

DynaOD: Dynamic Origin-Destination Flow Generation with Discrete-to-Continuous Temporal Semantic Modeling

Jie Zhao¹, Xianqi Dai², Jie Feng^{3*}, Huandong Wang¹ and Yong Li¹

¹Department of Electronic Engineering, BNRist, Tsinghua University

²Tsinghua Shenzhen International Graduate School

³Zhongguancun Academy

fengjie@bza.edu.cn, liyong07@tsinghua.edu.cn

Abstract

Dynamic origin-destination (OD) flow generation seeks to synthesize realistic mobility dynamics from temporal context alone, without relying on historical OD observations. A key challenge is to translate semantic temporal signals into temporally coherent OD patterns while preserving the inherent spatial heterogeneity of urban regions. We propose DynaOD, a semantic-driven framework that models temporal dynamics through two complementary perspectives: discrete directional trends that characterize qualitative shifts in urban activity patterns, and continuous temporal evolution that captures how such shifts unfold over time. By jointly encoding these temporal semantics, the framework constructs time-varying region representations that condition pretrained static OD generators in a lightweight and plug-and-play fashion. This modular design further supports scalable deployment and cross-city transferability. Extensive experiments on large-scale real-world datasets show that our method consistently outperforms representative baselines in both predictive accuracy and distributional fidelity. Code is publicly available at <https://github.com/csjiezhao/DynaOD>.

1 Introduction

Origin-destination (OD) flows describe collective mobility patterns between geographical regions and constitute a fundamental representation for urban spatial analysis [Andrienko *et al.*, 2016]. Accurate OD flows support a wide range of downstream applications, including mobility simulation [Zhang *et al.*, 2025], urban planning [Miller, 2018], and policy evaluation [van Wee *et al.*, 2023]. When fine-grained trajectory data are unavailable, synthesized OD flows offer an effective alternative for characterizing large-scale urban mobility dynamics [Zhang *et al.*, 2020; Li *et al.*, 2021]. However, urban mobility exhibits complex temporal dynamics shaped by recurring human routines as well as stochastic external events [Barroso *et al.*, 2020]. To capture such variability, dynamic OD flow generation aims to synthesize context-aware

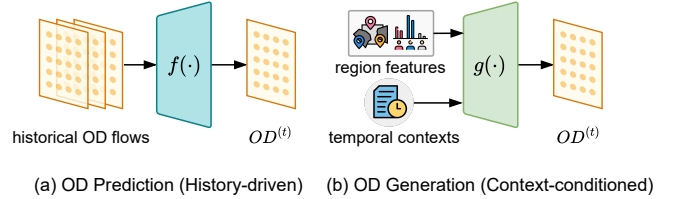


Figure 1: Comparison between history-driven OD prediction and context-conditioned OD generation. OD prediction extrapolates future flows from historical OD flows, whereas OD generation synthesizes flows directly from region attributes and temporal contexts.

OD flows without relying on historical OD observations or real-time traffic measurements. As illustrated in Figure 1, the task is formulated as a conditional generation problem grounded solely in regional semantics and temporal context.

Most existing OD generation methods focus on static OD matrices and primarily model spatial regularities from regional attributes without relying on historical OD observations [Rong *et al.*, 2024a]. However, such methods are inherently limited to time-aggregated settings and cannot capture evolving temporal mobility dynamics. Extending OD generation to dynamic temporal scenarios therefore remains largely underexplored [Rong *et al.*, 2024b; Yean *et al.*, 2025]. Existing efforts typically rely on real-time traffic observations or historical OD sequences [Bauer *et al.*, 2017], which restrict their applicability in cold-start and cross-city settings [Fang *et al.*, 2022; Li *et al.*, 2025]. Despite recent progress, dynamic OD generation under minimal supervision remains fundamentally challenging for three reasons.

First, temporal context influences OD flows through latent behavioral dynamics rather than explicit flow observations. These effects are often implicit and emerge only after spatial aggregation, making them difficult to capture from limited supervision. *Second*, OD flows exhibit strong spatial heterogeneity that varies across temporal contexts. Regions with similar functional roles may respond similarly under certain contexts yet diverge substantially under others, complicating the context-aware modeling of regional mobility dynamics. *Third*, achieving scalable dynamic OD generation across heterogeneous cities and temporal settings remains difficult. Practical systems must generalize without city-specific retraining while maintaining efficient deployment and low in-

*Corresponding author.

ference overhead.

To address these challenges, we propose **DynaOD**, a modular framework that models temporal context through semantic temporal abstractions. Specifically, DynaOD disentangles temporal semantics into two complementary components: discrete semantic trends that capture coarse-grained transitions in urban activity patterns, and continuous temporal dynamics that characterize how these transitions evolve over time. By jointly encoding these temporal dynamics, the framework constructs time-varying region features that condition pretrained OD generators for dynamic flow synthesis. A key advantage of this design is that temporal adaptation can be achieved without modifying the underlying OD generators, enabling seamless integration with diverse pretrained models. In addition, DynaOD incorporates a retrieval-enhanced mechanism to improve cross-city generalization, together with a lightweight distillation strategy for efficient and scalable deployment.

Our contributions are summarized as follows:

- We propose a semantic-driven framework that disentangles discrete temporal trends and continuous temporal dynamics for context-aware dynamic OD generation under minimal supervision.
- We design a retrieval-enhanced regional adaptation mechanism that enables the reuse of regional knowledge across temporal contexts, improving cross-city generalization and generation fidelity.
- We develop a lightweight plug-in architecture compatible with diverse pretrained OD generators, together with an efficient distillation strategy for scalable deployment.

2 Preliminaries

Definition 1 (Region). *An urban area is partitioned into a set of spatial regions $\mathcal{R} = \{r_i\}_{i=1}^N$, where each region corresponds to a basic spatial unit such as a census tract or administrative zone.*

Definition 2 (POI Distribution). *Each region $r_i \in \mathcal{R}$ is associated with a POI distribution vector $\mathbf{p}_i \in \mathbb{R}^{D_p}$ describing the composition of POI categories within the region.*

Definition 3 (Demographic Attributes). *Each region $r_i \in \mathcal{R}$ is associated with a demographic feature vector $\mathbf{d}_i \in \mathbb{R}^{D_d}$ containing demographic and socio-economic statistics.*

Definition 4 (Dynamic OD Flow Sequence). *Given a temporal horizon of length T , a dynamic OD flow sequence is represented as an ordered set of OD matrices $\{\mathbf{M}^{(t)}\}_{t=1}^T$, where $\mathbf{M}^{(t)} \in \mathbb{R}^{N \times N}$ denotes the OD flow matrix at time step t .*

Problem Statement. Given regions $\mathcal{R} = \{r_i\}_{i=1}^N$ with associated POI distributions $\mathbf{p}_i$ and demographic features $\mathbf{d}_i$, together with a temporal context sequence $\mathbf{c}_{1:T}$, the goal of dynamic OD flow generation is to synthesize a sequence of OD matrices $\{\mathbf{M}^{(t)}\}_{t=1}^T$ that captures realistic urban mobility dynamics under the specified temporal context.

3 Methodology

As illustrated in Figure 2, DynaOD leverages a large language model (LLM) to infer semantic directional trends from tem-

poral context and converts them into continuous temporal feature dynamics via a shape generator. The resulting features are then used to condition a pretrained static OD generator at each time step, enabling dynamic OD flow synthesis. In addition, the framework supports optional cross-city shape composition and LLM distillation to enhance generalization and deployment efficiency.

3.1 Discrete Semantic Modeling of Temporal Context

Dynamic OD generation requires modeling how urban functions and population activities respond to temporal context at a semantic level, rather than directly predicting numerical changes. Temporal cues such as weekdays or holidays often induce qualitative shifts in mobility behaviors [Grinberger and Shoval, 2019; Sparks *et al.*, 2022], whose effects on region attributes are indirect and difficult to capture with continuous predictors. Moreover, such directional controls do not have explicit ground-truth supervision and must instead be inferred from contextual semantics and commonsense priors. LLMs are well suited for this task because they encode rich world knowledge and commonsense reasoning capabilities. We therefore introduce a *directional control* mechanism that leverages LLMs to infer coarse-grained semantic trends in urban dynamics.

Discrete directional controls. For each region and date, the LLM generates two discrete control vectors: (i) a POI control vector $\mathbf{c}^{\text{poi}} \in \{-1, 0, 1\}^{D_p}$ and (ii) a demographic control vector $\mathbf{c}^{\text{demo}} \in \{-1, 0, 1\}^{D_d}$. Each element indicates whether the corresponding feature is expected to increase, remain stable, or decrease under the given spatio-temporal context. Rather than directly conditioning OD flows, these controls encode qualitative directional trends in urban dynamics. The POI control is inferred from regional context and temporal information, while the demographic control is further conditioned on inferred POI trends and demographic attributes, enabling dependency-aware modeling of urban functions and population dynamics.

Design rationale. This discrete formulation offers several advantages. It aligns naturally with commonsense reasoning, separates semantic trend inference from continuous magnitude modeling, and produces compact interpretable controls that generalize effectively across cities. Within DynaOD, the inferred directional controls are translated into continuous temporal feature dynamics through a shape generator, enabling pretrained static OD generators to adapt to diverse temporal contexts without retraining.

3.2 Continuous Temporal Evolution of Region Features

While directional control specifies *what* should change under a given temporal context, dynamic OD generation further requires modeling *how much* region attributes vary and how such changes evolve over time. We therefore introduce **ShapeNet**, a differentiable shape generator that translates discrete directional controls into continuous time-varying feature dynamics.

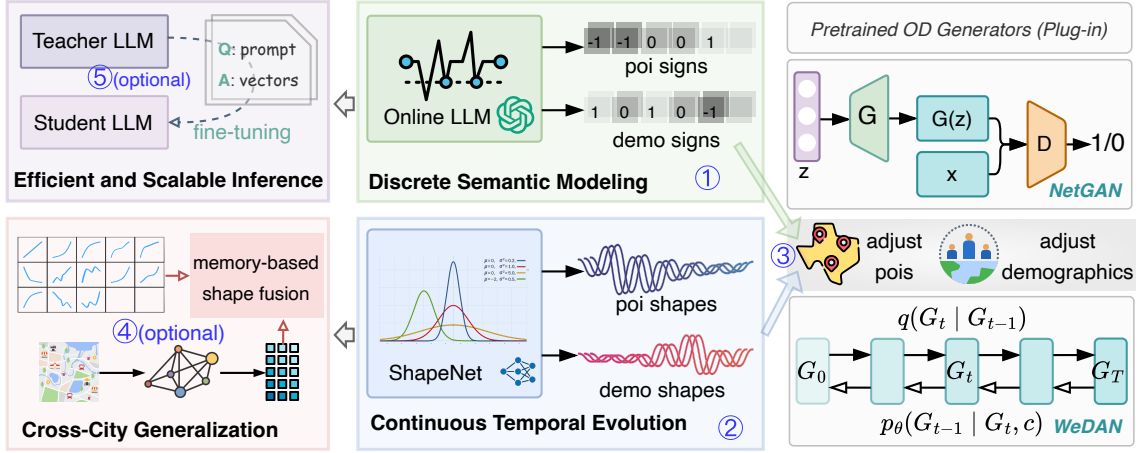


Figure 2: Architecture of DynaOD. (1) An LLM infers discrete directional signals from temporal context; (2) ShapeNet converts them into continuous multi-day feature evolution; (3) the resulting features modulate a pretrained OD generator stepwise to produce dynamic OD flows. Optional components include (4) retrieval-based shape composition for cross-city generalization and (5) LLM distillation for efficient inference.

Inputs and outputs. For each region i and day t within a window of length T , ShapeNet takes as input the static region features $\mathbf{x}_i = [\mathbf{x}_i^{\text{demo}}, \mathbf{x}_i^{\text{poi}}]$, directional controls $\mathbf{c}_{i,t}^{\text{poi}}$ and $\mathbf{c}_{i,t}^{\text{demo}}$, and a time-context encoding $\mathbf{u}_{i,t}$. It outputs Gaussian parameters for POI and demographic feature dynamics:

$$(\boldsymbol{\mu}_{i,t}^{\text{poi}}, \log \sigma_{i,t}^{2,\text{poi}}) \in \mathbb{R}^{D_p}, \quad (\boldsymbol{\mu}_{i,t}^{\text{demo}}, \log \sigma_{i,t}^{2,\text{demo}}) \in \mathbb{R}^{D_d}. \quad (1)$$

Shape sequences are sampled through the reparameterization trick and aggregated to obtain continuous temporal shapes $\mathbf{s}_{i,t}^{\text{poi}}$ and $\mathbf{s}_{i,t}^{\text{demo}}$.

Architecture. ShapeNet adopts a control- and time-dominant architecture, where static region features are used only for weak conditioning. Specifically, POI controls and demographic controls are first combined with their corresponding activation masks and time-context encodings to form per-day input tokens. These token sequences are independently processed by lightweight temporal encoders to capture day-to-day dependencies, and the resulting hidden representations are subsequently fused into shared temporal embeddings. To incorporate regional heterogeneity without overfitting to city-specific identities, ShapeNet further applies lightweight FiLM [Perez *et al.*, 2018] conditioning based on compact region codes derived from static region features.

Feature modulation. Given continuous shapes and discrete controls, DynaOD constructs time-varying region features via multiplicative modulation:

$$\begin{aligned} \tilde{\mathbf{x}}_{i,t}^{\text{poi}} &= \mathbf{x}_i^{\text{poi}} \odot \left(1 + \mathbf{s}_{i,t}^{\text{poi}} \odot \mathbf{c}_{i,t}^{\text{poi}}\right), \\ \tilde{\mathbf{x}}_{i,t}^{\text{demo}} &= \mathbf{x}_i^{\text{demo}} \odot \left(1 + \mathbf{s}_{i,t}^{\text{demo}} \odot \mathbf{c}_{i,t}^{\text{demo}}\right). \end{aligned} \quad (2)$$

The modulated region features are then fed into a frozen pre-trained static OD generator to synthesize dynamic OD flows at each time step.

Training and regularization. ShapeNet is trained by minimizing the discrepancy between generated and ground-truth OD flows over the temporal window, while the pretrained OD generator remains frozen. To make optimization tractable, gradients are propagated only through the final DDIM denoising steps during training. We further apply a weighted temporal smoothness regularizer to suppress unrealistic day-to-day variations, with stronger penalties when the corresponding directional controls indicate stability.

3.3 Cross-City Generalization via Composable Shape Priors

Although ShapeNet enables continuous temporal feature evolution, directly applying it to unseen cities may suffer from out-of-distribution inference and increased computational overhead. To support cross-city generalization and scalable deployment, we introduce **ShapeMem**, a retrieval-based memory of composable shape priors that synthesizes region-level shapes without invoking ShapeNet during inference.

ShapeMem construction. ShapeNet produces Gaussian temporal shape parameters for POI and demographic features. We aggregate these outputs by weekday to construct region-weekday shape priors, where each memory entry stores Gaussian statistics that capture typical temporal dynamics and their variability. These aggregated priors provide a compact and reusable representation of learned temporal evolution patterns.

Region representation and retrieval. To retrieve relevant priors for a target tract, we learn region embeddings using a graph-based representation learning approach [Thakoor *et al.*, 2022]. During inference, cosine similarity in the embedding space is used to identify top-5 similar tracts from the memory.

Composable prior fusion. For a target tract and weekday, retrieved shape priors are fused via similarity-weighted aver-

aging to produce tract-level temporal shapes:

$$\tilde{s}_{i,w}^{\text{poi}} = \sum_{j \in \mathcal{N}_K(i)} \alpha_{i,j} s_{j,w}^{\text{poi}}, \quad \tilde{s}_{i,w}^{\text{demo}} = \sum_{j \in \mathcal{N}_K(i)} \alpha_{i,j} s_{j,w}^{\text{demo}}, \quad (3)$$

where $\alpha_{i,j}$ denotes normalized similarity weights. This retrieval-based composition serves as a non-parametric approximation of ShapeNet inference for unseen cities.

Plug-and-play inference. The composed shapes are injected into the feature modulation step and fed into a pre-trained OD generator. Since ShapeMem operates independently of the OD generator, this mechanism is fully plug-and-play and enables efficient dynamic OD generation across new cities and temporal contexts.

3.4 Efficient and Scalable Inference

Dynamic OD generation requires producing directional control vectors for many regions over extended temporal horizons. Although LLMs are effective at inferring such high-level semantic controls, repeatedly querying high-capacity models incurs substantial computational overhead. To improve scalability, we adopt a knowledge distillation strategy that transfers directional control inference from a strong teacher LLM to a lightweight student model for efficient deployment [Cheng *et al.*, 2025; Fang *et al.*, 2026].

Teacher-student distillation. A strong online LLM is first used as a teacher to generate high-quality directional control vectors for POI and demographic attributes. These outputs are treated as pseudo-labels to supervise a compact student model through supervised fine-tuning. In our implementation, we instantiate the student with Qwen2.5-1.5B-Instruct and fine-tune it using LoRA, updating only the query and value projection matrices. The student is trained on POI and demographic instruction data to reproduce the teacher’s discrete directional decisions under identical region-date queries, without requiring explicit semantic descriptions for individual feature dimensions.

Structured directional prediction. Although directional control fundamentally relies on semantic understanding and commonsense inference, the distilled student model is trained under a structured discrete prediction formulation for efficient supervised fine-tuning. Specifically, the student predicts fixed-length vectors with entries in $\{-1, 0, 1\}$ for POI and demographic controls. POI and demographic prediction are treated as two separate instruction formats during training, which simplifies optimization and improves prediction stability while enabling unified deployment.

Plug-and-play deployment. After distillation, the student model fully replaces the teacher during inference and serves as a lightweight directional controller. Since the controller is decoupled from both ShapeNet and the OD generator, it can be independently updated or replaced without affecting the remaining components. This modular design enables efficient and scalable dynamic OD generation in large-scale deployment scenarios.

4 Experiments

4.1 Experimental Setup

Datasets

We construct a tract-level dynamic OD flow dataset by integrating publicly available U.S. mobility flows [Kang *et al.*, 2020] with regional attributes from a large-scale commuting benchmark [Rong *et al.*, 2025]. The dataset spans 500 U.S. counties in January 2019, where counties are treated as urban instances and census tracts as spatial regions. Following [Rong *et al.*, 2025], each tract is associated with POI distributions over 34 categories and 97-dimensional demographic attributes. Daily OD matrices are aligned with tract-level attributes to ensure consistent spatial correspondence.

Each date is associated with an 8-dimensional temporal context vector consisting of a weekday one-hot encoding and a holiday indicator. To evaluate spatial and temporal generalization, we adopt a strict two-axis split protocol. Counties and dates are partitioned into disjoint *seen* and *unseen* sets. Models are trained only on seen counties and seen dates, and evaluated on unseen counties at unseen dates. Under this protocol, only *ShapeNet* is optimized, while the pretrained OD generator and directional control signals remain fixed.

Metrics

We evaluate OD generation performance using complementary metrics that capture accuracy, structural consistency, and distributional similarity. Specifically, **RMSE** and **NRMSE** measure element-wise errors, where NRMSE is normalized by the empirical variance of real flows. **CPC** (Common Part of Commuting) measures the overlap between predicted and ground-truth OD flows, reflecting structural fidelity. In addition, **Jensen Shannon Divergence (JSD)** is computed over inflow, outflow, and full OD flow distributions to assess global mobility pattern similarity.

Baselines

We compare DynaOD with a diverse set of representative baselines covering physics-inspired, statistical learning, deep learning, and generative models for OD flow modeling. The compared methods include: (1) *Physics-inspired models*: two variants of the gravity model, **GM-P** and **GM-E** [Zipf, 1946]; (2) *Statistical learning models*: **Random Forest (RF)** [Pourebrahim *et al.*, 2019], **Support Vector Regression (SVR)** [Rodríguez-Rueda *et al.*, 2021], and **Gradient Boosting Regression Tree (GBRT)** [Robinson and Dilkina, 2018]; (3) *Deep learning models*: the **Deep Gravity Model (DGM)** [Simini *et al.*, 2021]; (4) *Generative graph models*: **NetGAN** [Bojchevski *et al.*, 2018]; and (5) *Graph diffusion-based models*: **WeDAN** [Rong *et al.*, 2025].

Adapting baselines to dynamic settings. For fair comparison, we augment all methods with the same temporal context features used by DynaOD, concatenated to region-level inputs whenever applicable, without modifying their original architectures. For NetGAN and WeDAN, we follow standard training protocols and use pretrained models [Rong *et al.*, 2025], denoted as NetGAN_{pt} and WeDAN_{pt} . These models are directly applied at inference time under different temporal contexts without additional fine-tuning. Unless otherwise

stated, DynaOD adopts WeDAN_{pt} as its underlying OD generator. To demonstrate modularity, we additionally replace the generator with NetGAN_{pt} in subsequent experiments.

4.2 Performance Comparison

Model	CPC $\uparrow$	RMSE $\downarrow$	NRMSE $\downarrow$	JSD-In $\downarrow$	JSD-Out $\downarrow$	JSD-OD $\downarrow$
GM-P	0.181	317.11	1.161	0.722	0.834	0.662
GM-E	0.179	309.23	1.163	0.731	0.838	0.671
RF	0.219	304.35	1.106	0.691	0.798	0.513
SVR	0.206	312.98	1.143	0.706	0.780	0.444
GBRT	0.196	310.41	1.127	0.720	0.806	0.550
DGM	0.214	312.29	1.133	0.747	0.827	0.502
NetGAN_{pt}	0.365	293.42	1.129	0.657	0.736	0.574
WeDAN_{pt}	0.331	<u>286.51</u>	<u>1.056</u>	<u>0.502</u>	0.741	<u>0.333</u>
DynaOD	0.492	253.09	0.940	0.322	0.421	0.220
Improve	34.7%	11.7%	11.0%	35.9%	42.8%	33.9%

Table 1: Performance comparison on the test set. $\uparrow(\downarrow)$ indicates that higher (lower) values are better. **Bold** and underline denote the best and second-best results, respectively.

Table 1 reports the performance comparison under the challenging setting of unseen cities and unseen dates. DynaOD consistently outperforms all baselines across every evaluation metric, demonstrating strong robustness to both spatial and temporal distribution shifts. Classical gravity-based and statistical learning methods exhibit limited performance, reflecting the difficulty of modeling dynamic urban mobility under temporally invariant assumptions. Learning-based regressors and DGM provide moderate improvements, yet still struggle to capture evolving temporal mobility patterns. Among generative baselines, pretrained NetGAN and WeDAN remain competitive due to their strong graph-based generation capabilities. Nevertheless, their reliance on static OD distributions limits generalization to dynamic scenarios with unseen spatial-temporal contexts. In contrast, DynaOD improves CPC by **34.7%** over the strongest baseline and reduces distributional divergence by up to **42.8%**, indicating substantially better alignment with real-world mobility dynamics. Overall, the results validate the effectiveness of semantic temporal control and continuous temporal adaptation in dynamic OD flow generation.

4.3 Ablation Study

Effectiveness of Directional Control

Figure 3 reports the ablation results on directional control in DynaOD. Removing all directional controls (*w/o ctrl*) leads to substantial degradation across every metric, suggesting that temporal evolution alone is insufficient to model dynamic mobility patterns without semantic guidance.

Introducing either POI control or demographic control consistently improves performance over the no-control variant, demonstrating that both contribute meaningful temporal conditioning signals. POI control mainly enhances reconstruction accuracy, whereas demographic control more effectively reduces distributional discrepancies in inflow and outflow patterns. When both controls are jointly incorporated, the

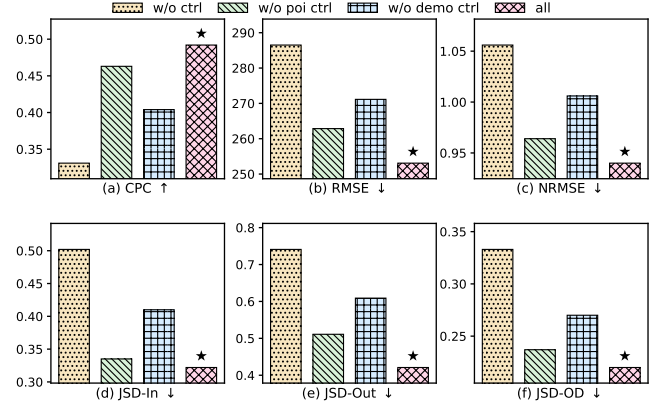


Figure 3: Ablation results on directional control. The star-shaped marker denotes the best performance among all variants.

model achieves the best overall performance across all metrics. The consistent gains over single-control variants indicate that urban functional dynamics and population behavior dynamics provide complementary information for modeling temporal mobility variations.

We further examine whether the effectiveness of directional control stems from semantic reasoning or merely from the discrete control formulation. To this end, we replace the LLM-based controller with a lightweight MLP classifier, resulting in the variant DynaOD_{CLS}. This replacement consistently reduces both reconstruction accuracy and distributional fidelity, with CPC decreasing from 0.492 to 0.413 and JSD-OD increasing from 0.220 to 0.274. These findings indicate that directional control benefits not only from discrete label prediction, but also from the semantic priors provided by LLMs. By contrast, the classifier-based variant relies solely on downstream OD supervision, limiting its ability to generalize across unseen spatial-temporal contexts.

Effectiveness of Shape Modeling

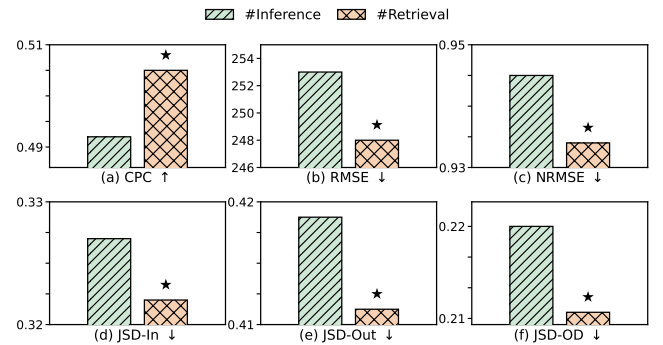


Figure 4: Ablation study on the effectiveness of shape modeling.

Figure 4 compares two strategies for continuous shape construction in DynaOD: direct ShapeNet inference (*#Inference*) and retrieval-based composition using similar regions (*#Retrieval*). Across all metrics, *#Retrieval* consistently achieves

better performance than *#Inference*. The improvements highlight two benefits of retrieval-based shape composition. First, retrieving shapes from similar regions yields more robust temporal dynamics under unseen spatial-temporal settings. Second, *#Retrieval* removes the need for ShapeNet inference during testing, substantially improving deployment efficiency and scalability. These results demonstrate that composable shape priors provide a scalable solution for cross-city dynamic OD generation.

4.4 Efficiency and Modularity Analysis

LLM Inference Overhead

We evaluate the computational overhead of LLM-based directional control using gpt-4o-mini. On average, a POI control query requires 1.94s with 415 input tokens and 110 output tokens (approximately \$0.00013), while a demographic control query requires 4.51s with 1436 input tokens and 287 output tokens (approximately \$0.00039). Despite the additional LLM inference cost, directional control generation remains practical because region-level queries are highly parallelizable. Furthermore, the proposed distillation strategy substantially reduces deployment overhead by replacing online LLM inference with a lightweight student model during testing.

Effectiveness of LLM Distillation

LLM	CPC $\uparrow$	RMSE $\downarrow$	NRMSE $\downarrow$	JSD-In $\downarrow$	JSD-Out $\downarrow$	JSD-OD $\downarrow$
Qwen-2.5-1.5B	0.403	274.43	1.010	0.411	0.637	0.276
Qwen-2.5-1.5B-SFT	0.476	260.86	0.959	0.336	0.475	0.234
GPT-4o-mini	0.492	253.09	0.940	0.322	0.421	0.220

Table 2: Performance comparison of DynaOD under different directional control generators (e.g., LLMs).

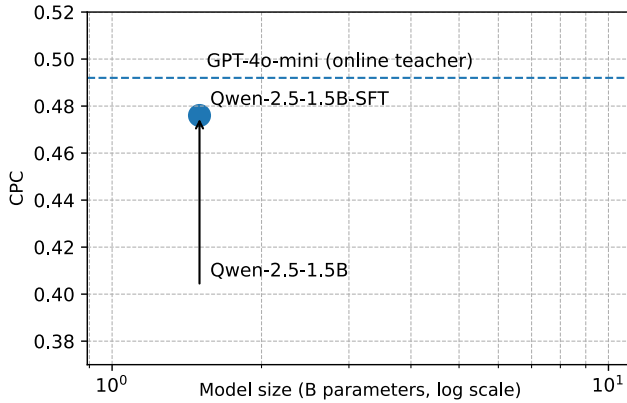


Figure 5: Performance-cost trade-off of directional control models.

Table 2 compares DynaOD under different directional controllers, including GPT-4o-mini, Qwen-2.5-1.5B, and its distilled variant Qwen-2.5-1.5B-SFT. All other components are kept fixed to isolate the effect of directional control quality. GPT-4o-mini achieves the strongest overall performance, whereas the vanilla Qwen-2.5-1.5B shows clear

degradation, especially in distributional metrics. This suggests that lightweight models struggle to reliably infer semantic directional signals under unseen spatial-temporal settings. After supervised fine-tuning, Qwen-2.5-1.5B-SFT substantially narrows the gap and achieves performance close to the teacher model.

Figure 5 further shows the performance-cost trade-off of different controllers. Distillation significantly improves the efficiency-quality trade-off of the lightweight model, enabling efficient local inference while preserving strong generation quality. Overall, the results demonstrate that the proposed distillation strategy effectively transfers semantic control capability from a strong LLM to a student model.

Generality Across OD Generators

Model	CPC $\uparrow$	RMSE $\downarrow$	NRMSE $\downarrow$	JSD-In $\downarrow$	JSD-Out $\downarrow$	JSD-OD $\downarrow$
NetGAN _{pt}	0.365	293.42	1.129	0.657	0.736	0.574
DynaOD _{NetGAN}	0.461	281.01	1.096	0.611	0.519	0.522
WeDAN _{pt}	0.331	286.51	1.056	0.502	0.741	0.333
DynaOD _{WeDAN}	0.492	253.09	0.940	0.322	0.421	0.220

Table 3: Generality of DynaOD across different pretrained OD generators. $\uparrow$ ($\downarrow$) indicates higher (lower) values are better.

Table 3 evaluates DynaOD across different pretrained OD generators, including NetGAN and WeDAN. In both settings, the generators remain frozen and DynaOD is applied without modifying the original architecture or training procedure. Across both generators, integrating DynaOD consistently improves performance in terms of CPC, error metrics, and distributional similarity. For the relatively weaker NetGAN, DynaOD leads to substantial gains, particularly in CPC and JSD metrics, indicating that directional control and shape modulation effectively adapt static graph generators to dynamic mobility scenarios. For the stronger WeDAN generator, DynaOD further improves performance across all metrics, achieving the best overall results.

Notably, while DynaOD consistently improves different OD generators, the absolute performance remains dependent on the capacity of the underlying generator. This indicates that DynaOD enhances temporal adaptability while preserving the core modeling capability of the base generator. As a result, DynaOD can naturally benefit from future advances in pretrained OD generators.

4.5 Case Study

To qualitatively illustrate how DynaOD translates temporal context into dynamic feature dynamics, we present a case study on a single region over a 7-day generation window. Figure 6 visualizes the discrete directional controls generated by the LLM and the corresponding continuous feature shapes produced by ShapeNet.

The left column shows the discrete POI and demographic control vectors. Although the controls are coarse-grained and temporally piecewise constant, they exhibit stable directional patterns across days. The right column shows the resulting continuous feature trajectories, where smooth temporal evolution naturally emerges from the discrete semantic controls.

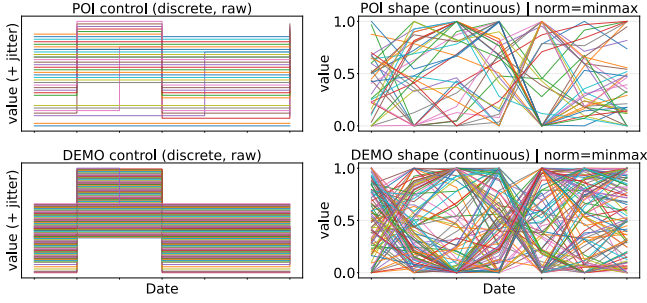


Figure 6: Case study of discrete directional control and continuous shape generation for a randomly selected tract over a 7-day window (2019-01-25 to 2019-01-31). Left: raw discrete POI and demographic control vectors generated by GPT-4o-mini (with jitter for visualization). Right: corresponding continuous feature shapes produced by ShapeNet (min-max normalized per dimension).

This example demonstrates how DynaOD decouples semantic temporal control from continuous numerical modeling: high-level directional controls determine *what* should change, while ShapeNet models *how* these changes evolve smoothly over time. Consequently, DynaOD can generate temporally coherent mobility dynamics without relying on historical OD observations.

5 Related Work

Urban spatio-temporal modeling has been widely studied [Li *et al.*, 2023; Xu *et al.*, 2023; Yuan *et al.*, 2024], among which OD flow modeling represents an important research direction. Existing studies can be broadly divided into static OD flow generation and dynamic OD flow modeling.

5.1 Static OD Flow Generation

Early work on static OD flow generation mainly relied on physical models, such as the gravity model [Zipf, 1946] and the radiation model [Simini *et al.*, 2012; Kang *et al.*, 2015], which model population movement through simplified physical assumptions. While computationally efficient, these approaches are constrained by strong inductive assumptions and often fail to capture the complex and heterogeneous nature of human mobility. Recent studies have shifted toward data-driven OD generation, leveraging machine learning and deep learning models to synthesize OD flows directly from urban attributes such as demographic, socioeconomic, and POI distributions [Liu *et al.*, 2020; Pourebrahim *et al.*, 2019; Robinson and Dilkina, 2018; Simini *et al.*, 2021; Rong *et al.*, 2023]. These approaches demonstrate that OD flows contain transferable spatial regularities and can be generated without direct trajectory observations. Nevertheless, most existing methods focus on static or time-aggregated settings and are evaluated within individual cities, limiting their ability to generalize across cities and evolving temporal contexts.

5.2 Dynamic OD Flow Modeling

Research on dynamic OD flows has primarily focused on forecasting, estimation, and scenario-driven simulation, rather than context-conditioned generation. Limited studies

consider scenario-based OD generation. For instance, [Yean *et al.*, 2025] uses conditional GANs to generate mobility flows under simulated urban scenarios with adaptive region definitions. However, such approaches still rely on historical mobility data for calibration and cannot generalize to unseen temporal contexts.

Most existing methods follow the forecasting paradigm, inferring future OD matrices from historical OD sequences using spatio-temporal learning models, including dynamic graph learning [Zhang and Xiao, 2021; Zhang *et al.*, 2022], continuous- and discrete-time representation learning [Xu *et al.*, 2023], hypergraph-based modeling [Shen *et al.*, 2024], and prototype-based hierarchical learning [Yuan *et al.*, 2025]. While effective for short-term forecasting, these approaches fundamentally depend on sufficient historical OD observations. Another line of research estimates or imputes dynamic OD flows from indirect traffic observations. Representative approaches integrate traffic states into OD modeling [Rong *et al.*, 2024b], transfer knowledge across heterogeneous data sources [Chen *et al.*, 2024], or formulate OD estimation as traffic-constrained optimization problems [Englezou *et al.*, 2024]. Despite reducing reliance on direct OD observations, these methods still require partial traffic measurements or calibrated physical assumptions.

Overall, existing dynamic OD modeling methods remain heavily dependent on historical OD data, traffic observations, or scenario-specific calibration. In contrast, dynamic OD generation under minimal supervision, where temporal context alone drives OD generation, remains largely underexplored.

6 Limitations

Despite its effectiveness, DynaOD has several limitations. First, the quality of discrete control depends on the reasoning capability of the language model, which may vary across regions with sparse or ambiguous contextual information. Second, ShapeNet currently models temporal evolution at a fixed daily granularity, and extending it to finer or irregular time scales remains an open problem. Finally, while DynaOD is generator-agnostic, its absolute performance is still bounded by the expressiveness of the underlying OD generator.

7 Conclusion

This paper introduces DynaOD, a framework for dynamic origin-destination flow generation that unifies semantic temporal control with generative mobility modeling. DynaOD uses LLMs to infer interpretable directional controls from temporal context and converts them into continuous feature dynamics through ShapeNet. These time-varying features are then used to adapt pretrained static OD generators for dynamic mobility synthesis without retraining. To improve scalability and cross-city generalization, DynaOD further integrates retrieval-based shape priors and a lightweight LLM distillation strategy. Extensive experiments on large-scale real-world datasets demonstrate that DynaOD consistently outperforms representative baselines under unseen spatial-temporal settings and remains compatible with different pretrained OD generators.

Acknowledgements

This work was supported by the the Science and Technology Innovation Program of Xiongan New Area (Grant No. 2025XAGG0041)

References

- [Andrienko *et al.*, 2016] Gennady Andrienko, Natalia Andrienko, Georg Fuchs, and Jo Wood. Revealing patterns and trends of mass mobility through spatial and temporal abstraction of origin-destination movement data. *IEEE transactions on visualization and computer graphics*, 23(9):2120–2136, 2016.
- [Barroso *et al.*, 2020] Joana Maia Fernandes Barroso, João Lucas Albuquerque-Oliveira, and Francisco Moraes Oliveira-Neto. Correlation analysis of day-to-day origin-destination flows and traffic volumes in urban networks. *Journal of Transport Geography*, 89:102899, 2020.
- [Bauer *et al.*, 2017] Dietmar Bauer, Gerald Richter, Johannes Asamer, Bernhard Heilmann, Gernot Lenz, and Robert Kölbl. Quasi-dynamic estimation of od flows from traffic counts without prior od matrix. *IEEE Transactions on Intelligent Transportation Systems*, 19(6):2025–2034, 2017.
- [Bojchevski *et al.*, 2018] Aleksandar Bojchevski, Oleksandr Shchur, Daniel Zügner, and Stephan Günnemann. NetGAN: Generating graphs via random walks. In *International conference on machine learning*, pages 610–619. PMLR, 2018.
- [Chen *et al.*, 2024] Peng Chen, Ziyan Wang, Bin Zhou, and Guizhen Yu. Dynamic origin-destination flow imputation using feature-based transfer learning. *IEEE Transactions on Intelligent Transportation Systems*, 25(11):17147–17159, 2024.
- [Cheng *et al.*, 2025] Jian Cheng, Haidong Kang, Yuxin Shao, Nan Li, Pengjun Chen, Rui Wang, Saiqin Long, Xiaochun Yang, and Lianbo Ma. Survey on efficient large language models: Principles, algorithms, applications, and open issues. *IEEE Transactions on Neural Networks and Learning Systems*, 2025.
- [Englezou *et al.*, 2024] Yirolanda Englezou, Stelios Timotheou, and Christos G Panayiotou. Path-based origin-destination matrix estimation utilizing macroscopic traffic dynamics. *IEEE Transactions on Intelligent Transportation Systems*, 25(8):8819–8836, 2024.
- [Fang *et al.*, 2022] Ziquan Fang, Dongen Wu, Lu Pan, Lu Chen, and Yunjun Gao. When transfer learning meets cross-city urban flow prediction: Spatio-temporal adaptation matters. In *IJCAI*, volume 22, pages 2030–2036, 2022.
- [Fang *et al.*, 2026] Luyang Fang, Xiaowei Yu, Jiazhang Cai, Yongkai Chen, Shushan Wu, Zhengliang Liu, Zhenyuan Yang, Haoran Lu, Xilin Gong, Yufang Liu, et al. Knowledge distillation and dataset distillation of large language models: Emerging trends, challenges, and future directions. *Artificial Intelligence Review*, 59(1):17, 2026.
- [Grinberger and Shoval, 2019] A Yair Grinberger and Noam Shoval. Spatiotemporal contingencies in tourists’ intra-diurnal mobility patterns. *Journal of Travel Research*, 58(3):512–530, 2019.
- [Kang *et al.*, 2015] Chaogui Kang, Yu Liu, Diansheng Guo, and Kun Qin. A generalized radiation model for human mobility: spatial scale, searching direction and trip constraint. *PloS one*, 10(11):e0143500, 2015.
- [Kang *et al.*, 2020] Yuhao Kang, Song Gao, Yunlei Liang, Mingxiao Li, and Jake Kruse. Multiscale dynamic human mobility flow dataset in the u.s. during the covid-19 epidemic. *Scientific Data*, pages 1–13, 2020.
- [Li *et al.*, 2021] Zhenlong Li, Xiao Huang, Tao Hu, Huan Ning, Xinyue Ye, Binghu Huang, and Xiaoming Li. ODT FLOW: Extracting, analyzing, and sharing multi-source multi-scale human mobility. *Plos one*, 16(8):e0255259, 2021.
- [Li *et al.*, 2023] Fuxian Li, Jie Feng, Huan Yan, Guangyin Jin, Fan Yang, Funing Sun, Depeng Jin, and Yong Li. Dynamic graph convolutional recurrent network for traffic prediction: Benchmark and solution. *ACM Transactions on Knowledge Discovery from Data*, 17(1):1–21, 2023.
- [Li *et al.*, 2025] Yudong Li, Jingyuan Wang, Xie Yu, Peiyu Wang, and Qian Huang. Cross city traffic flow generation via retrieval augmented diffusion model. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025.
- [Liu *et al.*, 2020] Zhicheng Liu, Fabio Miranda, Weiting Xiong, Junyan Yang, Qiao Wang, and Claudio Silva. Learning geo-contextual embeddings for commuting flow prediction. In *Proceedings of the AAAI conference on artificial intelligence*, volume 34, pages 808–816, 2020.
- [Miller, 2018] Eric J Miller. Accessibility: measurement and application in transportation planning. *Transport Reviews*, 38(5):551–555, 2018.
- [Perez *et al.*, 2018] Ethan Perez, Florian Strub, Harm De Vries, Vincent Dumoulin, and Aaron Courville. Film: Visual reasoning with a general conditioning layer. In *Proceedings of the AAAI conference on artificial intelligence*, volume 32, 2018.
- [Pourebbrahim *et al.*, 2019] Nastaran Pourebbrahim, Selima Sultana, Amirreza Niakanlahiji, and Jean-Claude Thill. Trip distribution modeling with twitter data. *Computers, Environment and Urban Systems*, 77:101354, 2019.
- [Robinson and Dilkina, 2018] Caleb Robinson and Bistra Dilkina. A machine learning approach to modeling human migration. In *Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies*, pages 1–8, 2018.
- [Rodríguez-Rueda *et al.*, 2021] PJ Rodríguez-Rueda, JJ Ruiz-Aguilar, J González-Enrique, and I Turias. Origin–destination matrix estimation and prediction from socioeconomic variables using automatic feature selection procedure-based machine learning model. *Journal of urban planning and development*, 147(4):04021056, 2021.

- [Rong et al., 2023] Can Rong, Jie Feng, and Jingtao Ding. GODDAG: Generating origin-destination flow for new cities via domain adversarial training. *IEEE Transactions on Knowledge and Data Engineering*, 35(10):10048–10057, 2023.
- [Rong et al., 2024a] Can Rong, Jingtao Ding, and Yong Li. An interdisciplinary survey on origin-destination flows modeling: Theory and techniques. *ACM Computing Surveys*, 57(1):1–49, 2024.
- [Rong et al., 2024b] Can Rong, Zhicheng Liu, Jingtao Ding, and Yong Li. Learning to generate temporal origin-destination flow based-on urban regional features and traffic information. *ACM Transactions on Knowledge Discovery from Data*, 18(6):1–17, 2024.
- [Rong et al., 2025] Can Rong, Jingtao Ding, Yan Liu, and Yong Li. A large-scale dataset and benchmark for commuting origin-destination flow generation. In *The Thirteenth International Conference on Learning Representations*, volume 2025, pages 96180–96205, 2025.
- [Shen et al., 2024] Loutao Shen, Junyi Li, Yong Chen, Chuanjia Li, Xiqun Chen, and Der-Hong Lee. Short-term metro origin-destination passenger flow prediction via spatio-temporal dynamic attentive multi-hypergraph network. *IEEE Transactions on Intelligent Transportation Systems*, 25(8):9945–9957, 2024.
- [Simini et al., 2012] Filippo Simini, Marta C González, Amos Maritan, and Albert-László Barabási. A universal model for mobility and migration patterns. *Nature*, 484(7392):96–100, 2012.
- [Simini et al., 2021] Filippo Simini, Gianni Barlacchi, Massimiliano Luca, and Luca Pappalardo. A deep gravity model for mobility flows generation. *Nature communications*, 12(1):6576, 2021.
- [Sparks et al., 2022] Kevin Sparks, Jessica Moehl, Eric Weber, Christa Brelsford, and Amy Rose. Shifting temporal dynamics of human mobility in the united states. *Journal of transport geography*, 99:103295, 2022.
- [Thakoor et al., 2022] Shantanu Thakoor, Corentin Tallec, Mohammad Gheshlaghi Azar, Mehdi Azabou, Eva L Dyer, Remi Munos, Petar Veličković, and Michal Valko. Large-scale representation learning on graphs via bootstrapping. *International Conference on Learning Representations (ICLR)*, 2022.
- [van Wee et al., 2023] Bert van Wee, Jan A Annema, David Banister, and Baiba Pudāne. *The transport system and transport policy: an introduction*. Edward Elgar Publishing, 2023.
- [Xu et al., 2023] Yi Xu, Liangzhe Han, Tongyu Zhu, Leilei Sun, Bowen Du, and Weifeng Lv. Continuous-time and discrete-time representation learning for origin-destination demand prediction. *IEEE Transactions on Intelligent Transportation Systems*, 25(3):2382–2393, 2023.
- [Yean et al., 2025] Seanglidet Yean, Jiazu Zhou, Bu-Sung Lee, and Markus Schläpfer. FloGAN: Scenario-based urban mobility flow generation via conditional gans and dynamic region decoupling. *arXiv preprint arXiv:2507.12053*, 2025.
- [Yuan et al., 2024] Yuan Yuan, Jingtao Ding, Jie Feng, Depeng Jin, and Yong Li. Unist: A prompt-empowered universal model for urban spatio-temporal prediction. In *Proceedings of the 30th ACM SIGKDD conference on knowledge discovery and data mining*, pages 4095–4106, 2024.
- [Yuan et al., 2025] Shilu Yuan, Xiaoyu Li, Wenqian Mu, Ji Zhong, Meng Chen, Haoliang Sun, and Yongshun Gong. Spatio-temporal prototype-based hierarchical learning for od demand prediction. In *Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence*, pages 3597–3605, 2025.
- [Zhang and Xiao, 2021] Dapeng Zhang and Feng Xiao. Dynamic auto-structuring graph neural network: A joint learning framework for origin-destination demand prediction. *IEEE Transactions on Knowledge and Data Engineering*, 35(4):3699–3711, 2021.
- [Zhang et al., 2020] Qi Zhang, Hong Chen, Hongchao Liu, Wei Li, and Yibin Zhang. Origin-destination-based travel time reliability under different rainfall intensities: An investigation using open-source data. *Journal of Advanced Transportation*, 2020(1):8816020, 2020.
- [Zhang et al., 2022] Ruixing Zhang, Liangzhe Han, Boyi Liu, Jiayuan Zeng, and Leilei Sun. Dynamic graph learning based on hierarchical memory for origin-destination demand prediction. In *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI-22*, pages 2383–2389, 2022.
- [Zhang et al., 2025] Yuheng Zhang, Yuan Yuan, Jingtao Ding, Jian Yuan, and Yong Li. Noise Matters: Diffusion model-based urban mobility generation with collaborative noise priors. In *Proceedings of the ACM on Web Conference 2025*, pages 5352–5363, 2025.
- [Zipf, 1946] George Kingsley Zipf. The $p \propto 1/p^d$ hypothesis: on the intercity movement of persons. *American sociological review*, 11(6):677–686, 1946.