는 FL 정확도 성능을 실험을 통해 확인하고자 한다.

### III. 실험 결과

본 실험에서는 하나의 글로벌 서버와 100개의 로컬 디바이스로 구성된 네트워크를 고려하며 모델 통합에 참여하는 로컬 디바이스의 비율은 0.2로 설정한다. 패킷 손실률  $p$ 는 일반적인 경우를 대표하기 위해 0.05에서 0.1 사이로 설정하며 (즉,  $p \in \{0.05, 0.1\}$ ) 네트워크는 UDP를 기반으로 동작함에 따라 재전송은 발생하지 않는다.

본 실험에서는 MNIST와 CIFAR-10 데이터셋을 사용하여 이미지 분류 작업을 수행하며 이때 각 데이터셋들은 non-IID(Independent and Identically Distribution)에 따라 분포되어있다고 가정한다. 본 실험에서 각 로컬 모델은 DNN(Deep Neural Network) 모델을 이용하여 MNIST와 CIFAR-10의 이미지 분류를 진행하며 DNN 모델은 2개의 hidden layer로 각각 100개의 hidden node로 구성한다. 기본 학습률은 0.001로 설정하고, 로컬 epoch는 20, 배치 크기는 10으로 설정한다. 데이터의 부분 손실이 FL의 정확도 성능에 미치는 영향을 확인하기 위해, FedAvg [6] 기법을 사용하여 성능 분석을 수행한다.

본 실험에서는 손실된 모델 파라미터를 채우는 0으로 채우는 zero-filling 기법과, 모델 파라미터를 압축한 후 UDP 환경에서 전송하는 방법으로 제안된 EDEN [7] 기법을 이용하여 FL 정확도 성능 저하를 분석한다. 이때 EDEN 기법은 손실된 패킷에 해당되는 압축된 정보를 0으로 채운 후에 원 모델 파라미터를 복원한다.

그림 2와 그림 3 통해 확인할 수 있듯이, 모델 파라미터의 부분적 손실이 발생한 경우 성능이 크게 저하되는 것을 관찰 할 수 있으며 패킷 손실률이 증가할 수록 성능 감소가 더 크게 일어나는 것을 확인할 수 있다. 뿐만 아니라 EDEN과 같이 모델 파라미터를 압축 기법을 이용하여 인코딩 한 후 전달된 경우에는 그 성능이 크게 저하될 뿐만 아니라 CIFAR-10 데이터 셋의 이미지 분류의 경우 학습이 불가능한 것을 확인할 수 있다. 따라서 데이터의 부분적 손실로 인한 성능 저하를 방지하면서 안정적으로 수렴할 수 있도록 하는 데이터 전송 방법 개발이 매우 필수적임을 알 수 있다.

### IV. 결론

본 논문에서는 불안정한 네트워크 환경에서의 데이터 부분 손실이 연합학습의 성능에 미치는 영향을 실험적으로 분석하였다. 데이터 전송의 효율성을 고려할 때 발생하는 부분 손실을 간과하면 모델의 정확도가 심각하게 저하됨을 확인하였고, 모델 파라미터에 대한 추가적인 인코딩 과정이 포함된 데이터는 학습이 불가능해질 수 있음을 확인하였다. 이러한 결과를 통해 패킷 손실로 인한 성능 저하를 방지할 수 있는 FL 시스템 개발의 필수적임을 알 수 있었다.

### ACKNOWLEDGMENT

이 논문은 2021년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임(No.2021-0-00739, 분산/협력 AI 기반 5G+ 네트워크 데이터 분석 기능 및 제어 기술 개발, RS-2024-00344830-(총괄10-세부3) 6G 네트워크 구조/산업융합 표준기술 개발 및 표준화)

### 참고 문헌

- [1] 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," *IEEE Transactions on Wireless Communications*, vol. 20, no. 1, pp. 269-283, 2020.
- [2] L. Cui, X. Su, Y. Zhou, and L. Zhang, "ClusterGrad: Adaptive gradient compression by clustering in federated learning," in *IEEE Global Communications Conference*. IEEE, 2020, pp. 1-7.
- [3] Z. Qin, D. Chen, B. Qian, B. Ding, Y. Li, and S. Deng, "Federated full-parameter tuning of billion-sized language models with communication cost under 18 Kilobytes," *arXiv preprint arXiv:2312.06353*, 2023.
- [4] G. Cleland, D. Wu, R. Ullah, and B. Varghese, "FedComm: Understanding communication protocols for edge-based federated learning," in *2022 IEEE/ACM 15th International Conference on Utility and Cloud Computing (UCC)*. IEEE, 2022, pp. 71-81.
- [5] H. Ye, L. Liang, and G. Y. Li, "Decentralized federated learning with unreliable communications," *IEEE Journal of Selected Topics in Signal Processing*, vol. 16, no. 3, pp. 487-500, 2022.
- [6] Y. Zhou, Q. Ye, and J. Lv, "Communication-efficient federated learning with compensated overlap-fedavg," *IEEE Transactions on Parallel and Distributed Systems*, vol. 33, no. 1, pp. 192-205, 2021.
- [7] S. Vargaftik, R. B. Basat, A. Portnoy, G. Mendelson, Y. B. Itzhak, and M. Mitzenmacher, "Eden: Communication-efficient and robust distributed mean estimation for federated learning," in *International Conference on Machine Learning (ICML)*. PMLR, 2022, pp. 21984-22014.