

# 다중 접속 채널 기반 무선 분할 학습

김선중, 전요셉  
포항공과대학교

{ksj8970, yoseb.jeon}@postech.ac.kr

## Wireless Split Learning based on Multiple Access Channels

Seonjung Kim, Yo-Seb Jeon  
Pohang Univ. of Science and Technology (POSTECH)

### 요약

본 논문은 다중 접속 채널(multiple access channel, MAC) 환경에서 분할 학습(split learning)을 위한 통신 지연을 감소시키고 학습의 효율을 높이는 무선 분할 학습 구조를 제안한다. 제안하는 구조는 서버 모델의 비동기식 업데이트와 주파수 분할 이중화(frequency division duplex, FDD)를 활용한다. 제안된 복합 다중 접속 방식은 시분할 다중 접속과 비직교 다중 접속을 융합하여 통신 지연 시간을 단축시키고, 서버의 비동기 업데이트를 가능하게 한다. 이를 통해 서버는 그룹별로 활성화 값을 순차적으로 받으면서 모델을 효율적으로 업데이트할 수 있다. 또한, 이전 연구의 그룹화 알고리즘을 일반적인 크기의 그룹에 적용할 수 있도록 확장한다. 모의 실험을 통해, 제안된 구조가 분할 학습의 빠른 수렴을 달성함을 입증한다.

### I. 서론

분할 학습(split learning)은 분산형 기계 학습 기법의 일종으로 글로벌(global) 모델을 분할 층을 기준으로 분할하여 앞부분을 무선 단말기들이 가지고 뒷부분을 서버가 가져간 후 서로 협력하여 모델을 학습한다. 순전파 과정에서 단말기들은 데이터를 신경망의 앞부분에 넣어 분할층에서 나오는 활성화 값을 서버에게 보내고 서버는 받은 활성화 값을 신경망의 뒷부분에 넣어 최종 출력을 얻는다. 역전파 과정에서는 서버에서 손실함수를 미분하여 경사하강법을 이용해 분할층까지 계산한 그라디언트 값을 단말기로 보내고 단말기에서는 받은 그라디언트 값을 이용해 입력층까지 그라디언트 값을 계산할 수 있다. 계산한 그라디언트를 이용해 모델을 업데이트 하는 과정을 단말기별로 번갈아 가며 모델이 수렴할 때까지 진행한다. 이러한 학습 방식은 연합 학습과 마찬가지로 학습을 위한 데이터를 직접 보내지 않기 때문에 보안에 강하다는 장점이 있고, 글로벌 모델의 크기에 비해 통신 오버헤드가 증가하는 연합 학습의 문제를 해결할 수 있다 [1]. 그럼에도 여전히, 단말기와 서버 간의 빈번한 통신을 필요로 하는 분산형 기계 학습의 특성상 통신 지연 시간을 줄이는 것은 여전히 중요한 문제이다.

이전 연구 [2]에서는 비직교 다중 접속 채널에서 통신 지연을 감소시키는 효율적인 단말기 그룹화 방식을 제안하였다. 본 논문에서는 다중 접속 채널의 특징을 활용한 서버 모델의 비동기식 업데이트와 주파수 분할 이중화(FDD)를 적용한 새로운 무선 분할 학습 구조를 제안한다. 또한, 기존 연구에서 그룹 크기를 2로 제한했던 그룹화 방식을 2보다 큰 그룹에도 적용할 수 있도록 개선한다.

### II. 본론

본 논문에서는 서버가  $N$ 개의 단말기들( $\mathcal{S} = \{1, \dots, N\}$ )을  $K$ 개씩 묶어  $G$ 개의 클러스터로 나누어 신경망 모델을 훈련시키는 분할 학습을 고려한다. 이 때, 단말기들의

SNR은  $(\text{SNR}_1, \dots, \text{SNR}_N)$ 로 주어지고,  $\text{SNR}_1 \leq \dots \leq \text{SNR}_K$ 의 순서로 정렬되어 있다고 가정한다. 이 때, 각각의 클러스터는 다중 접속을 이용해 서버와 통신하게 되는데 본 논문은 시분할 다중 접속과 비직교 다중 접속을 융합한 복합 다중 접속을 제안한다. 구체적으로, 클러스터 내부에서  $K$ 개의 단말기를  $L$ 개씩 묶어  $C$ 개의 그룹으로 나누고 같은 그룹에 속한 단말기들은 비직교 다중 접속을 이용해 동시에 서버와 통신하고 그룹들 간에는 시분할 다중 접속을 이용해 하나의 그룹이 활성화 값을 전부 보내면 순차적으로 다음 그룹이 활성화 값을 보낸다. 복합 다중 접속을 통해 얻는 이득에는 크게 3가지를 제시한다.

첫 번째로, 직교 다중 접속보다 짧은 지연시간을 성취할 수 있다는 것이다. 비직교 다중 접속은 시간이 스펙트럼과 같은 통신 자원의 활용성이 더 높기 때문에, 최적의 자원 배분 상황에서 더 짧은 통신 지연을 얻을 수 있다. 특히 시분할 다중 접속 대비 더 짧은 지연 시간을 얻는 것은 지난 연구에서 증명되었고, 주파수 분할 다중 접속 대비 우수성은 모의 실험결과를 통해 입증된다.

두 번째로, 그룹 간 시분할 다중 접속을 활용해서 서버 모델에 비동기식 업데이트를 적용할 수 있다. 클러스터 내의 모든 단말기가 동시에 비직교 다중 접속을 사용해 통신을 하게 되면 서버는 순전파를 하기 전에 모든 단말기의 활성화 값을 수신해야 한다. 하지만 제안한 복합 다중 접속에서는 단말기가  $L$ 개씩 묶여 동시에 상향 링크 통신을 하기 때문에 서버도 활성화 값을  $L$ 개씩 받게 된다. 서버는 처음 그룹으로부터  $L$ 개의 활성화 값을 받았을 때, 받은 활성화 값을 이용해 순전파 및 모델 업데이트를 진행하면서 다음  $L$ 개의 활성화 값을 수신할 수 있다. 이 방식은 단말기의 모델보다 서버의 모델을 자주 업데이트하기 때문에 비동기식 업데이트로 간주된다. 비동기식 업데이트를 통해서 서버 모델의 파라미터들은 이전 그룹의 데이터를 이용해 더 좋은 시작점으로 이동한 후, 다음 그룹의 활성화 값을 기다리는 효과를 얻게 된다. 일반적으로 서버의 계산 능력이 높고 단말기의 송신파워가 작은 상황에서는 상향 링크 지연 시간이 서버의 계산 시간을 압도하기 때문에

비동기식 업데이트의 도입은 추가적인 지연시간을 발생시키지 않는다.

세 번째로, 주파수 분할 이중화 (FDD)를 이용해 서버가 활성화 값을 수신함과 동시에 이전 그룹에게 그라디언트 값을 보낼 수가 있다. 이는 복합 다중 접속을 이용해 서버가 비동기식 업데이트가 가능하기 때문에 얻는 이득으로, 서버가 업데이트를 진행하기 위해 모든 활성화 값을 받아야 하는 비직교 다중 접속 상황에서는 이용할 수 없다. 다만, 비동기식 업데이트와 FDD 를 이용한 동시 전송은 클러스터의 크기가 그룹의 크기보다 큰 경우 ( $K > L$ )에만 이용할 수 있다. 이는  $K > L$  인 경우에만 하나의 클러스터 내에 여러 개의 그룹이 존재하기 때문이다.

마지막으로, 제안된 복합 다중 접속 상황에서 상황 링크 시간을 최소화할 수 있는 그룹화 알고리즘을 제시한다. 우선 그룹의 크기가 2 인 경우에 대해서는 이전 연구에서 효율적인 알고리즘이 제시되었다. 이 결과를 활용하고자,  $N$ 개의 단말기들을 우선 아래와 같이  $L/2$ 개씩 묶어서  $2N/L$ 개의 부분집합들로 쪼갬다.

$$\begin{aligned} \mathcal{H}_1 &= \left\{1, 2, \dots, \frac{L}{2}\right\}, \mathcal{H}_2 = \left\{\frac{L}{2} + 1, \dots, L\right\}, \dots \\ \mathcal{H}_{2N/L} &= \left\{N - \frac{L}{2} + 1, \dots, N\right\} \end{aligned} \quad (1)$$

여기서 각각의 부분집합  $\mathcal{H}_k$ 에 소속된 단말기들의 평균 SNR을  $\mathcal{H}_k$ 의 SNR로 생각하면 각각의 부분집합들을 하나의 단말기처럼 생각할 수 있다. 이제  $\mathcal{H}_1, \mathcal{H}_2, \dots, \mathcal{H}_{2N/L}$ 에 이전 연구에서 제안한 그룹 크기 2 에 대한 단말기 그룹화를 적용하면 그룹 크기가  $L$  인  $N/L$ 개의 그룹을 얻을 수 있다. 이제 그룹을  $C$ 개씩 하나로 묶으면  $G$ 개의 클러스터를 구성할 수 있고, 클러스터 별로 순차적으로 제안된 무선 분할 학습 방식은 진행할 수 있다.

모의 실험은 기지국을 중심으로 반경 10m 에서 1km 사이에 단말기를 균일한 분포로 생성시켜 실험하였다. 데이터셋은 불균일하게 분포된 MNIST 데이터셋을 사용하였고 12 개의 층을 가진 합성곱 신경망 중 단말기는 앞의 3 개의 층을 서버는 나머지 9 개의 층을 나누어 가지고 분할 학습을 진행하였다. 학습률은 학습 기법마다 가장 빠른 수렴 속도를 보여주는 학습률을 사용하였다.

그림 1 은 총 20 개를 단말기가 학습에 참여하였을 때, 학습 시간 대비 테스트 성능을 측정한 결과를 나타낸다. 클러스터의 크기는  $K = 2$  와  $K = 4$  두 가지를 고려하였고, 제안된 기법에서 그룹의 크기는  $L = 2$ 로 제한하였다. 비교 기법으로는 주파수 분할 다중 접속 (FDMA)을 이용한 클러스터 구조를 사용하였다. 실험 결과를 보면 FDMA 보다 제안한 구조가 빠르게 수렴하는 것을 통해 복합 다중 접속의 우수성을 확인할 수 있다. 또한, 제안된 기법 내에서도  $K = 4$  인 경우가  $K = 2$  인 경우보다 빠른 수렴 속도를 보여주는데, 이는  $K = 4$  일 때에만  $K > L$  이 성립하여 비동기식 업데이트와 FDD 를 이용해 보다 분할 학습의 효율성이 높아졌음을 입증한다.

그림 2 는 총 24 개의 단말기가 학습에 참여하였을 때, 다양한 그룹 크기들 간의 학습 시간 대비 테스트 성능을 측정한 결과를 나타낸다. 그룹 크기에 의한 효과만을 보기 위해 클러스터의 크기는  $K = 24$  로 설정하여 하나의 클러스터만 사용하였다. 실험 결과를 보면  $L = 2$  인 경우가 가장 빠른 수렴 속도를 보여주는 것을 확인할 수 있다.

### III. 결론

본 논문에서는 분할 학습의 학습 시간을 감소시키는 효율적인 학습 구조를 제안한다. 제안된 구조는 시분할 다중 접속과 비직교 다중 접속을 융합한 복합 다중 접속 구조를

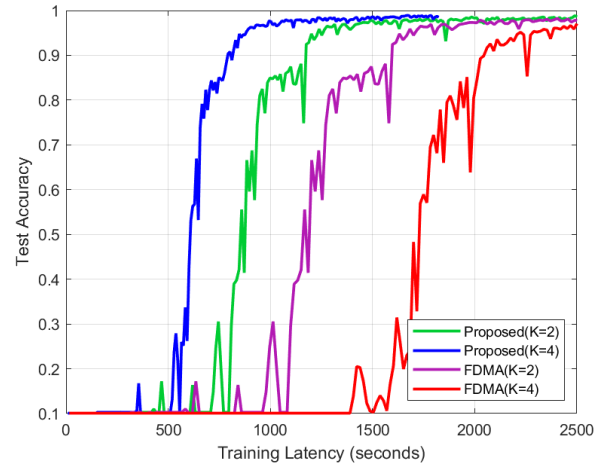


그림 1. 다양한 분할 학습 기법들 간의 학습 시간 대비 테스트 성능 비교.

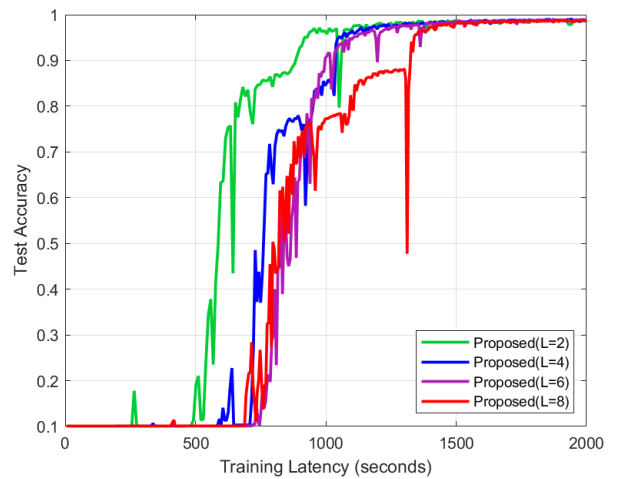


그림 2. 다양한 그룹 크기들 간의 학습 시간 대비 테스트 성능 비교.

도입하여 서버 모델의 비동기식 업데이트를 가능하게 한다. 이러한 비동기식 업데이트 모델 학습의 성능을 높일 뿐 아니라 FDD 를 이용해 그라디언트 전송에 필요한 통신 지연을 감소시키는 역할도 한다. 또한, 이전 연구에서 제안한 그룹화 방식을 일반적인 크기의 그룹에 대해서 적용할 수 있도록 확장한다. 모의 실험을 통해, 제안된 구조가 분할 학습의 빠른 수렴을 달성함을 입증한다.

### ACKNOWLEDGMENT

이 논문은 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No. 2022R1C1C1010074).

### 참고 문헌

- [1] A. Singh, P. Vepakomma, O. Gupta, and R. Raskar, "Detailed comparison of communication efficiency of split learning and federated learning," arXiv:1909.09145 [cs.LG], Sep. 2019, [Online]. Available: <https://doi.org/10.48550/arXiv.1909.09145>
- [2] S. Kim and Y.-S. Jeon, "Device Grouping for Split Learning over Multiple Access Channel," in Proc. KICS Int. Conf. Commun. 2023, pp. 639-640, Jeju Island, Korea, June 2023.