

Towards Routing Foundation Models for Commercial Vehicles

Bo Tang*

Massachusetts Institute of Technology
Cambridge, Massachusetts, USA
botang@mit.edu

Yunfei Ma✉*

Geotab
Oakville, Ontario, Canada
yunfeimai@geotab.com

Zhengzhong Ricky You

Tsinghua University
Beijing, China
ricky.you.or@gmail.com

Abstract

Commercial fleet trajectories record recurring operational decisions across vehicles, regions, and time, offering a resource for pretraining routing foundation models. Classical optimization methods solve specified routing instances, while learning-based routing methods train largely on synthetic instances. Longitudinal commercial trajectories offer a complementary source of operational experience for learning local movement patterns and recurring service sequences. We propose a research roadmap connecting this resource to routing through controlled studies of scaling, transfer, and privacy. We will compare spatiotemporal representations and separate the effects of data volume and coverage from model capacity and compute. Downstream studies will test whether pretraining improves routing quality, feasibility, and adaptation efficiency when task-specific stops and constraints are supplied. Finally, we will evaluate whether synthetic trajectories transfer these benefits to independent models while assessing disclosure risks. The proposed studies ask whether additional fleet experience yields predictable gains that persist across regions and routing tasks and can be shared beyond a single model architecture.

CCS Concepts

• Information systems → Spatial-temporal systems; • Applied computing → Transportation.

Keywords

commercial fleet trajectories, foundation models, route planning, scaling laws, transfer learning

ACM Reference Format:

Bo Tang, Yunfei Ma, and Zhengzhong Ricky You. 2026. Towards Routing Foundation Models for Commercial Vehicles. In *The 2nd ACM SIGSPATIAL International Workshop on Urban Mobility Foundation Models (UMFM '26)*, November 03–06, 2026, Riverside, CA, USA. ACM, New York, NY, USA, 2 pages. <https://doi.org/10.1145/3849729.3856134>

1 Introduction

Vehicle routing is a core operational problem in parcel delivery, field service, freight transportation, waste collection, and many other fleet systems. Classical routing methods provide mature tools for well-specified routing instances [8]. Real-world deployment, however, presents challenges at two levels: computational, because

large or frequently updated instances may exceed the available solution time, and modeling, because the mathematical model may omit relevant operating conditions. Travel times change, access and parking conditions are only partly observed, and drivers often rely on local knowledge that is difficult to encode. Dynamic and stochastic formulations capture some of this variability only when the relevant states or distributions are specified. The Amazon Last Mile Routing Challenge, which scored routes by their similarity to driver-executed sequences, illustrates the value of operational history: its winning method extracted zone-order constraints from historical executed routes and enforced them within a local-search solver [3]. Historical operations can therefore provide useful routing structure that is not fully expressed by a standard routing objective.

These limitations motivate learning-based routing. Neural combinatorial optimization learns construction or improvement heuristics for mathematically specified routing instances, often on synthetic Euclidean instances, but generalization across sizes and distributions remains difficult [10]. RouteFinder transfers across combinations of vehicle routing problem (VRP) attributes, while RRNCO introduces road-network-derived asymmetric distance and duration matrices from 100 cities [1, 7]. These methods broaden instance coverage and realism but do not study pretraining on longitudinal commercial fleet records or scaling with observed fleet experience.

Trajectory foundation models provide a complementary path based on observed mobility. UniTraj pretrains on worldwide traces and studies transfer across trajectory tasks, while MoveGPT explores scaling across data, model capacity, and geographic coverage [5, 11]. These results suggest that large trajectory corpora can support transferable spatiotemporal representations, although their value for downstream routing remains largely untested. Record count alone does not guarantee effective scale; results at up to 10^{11} samples show that representation and sample quality affect scaling [6].

Geotab offers a distinctive setting in which to test this premise. The company reports approximately six million connected vehicles and assets and 100 billion data points processed daily [4]. Commercial fleet trajectories are collected continuously, linked to vehicles over time, span multiple regions, and reflect recurring and diverse commercial routing behavior. This structure makes it possible to measure data growth across unique vehicles, vehicle-days, regions, and temporal coverage rather than raw GPS points alone. We propose to compare multi-resolution spatiotemporal tokenizations and estimate scaling relationships while varying data quantity, quality, diversity, model capacity, and compute. We will test whether additional operational experience yields predictable gains with scale that persist across regions and routing tasks and transfer through synthetic data to independent architectures.

*These authors contributed equally to this research.



This work is licensed under a Creative Commons Attribution 4.0 International License. *UMFM '26, Riverside, CA, USA*

© 2026 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-3055-9/2026/11

<https://doi.org/10.1145/3849729.3856134>

2 Scaling and Transfer

A large training corpus alone does not establish a routing foundation model. Additional fleet data must yield measurable gains that improve routing when task-specific stops, constraints, and objectives are supplied. We will first characterize coverage and sampling continuity, then compare multi-resolution spatiotemporal tokenizations using next-location, leg travel-time, and idling-duration prediction signals. Full trajectories retain movement between stops but require longer sequences; stop-based representations shorten the context while discarding some movement detail. We will compare the routing information retained at each resolution against vocabulary size and sequence length. Controlled scaling studies will guide routing adaptation, followed by synthetic transfer and privacy evaluation.

Scaling laws. Lin et al. [6] attributed the lack of expected scaling in conventional place-visitation modeling at up to 10^{11} samples to sample-quality problems; their spatial-interaction behavior formulation recovered power-law scaling. We will measure data growth across unique vehicles, vehicle-days, regions, and temporal coverage. For each tokenization, data comparisons will hold model architecture, capacity, and training compute fixed while varying data quantity, quality, and diversity. Separate studies will vary model capacity and compute independently. We will fit scaling relationships on smaller runs and test extrapolations on larger runs, using disjoint vehicle, region, and time splits to evaluate generalization.

Transfer learning. Better trajectory prediction does not necessarily lead to better routing. UniTraj transfers across trajectory tasks and RouteFinder across vehicle routing problem (VRP) variants [1, 11]. Historical zone order and precedence information also guided the winning Amazon solver [3]. Fleet trajectories do not by themselves specify the required stops, objective, or constraints of a routing instance. We will adapt pretrained models using task-specific datasets that supply this information. The model will rank candidate orders or propose initial routes using only the required stops. A conventional solver will check feasibility and refine these solutions. Learned route preferences will guide the search, while the routing instance determines customer coverage, capacity limits, and time windows. Under a fixed solve budget, we will compare against the same architecture trained from scratch and strong routing baselines. We will report route cost, feasibility, and runtime, and measure adaptation efficiency by the routing data and compute needed to reach matched performance. Alignment with observed operational patterns will be assessed separately from route quality. We will test whether routing gains increase with pretraining scale and persist across tasks and regions.

Synthetic data and privacy. Synthetic trajectories could complement model weights by transferring learned fleet structure across architectures without releasing raw records. PrivTrace shows that differentially private trajectory synthesis can retain selected aggregate mobility statistics [9]. It remains unknown whether synthetic data from larger pretrained models provide greater benefits to independent models on real routing tasks. Establishing this relationship would allow researchers to study the value of commercial fleet experience using architectures suited to their own tasks. We will

generate equal data volumes at each pretraining scale to train independent receiver models with architectures distinct from the generator architecture. Each receiver architecture, its training compute, and the adaptation protocol will remain fixed across generator scales, with transfer evaluated on held-out real routing data. Privacy evaluation will include membership inference, trajectory-user linkage, reconstruction attacks, and sensitive-location disclosure [2]. These empirical tests will characterize privacy risks without establishing a privacy guarantee.

3 Outlook

We aim to contribute the pretrained model and privacy-evaluated synthetic trajectories as common starting points for routing research. Public trajectory datasets have supported progress in mobility learning, but their scale, continuity, and geographic coverage can constrain foundation-model research. The model could support adaptation to new regions and tasks, comparisons across architectures, and continued research on scaling and transfer. Synthetic trajectories could make learned structure available without releasing raw fleet records. The broader goal is shared scientific infrastructure for developing more adaptive, efficient, and reliable routing systems.

References

- [1] Federico Berto, Chuanbo Hua, Nayeli Gast Zepeda, André Hottung, Niels Wouda, Leon Lan, Junyoung Park, Kevin Tierney, and Jinkyoo Park. 2025. RouteFinder: Towards Foundation Models for Vehicle Routing Problems. *Transactions on Machine Learning Research* (2025). <https://openreview.net/forum?id=QzGLOaOPiY>
- [2] Erik Buchholz, Alsharif Abuadbba, Shuo Wang, Surya Nepal, and Salil S. Kanhere. 2024. SoK: Can Trajectory Generation Combine Privacy and Utility? *Proceedings on Privacy Enhancing Technologies* 2024, 3 (2024), 75–93. doi:10.56553/popets-2024-0068
- [3] William Cook, Stephan Held, and Keld Helsgaun. 2024. Constrained Local Search for Last-Mile Routing. *Transportation Science* 58, 1 (2024), 12–26. doi:10.1287/trsc.2022.1185
- [4] Geotab Inc. 2026. About Geotab. <https://www.geotab.com/about/>. Accessed August 28, 2026.
- [5] Chonghua Han, Yuan Yuan, Jintao Ding, Jie Feng, Fanjin Meng, and Yong Li. 2025. MoveGPT: Scaling Mobility Foundation Models with Spatially-Aware Mixture of Experts. arXiv:2505.18670 [cs.AI] <https://arxiv.org/abs/2505.18670>
- [6] Lifeng Lin, Hangli Ge, Xiang Zhang, Yao Yao, Kazuma Hatano, Ryosuke Shibasaki, and Noboru Koshizuka. 2026. The Scaling Bottleneck of Human Mobility Modeling. In *Proceedings of the 32nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining V.2*. ACM, 11411–11421. doi:10.1145/3770855.3818976
- [7] Jiwoo Son, Zhikai Zhao, Federico Berto, Chuanbo Hua, Zhiguang Cao, Changhyun Kwon, and Jinkyoo Park. 2026. RRNCO: Towards Real-World Routing with Neural Combinatorial Optimization. In *International Conference on Learning Representations*, C. Vondrick, B. Hariharan, C. Raffel, L. Pinto, D. Yang, and A. Faust (Eds.), Vol. 2026. 91794–91818. https://proceedings.iclr.cc/paper_files/paper/2026/file/9446c291a8744a125a0bda5b18f4d5a1-Paper-Conference.pdf
- [8] Thibaut Vidal, Gilbert Laporte, and Piotr Matl. 2020. A Concise Guide to Existing and Emerging Vehicle Routing Problem Variants. *European Journal of Operational Research* 286, 2 (2020), 401–416. doi:10.1016/j.ejor.2019.10.010
- [9] Haiming Wang, Zhikun Zhang, Tianhao Wang, Shibo He, Michael Backes, Jiming Chen, and Yang Zhang. 2023. PrivTrace: Differentially Private Trajectory Synthesis by Adaptive Markov Models. In *32nd USENIX Security Symposium (USENIX Security 23)*. USENIX Association, 1649–1666. <https://www.usenix.org/conference/usenixsecurity23/presentation/wang-haiming>
- [10] Jianan Zhou, Yaoxin Wu, Wen Song, Zhiguang Cao, and Jie Zhang. 2023. Towards Omni-generalizable Neural Methods for Vehicle Routing Problems. In *Proceedings of the 40th International Conference on Machine Learning (Proceedings of Machine Learning Research, Vol. 202)*. PMLR, 42769–42789. <https://proceedings.mlr.press/v202/zhoul23o.html>
- [11] Yuanshao Zhu, James Yu, Xiangyu Zhao, Xun Zhou, Liang Han, Xuetao Wei, and Yuxuan Liang. 2025. UniTraj: Learning a Universal Trajectory Foundation Model from Billion-Scale Worldwide Traces. In *Advances in Neural Information Processing Systems*, Vol. 38. Curran Associates, Inc., 47228–47265. doi:10.52202/085713-1407