Beihang University 北京航空航天大學 | Universidade de Macau 澳門大學 University of Macau | CARS 铁科院 | KDD2025 August 3-7, 2025

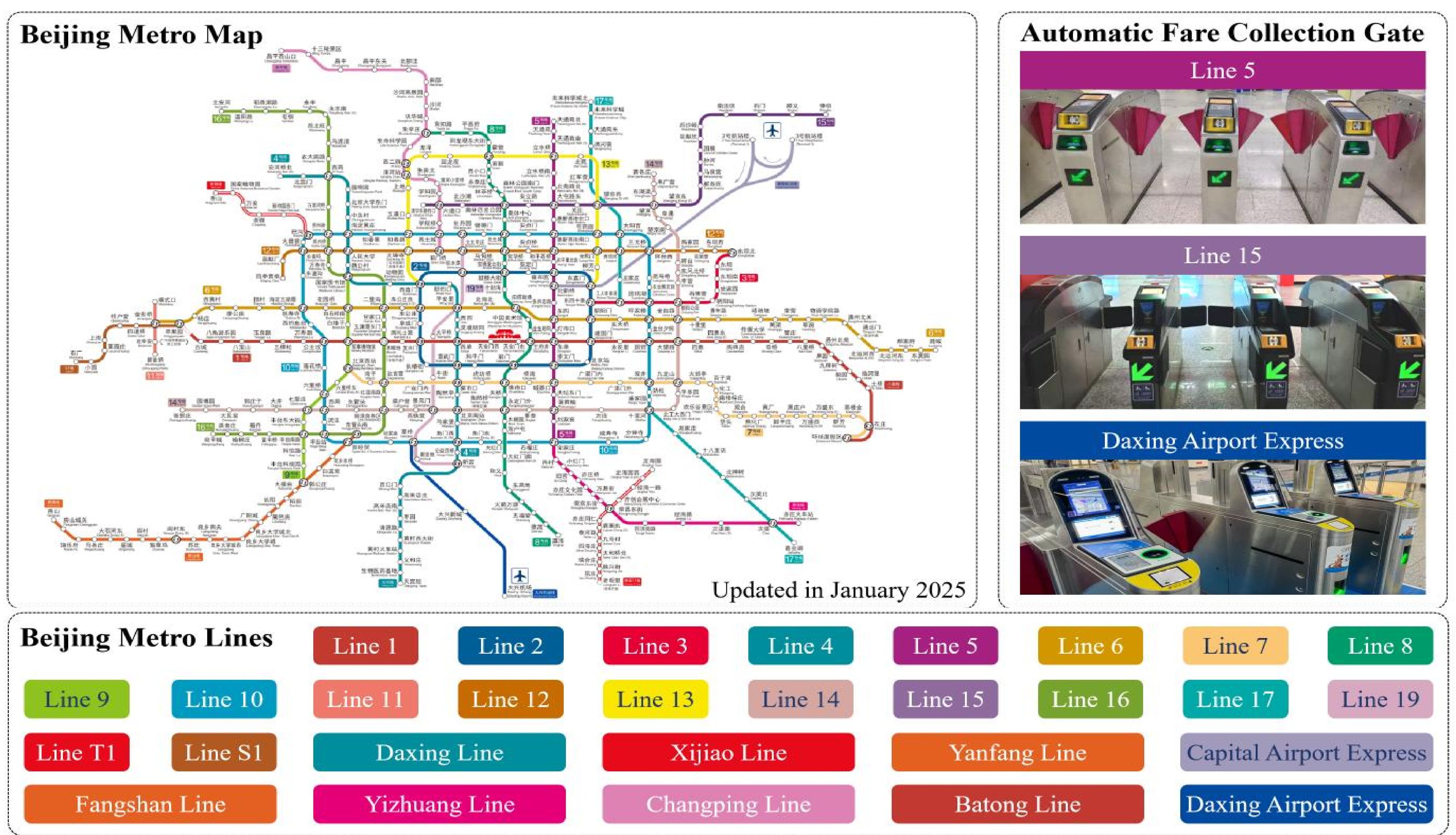
FedMetro: Efficient Metro Passenger Flow Prediction via Federated Graph Learning

Tianlong Zhang[1], Xiaoxi He[2], Yuxiang Wang[1], Yi Xu[1], Rendi Wu[1], Zhifei Wang[3], Yongxin Tong[1]

[1] Beihang University, [2] University of Macao, [3] China Academy of Railway Sciences

Background

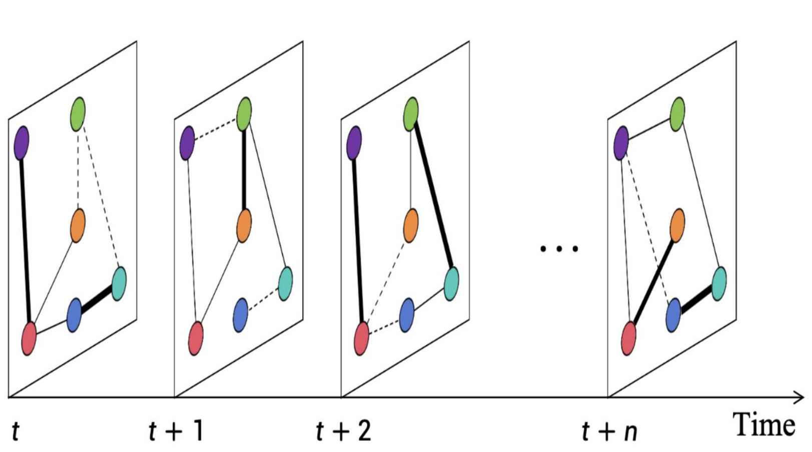
Metro passenger flow prediction plays a crucial role in urban transportation management, enabling applications such as station congestion alerts, timetable optimization, and transportation recommendations. Accurate metro passenger flow prediction relies heavily on Spatial-Temporal Graph Neural Networks (STGNNs), which leverage the spatial-temporal patterns embedded in Automatic Fare Collection (AFC) data to model complex correlations across metro networks.



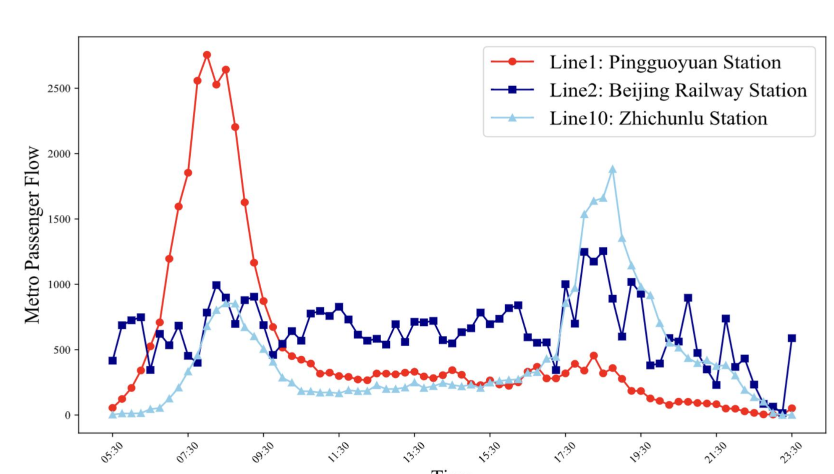
Challenges

Federated graph learning has emerged as a promising paradigm for privacy-preserving training of STGNNs, enabling collaborative modeling of spatial-temporal correlations. However, existing federated graph learning approaches are not directly applicable to the accurate prediction of metro passenger flow, as they face the following three significant challenges.

- **Time-evolving Spatial Correlations**
- **Heterogeneous Temporal Correlations**
- **Communication Bottlenecks in Federated Inference**



Time-evolving spatial correlations

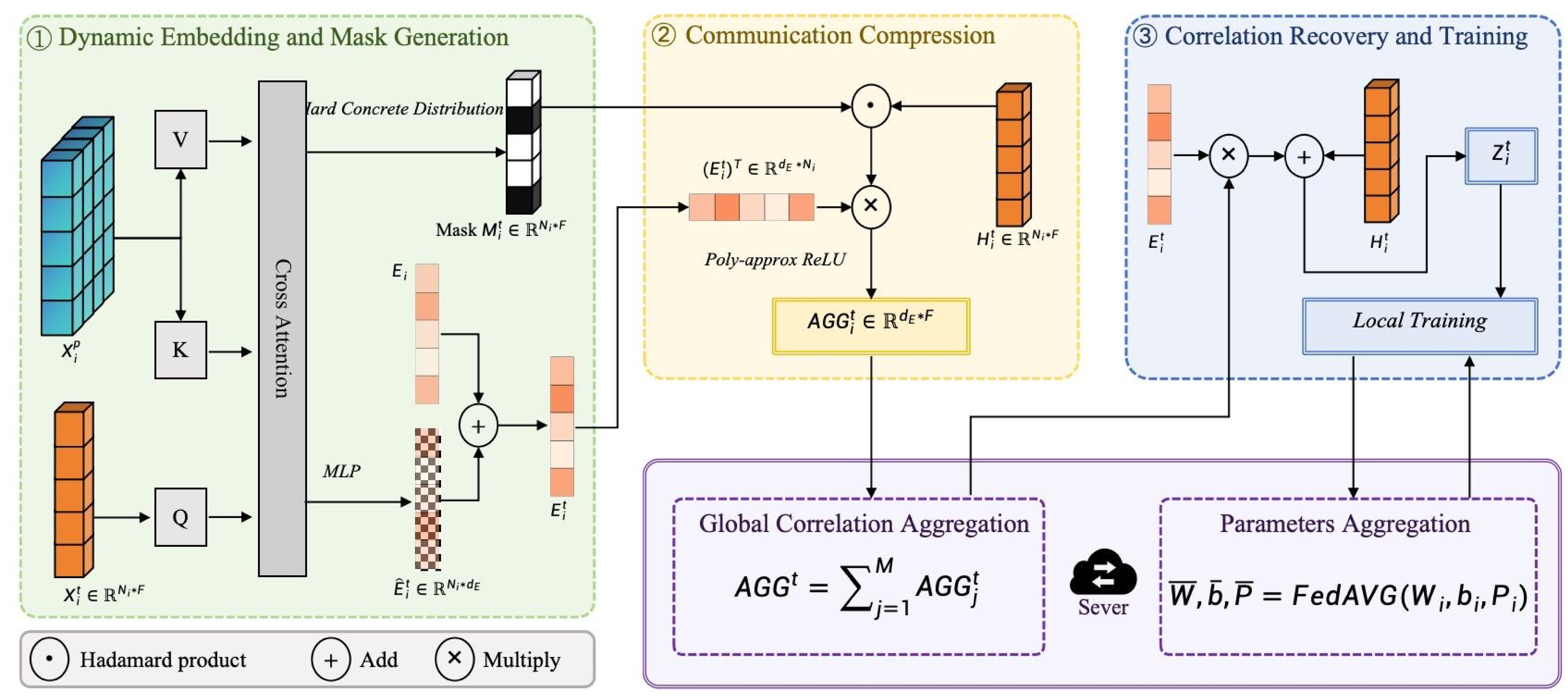


Heterogeneous temporal correlations

Design of FedMetro

We propose a federated dynamic graph learning method that employs cross-attention to generate dynamic node embeddings for each station, capturing the evolving passenger flow spatial-temporal correlations without sharing raw AFC data, ensuring compliance with privacy regulations.This method models the global metro passenger flow correlations, overcoming the challenges of time-evolving spatial correlations and the heterogeneity of temporal correlations, thereby improving the accuracy of metro passenger flow predictions. Furthermore, we significantly reduce communication overhead by dynamically sparsifying the spatial correlation graph, while maintaining high prediction accuracy and effectively addressing communication bottlenecks in federated inference. We use a client-server federated learning framework. The server is responsible for global correlations and training parameters aggregation, while the client consists of three modules:

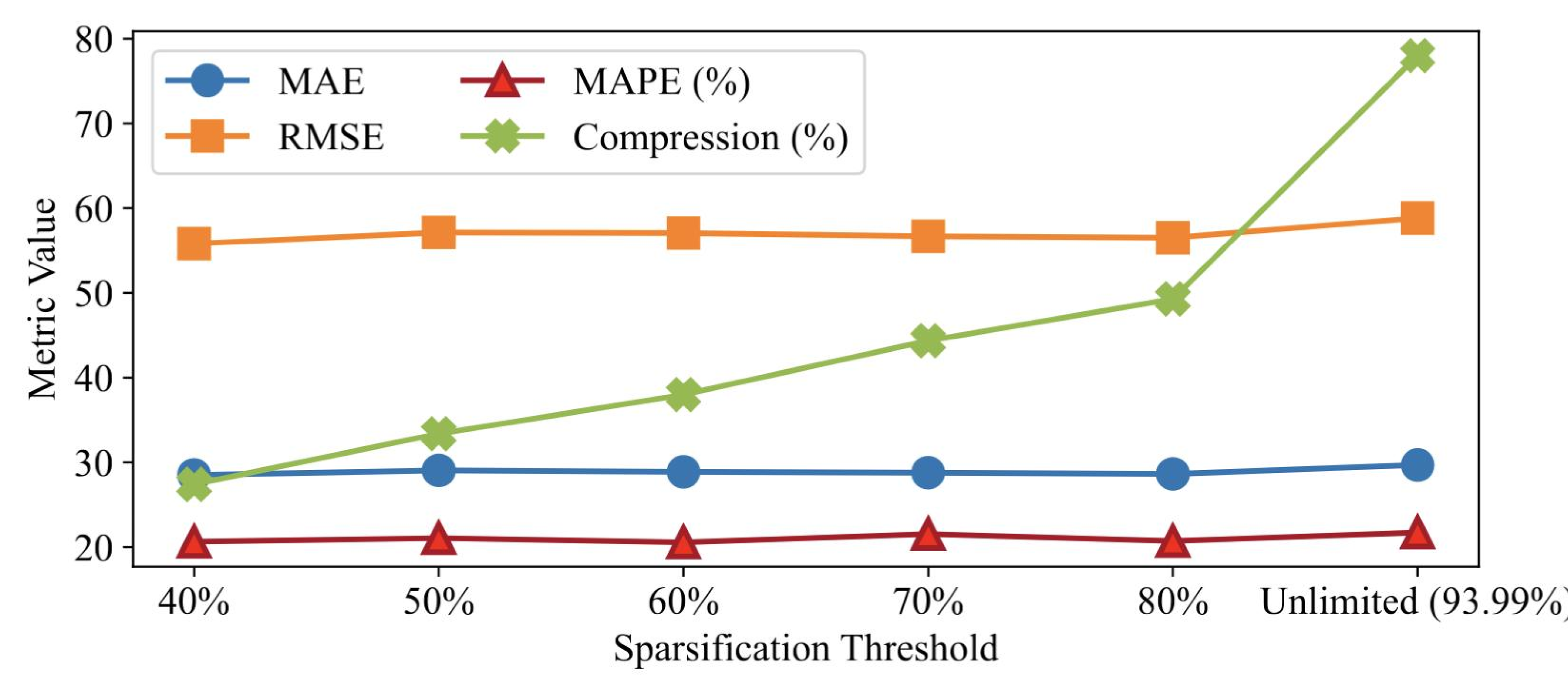
- **Dynamic Embedding and Mask Generation**
- **Communication Compression**
- **Correlations Recovery and Training**



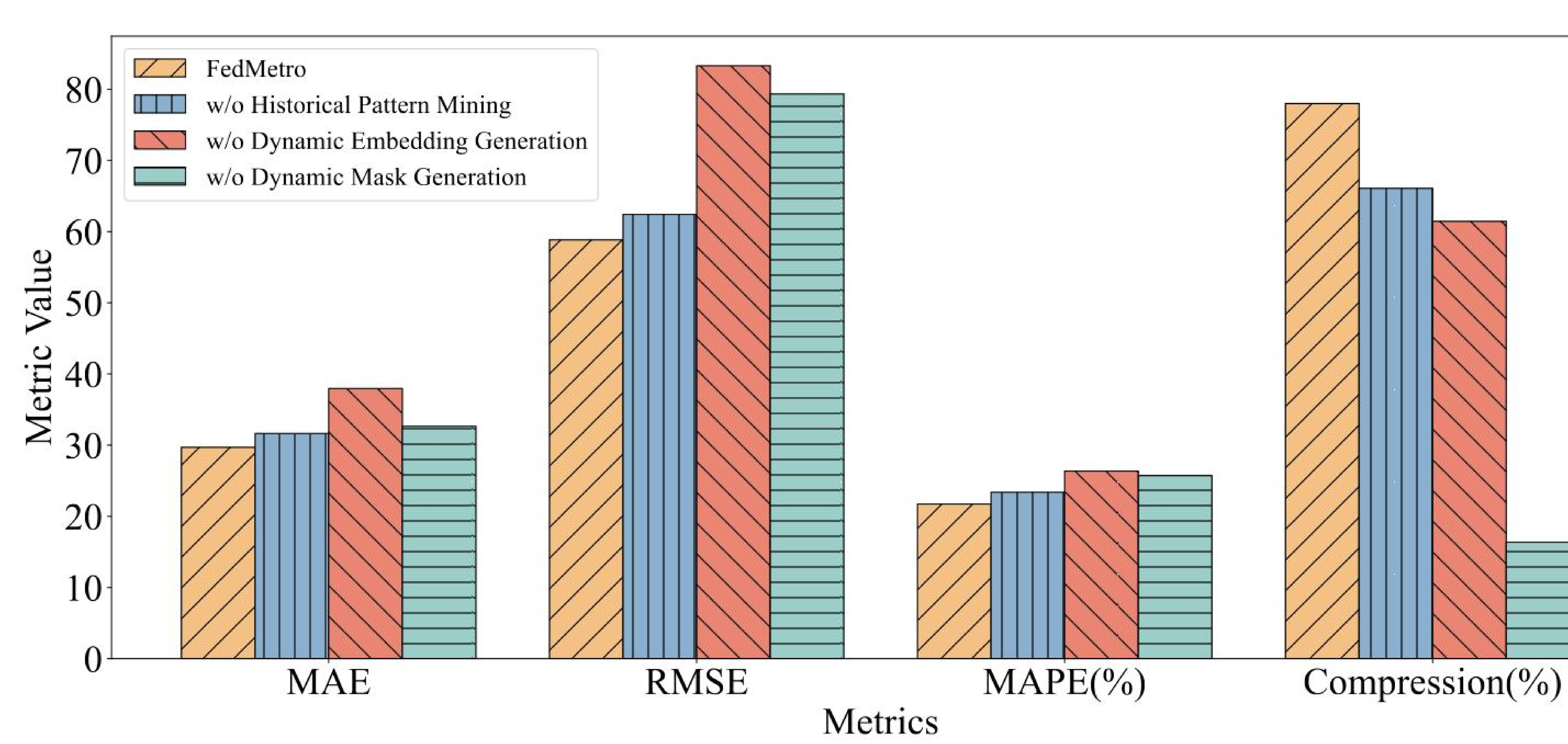
FedMetro System Overview

Results

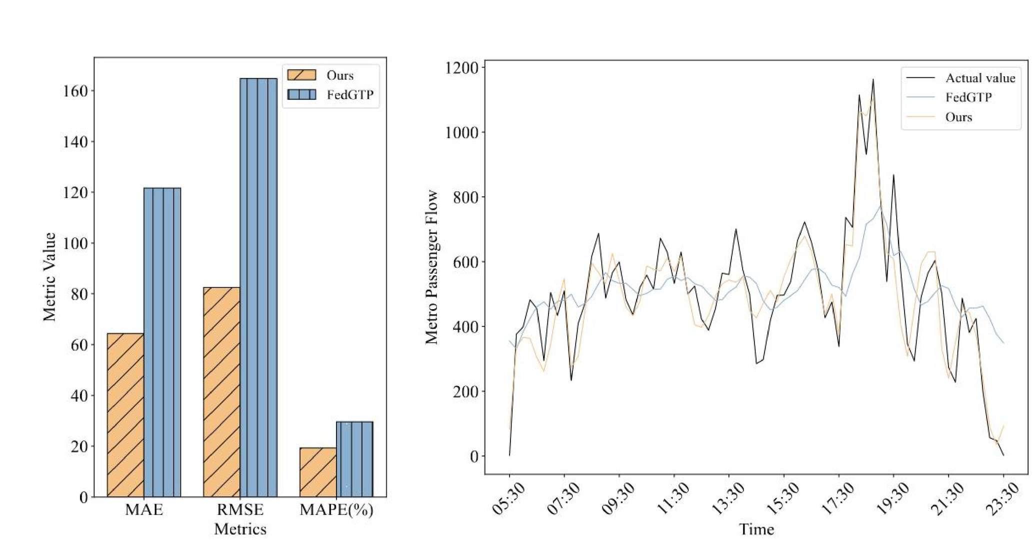
Evaluations on three real-world metro AFC datasets demonstrate that FedMetro significantly outperforms baseline methods, achieving up to **17.08%** higher accuracy while reducing federated inference communication overhead by **77.99%**. Practical deployments further confirm its effectiveness in delivering accurate station-level predictions across metro lines.



The results of communication compression study



Ablation study on BJMetro



(a) Comparison of metric value (b) Comparison of predicted and actual value

Comparison of deployment performance

Visualization of prediction results

Big Data Analysis Group

yxtong@buaa.edu.cn
tianlong@buaa.edu.cn

https://github.com/AlexMufeng/FedMetro