

Dual-BERT: A Multimodal Representation Learning Framework for Demographic-Behavior Alignment in Sparse Mobility Data

Brian Yueshuai He

Assistant Professor

University of Louisville

Abstract: Urban mobility systems are increasingly complex and challenging to model. Location-Based Services (LBS) data offer a promising alternative but suffer from key limitations, including sparse observations and missing demographic context. This study develops Dual-BERT, a dual-stream representation learning framework that jointly models activity sequences and demographic attributes using National Household Travel Survey data, addressing two fundamental challenges in LBS: reconstructing incomplete activity sequences and uncovering demographic-mobility associations. The methodology employs a two-stage training approach, combining masked language modeling (MLM) pre-training followed by contrastive alignment fine-tuning, enabling the model to learn rich representations capturing both temporal activity patterns and their demographic context. Validation shows that Dual-BERT consistently outperforms single-modal baselines, achieving performance comparable to ideal cases with complete information. The dual-stream architecture successfully reconstructs routine activities (Home, Education, Work) with F1 scores of 0.90–0.94 and predicts activity durations with around 10% error. Demographic inference demonstrates the model's ability to recover real demographic distributions across multiple behavioral groups, achieving near-perfect alignment for attributes such as Urban/Rural location, Gender, and Education, while Household Size and Occupation remain more challenging. This framework provides transportation researchers and urban planners with data-driven tools to mitigate demographic bias and address data sparsity in mobility modeling.

Bio: Dr. Brian Yueshuai He is an assistant professor at the University of Louisville with research expertise in transportation system modeling, human mobility data mining and generation, and artificial intelligence in transportation. His work focuses on developing advanced deep learning and computational simulation methods to better understand, model, and predict human mobility and transportation system dynamics. He earned his Ph.D. in Transportation Planning and Engineering from New York University in 2020 and previously served as a research scientist at the University of California, Los Angeles before joining UofL.

108 SN

January 15, 2026

12pm - 1pm EST

