GFedCL: Graph-Based Federated Continual Learning with Spatial and Temporal Awareness

Qingyang Yu[1], Yang Hua[2], Qizhen Zhang[3], Hao Wang[1]
Stevens Institute of Technology[1], Queen's University Belfast[2], University of Toronto[3]

zoom IntelliSys Lab

ICML 2026

1. Motivation

- Auxiliary conditioning helps in replay.

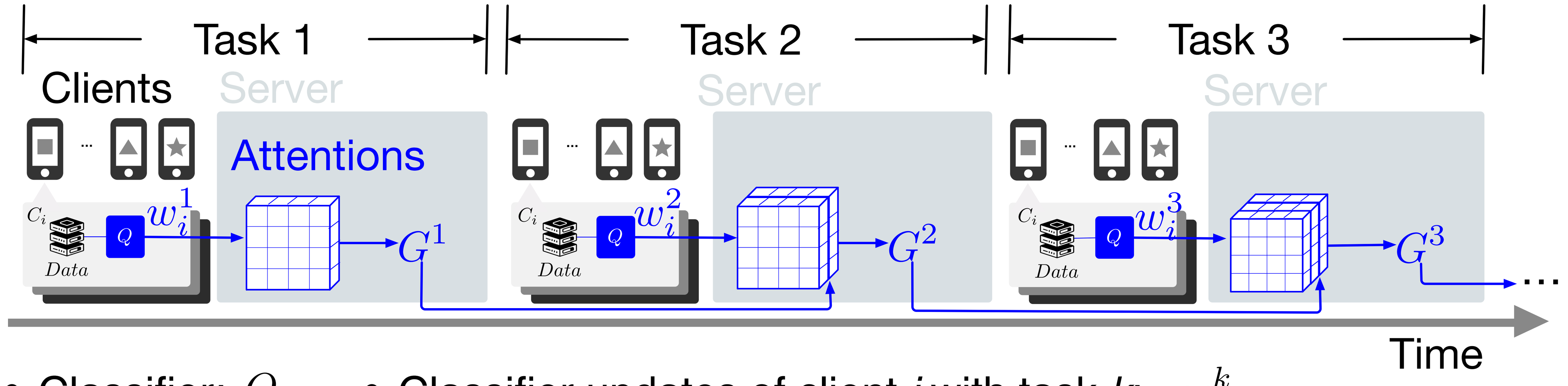
What auxiliary conditioning exists in FCL?

- Spatial and Temporal variation patterns.
- For example, patterns of client *i* of task *k* (represented as a graph row):

$$\mathbf{g}_i^k \triangleq G_{i,:}^k \in \mathbb{R}_+^N,$$

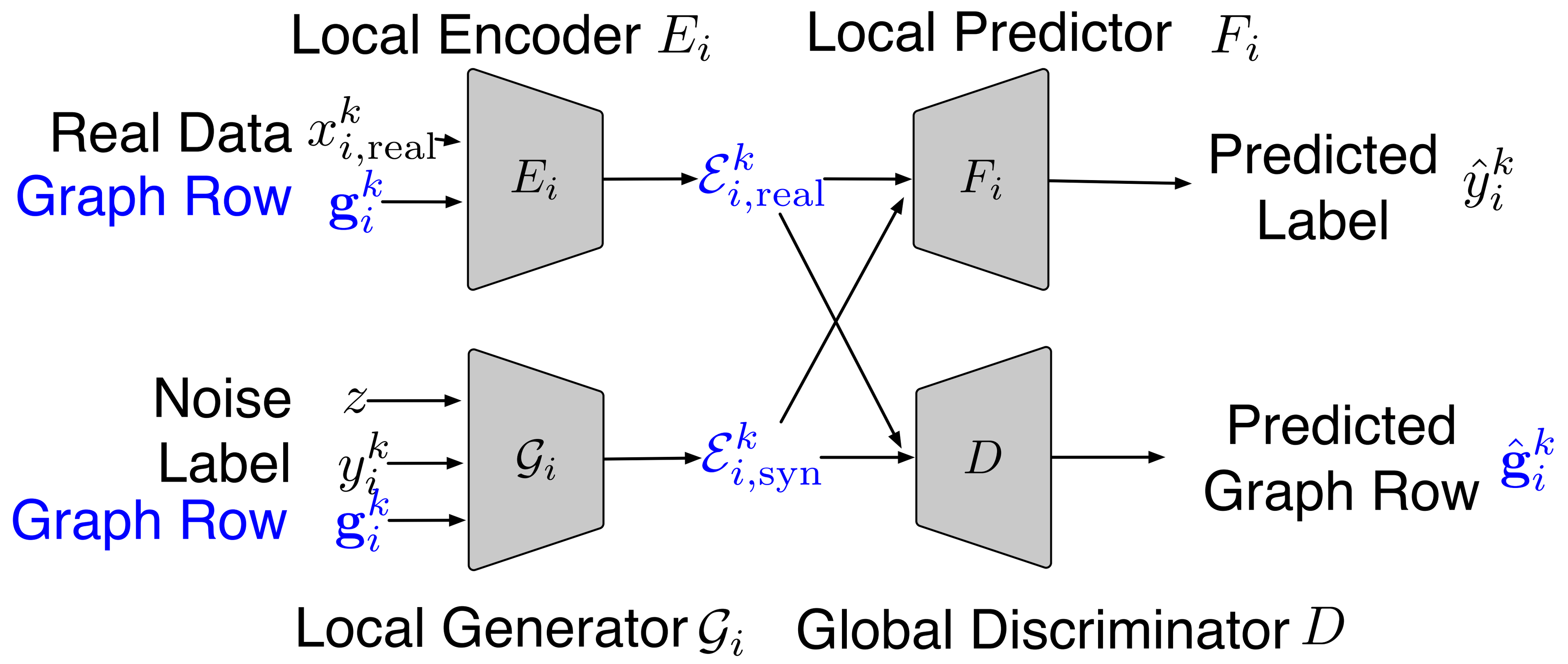
where $G_{ij}^k = \underbrace{\alpha_{ij}^k}_{\text{Spatial Attention}} \cdot \underbrace{\beta_{ij}^k}_{\text{Temporal Attention}} \in (0,1).$

2. Graph Generation



- Classifier: Q.
- Classifier updates of client *i* with task *k*: w_i^k.
- Spatial attention: $\alpha_{ij}^k = \mathrm{softmax}_j\left(a(w_i^k, w_j^k)\right) \in (0,1).$
- Temporal attention: $\beta_{ij}^k = \mathrm{softmax}_j\left(b(H_i^k, H_j^k, m)\right) \in (0,1)$, where $H_i^k = \{a_{i,:}^t\}_{t=1}^k$.

3. GFedCL



- Graph serves as conditioning in the adversarial training.

4. Representation Alignment between Real and Synthetic Data Embeddings

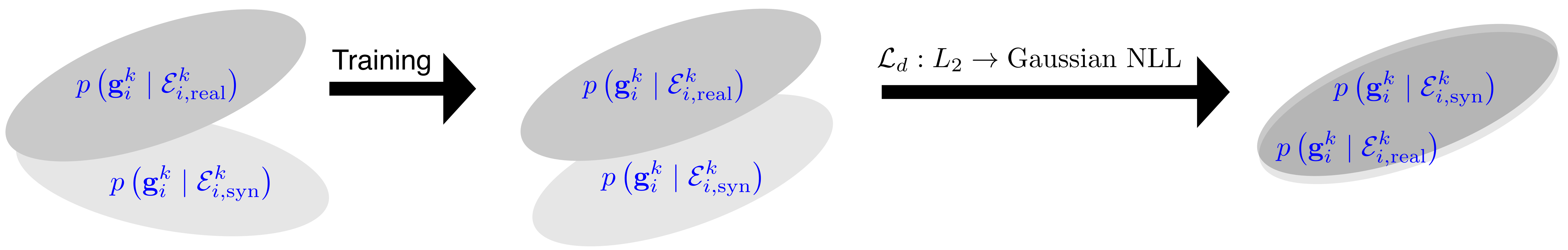
Before Training

After Training

Mean Alignment: $\forall i \in [N], \forall k \in [K],$

$$\mathbb{E}[\mathbf{g}_i^k | \mathcal{E}_{i,\mathrm{real}}^k] = \mathbb{E}[\mathbf{g}_i^k | \mathcal{E}_{i,\mathrm{syn}}^k] = \mathbb{E}[\mathbf{g}_i^k] = \mathbf{g}^*.$$

Mean and Variance Alignment

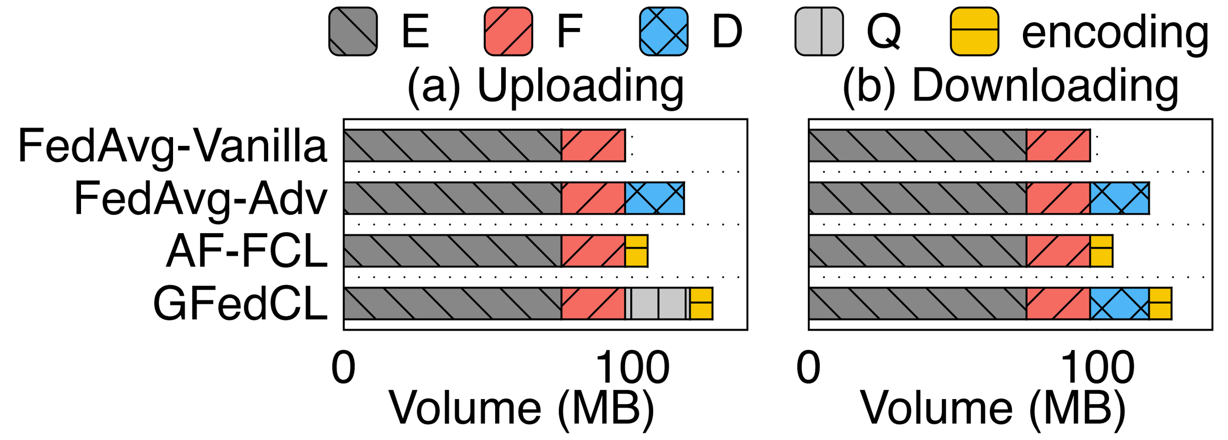


- Representation alignment alleviate catastrophic forgetting and data heterogeneity simultaneously.

6. Evaluation

Comm. Overhead

Main Results



Method	EMNIST-Letters		CIFAR-100		TinyImageNet	
	Acc (%)↑	Forgetting (%)↓	Acc (%)↑	Forgetting (%)↓	Acc (%)↑	Forgetting (%)↓
GLFC	43.47 ± 2.14	8.74 ± 2.78	46.08 ± 2.70	5.66 ± 1.74	42.93 ± 2.88	7.64 ± 2.91
FedAvg-Adv	25.67 ± 4.49	9.77 ± 3.17	32.05 ± 4.51	8.92 ± 2.01	25.41 ± 6.55	9.74 ± 3.29
FedCIL	38.95 ± 4.53	8.87 ± 2.31	39.47 ± 3.24	6.54 ± 1.72	45.75 ± 5.26	7.51 ± 1.33
MFCL	52.66 ± 3.24	6.76 ± 1.24	37.09 ± 3.76	5.26 ± 1.87	47.82 ± 3.09	8.11 ± 1.07
AF-FCL	41.22 ± 4.51	6.58 ± 1.44	43.27 ± 4.77	5.87 ± 2.35	54.91 ± 3.72	7.84 ± 2.02
C^2Prompt	68.47 ± 3.81	7.42 ± 1.80	<u>67.93</u> ± 5.19	6.94 ± 2.77	<u>72.41</u> ± 4.53	6.72 ± 1.84
AFL	<u>75.17</u> ± 1.31	**2.94** ± 0.79	55.33 ± 1.55	**3.47** ± 0.85	49.44 ± 1.35	**3.95** ± 1.59
GFedCL	**77.23** ± 2.76	<u>3.34</u> ± 1.21	**73.67** ± 3.42	<u>4.24</u> ± 1.45	**77.39** ± 3.02	<u>5.57</u> ± 1.88

The work was supported in part by the United States National Science Foundation under Grants 2523997, 2315612, and 2332638, and bythe AWS Cloud Credit for Research program, and by the Natural Sciences and Engineering Research Council of Canada ALLRP 588144-23.

NSF aws NSERC CRSNG