

# 부분 학습 기반 연합학습에서 수렴 안정화를 위한 EMA 기반 서브모델 선택 기법

이희재, 이주형\*  
가천대학교

[leeheejae@gachon.ac.kr](mailto:leeheejae@gachon.ac.kr), [j17.lee@gachon.ac.kr](mailto:j17.lee@gachon.ac.kr)\*

## EMA-based Stable Submodel Selection for Partial Training in Federated Learning

Heejae Lee, Joohyung Lee\*  
Gachon Univ.

### 요약

연합 학습(Federated Learning) 환경에서 클라이언트의 통신 및 계산 비용을 절감하기 위해 글로벌 모델의 일부만을 학습하는 부분 학습(Partial Training) 기법이 연구되고 있다. 그러나 데이터 불균형(Non-IID)환경에서는 파라미터의 크기가 불연속적으로 변화하기 때문에, 단순 크기 기반 파라미터 선택은 서브모델의 일관성을 저해하여 전역 모델의 수렴 불안정성과 성능 진동을 유발한다. 본 논문은 이를 해결하기 위해 시간적 일관성을 고려한 서브모델 구성 기법을 제안한다. 제안 기법은 지수 이동 평균을 적용한 파라미터 누적 중요도와 변동성 지표를 결합하여 선별 안정성을 높이고, 서버 측 학습률 감소 및 Warm-up 전략을 도입한다. 실험 결과, 제안 기법은 기존 방식 대비 Non-IID 환경에서도 전역 모델의 안정적인 수렴 성능을 달성함을 확인하였다.

### I. 서론

연합 학습(Federated Learning, FL)은 데이터를 서버에 모으지 않고 전역 모델을 학습하는 프레임워크로, 데이터 프라이버시 보호 및 분산 학습 환경에서의 활용 가능성으로 인해 주목받고 있다 [1]. 그러나 실제 환경에서는 클라이언트 간 연산 및 통신 자원 불균형 문제로 인해 모든 클라이언트가 동일한 규모의 모델을 학습하는 데 한계가 존재한다 [2, 3, 4].

이러한 한계를 극복하기 위해 글로벌 모델의 일부만을 학습하는 부분 학습(Partial Training) 기반 연합 학습 기법이 제안되었다 [2, 3, 4]. 대표적인 예시로써, FIARSE [2]는 파라미터의 가중치 크기를 기반으로 중요도가 높은 파라미터를 선택하여 서브모델을 구성한다. 그러나 데이터 Non-IID 환경에서 파라미터별 업데이트 빈도와 방향의 불균형으로 인해 중요도 추정과 서브모델 추출 불안정성이 심화된다. 이는 전역 모델 집계 과정에서 업데이트 노이즈를 키워 수렴 과정 전반에서 성능 불안정을 유발한다.

본 논문은 이러한 문제를 해결하기 위해 (i) 지수 이동 평균(Exponential Moving Average, EMA) 기반 가중치 크기 누적을 통해 변동을 완화하고, (ii) 누적 크기 대비 현재 가중치 차이를 파라미터의 안정성 지표로 활용하여 불안정한 파라미터의 선택을 억제하며, (iii) 서버 측 지수적 학습률 감소를 적용하여 수렴 후반의 오버피팅을 완화하는 안정화 기법을 제안한다. 또한 학습 초기에 Warm-up 라운드를 도입하여 중요도 통계의 초기

편향을 완화한다. CIFAR-10 을 이용한 실험 결과, 제안 기법은 Non-IID 환경에서 기존 방식 대비 라운드 간 글로벌 정확도 진동을 완화하였다.

### II. 본론

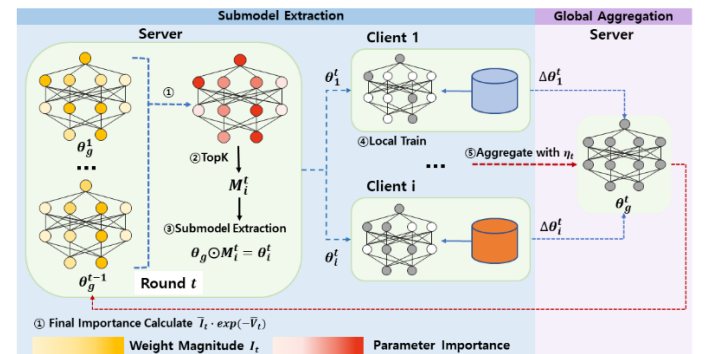


그림 1. 제안 기법의 전체 프레임워크.

#### 2.1 메서드

파라미터 선택 변동에 따른 수렴 불안정성을 완화하고자, 본 논문에서는 시간적 누적 중요도를 도입한다. 학습 초기에는 전체 모델에 대해 Warm-up 라운드를 도입하여 중요도 통계의 초기 편향을 완화한다.

라운드  $t$ 에서의 단일 시점 중요도  $I_t = |w_t|$ 를 기반으로, EMA 기반 누적 중요도  $\bar{I}_t$ 는 다음과 같이 정의된다 [5]:

$$\bar{I}_t = \alpha \bar{I}_{t-1} + (1 - \alpha) I_t. \quad (1)$$

여기서  $\alpha \in [0,1)$  는 과거 중요도의 반영 비율을 조절하는 계수이다. 이를 통해 순간적인 가중치 변동의 영향을 완화하고 중요도 추정의 분산을 감소시킨다. 또한, 파라미터의 가중치 크기 변화의 불안정성을 측정하기 위해 편차  $d_t = |I_t - \bar{I}_t|$  를 정의하고 이를 누적한  $V_t$  를 사용한다. 누적된 라운드에 대해서 파라미터의 불안정성  $V_t$  는 다음과 같이 정의된다:

$$\bar{V}_t = \alpha \bar{V}_{t-1} + (1 - \alpha) d_t. \quad (2)$$

$V_t$  는 학습 전반에 걸쳐 파라미터의 가중치 크기 얼마나 안정적으로 유지되는지를 나타낸다. 따라서 본 논문에서는 가중치 크기와 안정성을 결합한 최종 중요도  $Importance_t^{final}$  를 다음과 같이 정의한다:

$$Importance_t^{final} = \bar{I}_t \cdot \exp(-\bar{V}_t). \quad (3)$$

서버는 최종 중요도를 기준으로 Top k 개의 파라미터를 선택하여 각 클라이언트에 대해 마스크  $M_t^f$  를 생성하고, 서버모델  $\theta_t^f$  를 배포한다. 클라이언트는 로컬 데이터로 지도 학습 손실 함수를 최소화하도록 서버모델의 로컬 학습을 수행한다. 서버는 클라이언트의 로컬 학습 결과  $\Delta\theta_t^f$  를 집계하여 전역 모델  $\theta_g^f$  을 업데이트 한다. 집계 단계에서 수렴 후반 전역 모델 성능 진동을 완화하기 위해, 서버 측 학습률을 점진적으로 감소시키는 전략을 적용한다 [4]. 라운드  $t$ 에서의 서버 학습률  $\eta_t$  는 다음과 같다:

$$\eta_t = \eta_0 \cdot \gamma^t. \quad (4)$$

여기서  $\eta_0$  는 초기 서버 학습률이며,  $\gamma \in (0,1)$  는 감쇠 계수이다. 해당 학습률은 서버가 수신한 로컬 업데이트  $\Delta\theta_t^f$  를 전역 모델에 반영하는 보폭을 조절한다. 이를 통해 학습 초반에는 넓은 보폭으로 빠른 탐색을 유도하고, 수렴 후반부에는 업데이트 크기를 줄여 성능 진동을 억제하고 안정적인 수렴을 유도한다.

## 2.2 실험 결과

제안 기법을 검증하기 위해, CIFAR-10 데이터셋을 사용하여 이미지 분류 실험을 수행하였다. 데이터 Non-IID 환경을 재현하기 위해, 학습 데이터를 Dirichlet 분포 ( $\alpha=0.3$ ) 에 따라 100 개의 클라이언트에 분할하였으며, 각 라운드마다 10 개의 참여 클라이언트를 무작위 추출한다. 학습 모델은 BetaResNet18\_sbn 을 사용하였으며, 글로벌 모델 파라미터 수에 대한 서버모델 추출 비율은 1.0, 0.25, 0.0625, 0.015625 로 설정하였다.

실험 결과, 제안 기법은 학습 후반부(Late) 정확도 변화량의 표준편차( $Std(\Delta Acc)$ )를 Baseline 대비 35% 감소시켜 학습 안정성을 향상시켰다. 특히 정확도가 감소하는 경우의 편차(Downside)와 전체 학습 라운드의 최대 낙폭(MaxDrop)을 37% 낮춤으로써, Non-IID 환경의 치명적인 성능 급락과 수렴불안정을 효과적으로 방어함을 확인하였다. 최고 정확도의 소폭 감소는 수렴 안정성 확보를 위한 허용 가능한 상충 관계이며, 초기

Method	Max(Whole)	Max(Late)	Final
Baseline(FIARSE)	56.44%	56.44%	54.51%
Ours	55.03%	55.03%	53.97%

표 1. CIFAR-10 데이터셋에서 Baseline 과 제안 방법의 전역 정확도 비교.

Segment (round)	Method	Std ( $\Delta Acc$ )↓	Std(Down side)↓	MaxDrop↓
Early	Baseline	0.0287	0.0106	0.0396
	Ours	0.0284	0.0145	0.0511
Late	Baseline	0.0207	0.0227	0.0776
	Ours	<b>0.0134</b>	<b>0.0129</b>	<b>0.0486</b>

표 2. 라운드 간 전역 정확도 변화량 기준 Baseline 과 제안 방법의 수렴 안정성 비교.

변동성의 유지는 Warm-up 전략이 학습 초기 탐색 성능을 저해하지 않음을 시사한다.

## III. 결론

본 논문에서는 데이터 분포가 불균형한 연합학습 환경에서 부분 학습 적용 시 발생하는 가중치 크기 기반 파라미터 선택의 불안정성과 이로 인한 전역 모델의 성능 진동 문제를 해결하기 위해 EMA 기반의 누적 중요도와 안정성 지표를 활용한 가중치 안정화 기법을 제안하였다. 또한 서버 측 학습률 감소와 Warm-up 전략을 통해 학습 초기에 충분한 탐색을 보장하고, 후반에는 안정적인 수렴을 유도하는 단계별 최적화를 수행하였다. 이를 통해 가중치 크기 기반 서버모델 추출 방식을 유지하면서 수렴 불안정성을 완화할 수 있는 토대를 마련하였다.

## ACKNOWLEDGMENT

“본 연구는 과학기술정보통신부 및 정보통신기획평가원의 2026년도 SW 중심대학사업의 결과로 수행되었음”(2021-0-01389)

## 참고 문헌

- [1] McMahan, B., Moore, E., Ramage, D., Hampson, S., and Aguera y Arcas, B., “Communication-efficient learning of deep networks from decentralized data,” Proc. International Conference on Artificial Intelligence and Statistics (AISTATS), pp. 1273-1282, 2017.
- [2] F. Wu, X. Wang, Y. Wang, T. Liu, L. Su, and J. Gao, “FIARSE: Model-Heterogeneous Federated Learning via Importance-Aware Submodel Extraction,” in Proc. Advances in Neural Information Processing Systems (NeurIPS), vol. 37, pp. 115615-115651, 2024.
- [3] C. Diao, J. Lin, Y. Li, J. Liu, and J. Li, “HeteroFL: Computation and Communication Efficient Federated Learning for Heterogeneous Clients,” in Proc. ICLR, 2021.
- [4] Liang, J., Zhang, L., Qu, X., and Wang, J., “FedEcover: Fast and stable converging model-heterogeneous federated learning with efficient-coverage submodel extraction,” Proc. IEEE 41st International Conference on Data Engineering (ICDE), pp. 2575-2587, 2025.
- [5] Polyak, B. T. and Juditsky, A. B., “Acceleration of stochastic approximation by averaging,” SIAM J. Control Optim., vol. 30, no. 4, pp. 838-855, 1992.