Modeling and Predicting Agent Trajectory in Urban Road Networks

Julien Baudru[1 [0000-0002-8771-2494]] and Hugues Bersini[1 [0000-0001-5820-7360]]

Université Libre de Bruxelles (ULB), Brussels, Belgium
julien.baudru@ulb.be

Abstract. Accurate next-node prediction in road networks supports efficient carpooling, routing, and traffic management in urban systems. We developed an agent-based simulator of the Brussels road network to generate synthetic mobility data using five source-target selection patterns. Using this data, we evaluate LSTM, Transformer, and GNN models for next-node prediction. Results show that models perform best on Activity and Hub-and-Spoke datasets (around 94% accuracy), while the Zone dataset poses greater challenges, especially for GNNs, highlighting the importance of dataset complexity over size in prediction accuracy.

Keywords: Urban Mobility · Next Location Prediction · Road Network Simulation · Deep Learning · Graph Neural Networks · Transformer Models · LSTM · Agent-Based Simulation.

1 Introduction

Understanding and predicting urban mobility patterns is critical for designing efficient, low-emission transportation systems in dense cities. Accurate travel demand prediction enables better planning, reduces congestion, and can significantly lower computational costs for tasks such as car-pooling transfer point computation and vehicle routing in large-scale road networks [1]. Deep learning approaches, particularly those leveraging spatio-temporal data, have demonstrated promise in capturing the complex dependencies inherent in transportation systems [8]. In Brussels, multimodal travel is essential, with 36% of trips made on foot, 9% by bicycle, 22% by metro/tram/bus, and 27% by car, highlighting the need for integrated, predictive transport systems to support the shift toward sustainable mobility [6]. To address these challenges, we develop a simulator based on the Brussels road network, where agents simulate daily movements using among five source-target ($s-t$) selection models. Using this synthetic, we train and evaluate three predictive models, a LSTM, a Transformer, and a Graph Neural Networks (GNN), to assess their effectiveness in learning urban mobility patterns for practical downstream applications such as transfer point estimation and dynamic car-pooling optimization, as well as classical traffic forecasting, demand-responsive transit scheduling, congestion management, or real-time route recommendation.

2 Related work

Traditional mobility models focus on destination choice given origins. Recent work integrates spatial interaction with data-driven methods. Yan et al. proposed a destination choice game incorporating congestion and crowding, improving flow predictions over gravity and radiation models [9]. Deep learning approaches, such as Hu and Zhang's GraphResLSTM, combine graph convolution, residual networks, and LSTM to predict origin-destination demand on SUMO-simulated road networks, outperforming alternatives [3]. Traffic simulators like SUMO and MATSim enable scalable, reproducible evaluations with realistic synthetic trajectories [3]. Next location prediction uses recent trajectories to forecast immediate destinations, classical Markov models struggle with long-range dependencies, whereas LSTMs yield better sequential modeling [7]. Transformer-based models further are supposed to improve accuracy by leveraging travel modes and spatiotemporal context, as shown by Hong et al. with 5–8% F1 score gains [2]. Graph neural networks capture complex spatial relations in sparse urban settings, enhancing performance, as surveyed by Yu et al. [10].

3 Simulation Framework

To generate realistic mobility data for prediction tasks, we developed a simulation framework using the Brussels road network $G = (V, E)$ with $|V| = 18547$ road intersection and $|E| = 40890$ road segments, modeling agents navigating the network with dynamic, behavior-driven trip generation. Each agent in the simulation is assigned a travel speed for each edge and experiences congestion based on the presence of other moving agents, modeled using the Bureau of Public Roads (BPR) congestion function [4]. Specifically, the speed on an edge decreases as the volume-to-capacity ratio increases, ensuring that agents experience slower travel during peak congestion while maintaining a minimum operational speed of 10% of the normal speed. For all $s-t$ selection models, the probability of an agent starting a new trip depends on the 24-hour simulation time (6 AM to 6 AM next day). This temporal pattern reflects real-world mobility, with peaks during morning and evening rush hours and moderate activity at midday. We implemented five source-target $(s-t)$ selection models to reflect diverse travel behaviors.

Random Model: Uniformly random source and target selection for baseline comparison.

Activity-Based Model: Agents have predefined home and work locations and follow behavioral patterns based on type (commuters, delivery, leisure, business), ensuring realistic alternation between common activity nodes.

Zone-Based Model: The network is divided into a 3x3 grid, with 60% of trips occurring within zones and 40% between zones, following Transportation Analysis Zones (TAZ) [5] methodology.

Gravity Model: Implements a classical gravity model where trip probabilities are proportional to node importance (centrality) and inversely proportional to distance, capturing the tendency to travel to significant locations while avoiding excessively long trips.

Hub-and-Spoke Model: Reflects real-world concentration of traffic by ensuring that 70% of trips involve major hub nodes with high centrality, modeling flows toward and from central urban points.

Building our own simulator ensures synthetic data closely aligns with the targeted urban context. This improves the relevance and robustness of downstream predictive modeling.

4 Prediction models

We evaluate three neural architectures for next-node prediction in road network trajectories, where the task is to predict the next visited node given a fixed-length sequence of previously visited nodes.

LSTM Model: Uses a 3-layer bidirectional LSTM with 512 hidden units per direction, preceded by a 256-dimensional node embedding layer. Dropout (0.3) is applied between layers for regularization. The final hidden states are concatenated (1024-dim) and passed to a fully connected classification layer producing a softmax distribution over all road network nodes.

Transformer Model: Implements a 6-layer encoder with 8 attention heads per layer and 512-dimensional node embeddings with sinusoidal positional encodings. Causal masking enforces autoregressive prediction. Layer normalization is applied pre-activation for stability, and multi-head self-attention captures long-range dependencies across node sequences.

GNN Model: Utilizes a 4-layer Graph Attention Network (GAT) with 4 heads per layer to process the full road network graph, learning topology-aware node embeddings (128-dim). These enriched embeddings feed into a 2-layer bidirectional LSTM (256 hidden units per direction) to encode temporal patterns before classification.

All models are trained as multi-class classifiers using cross-entropy loss with 0.1 label smoothing and optimized with AdamW using a cosine annealing scheduler. Batch sizes and learning rates are tuned per architecture to account for computational differences: LSTM ($batch = 64$, $lr = 0.001$), Transformer ($batch = 32$, $lr = 0.0001$), and GNN ($batch = 48$, $lr = 0.0005$). Early stopping ($patience = 10$) and gradient clipping ($maxnorm = 1.0$) are employed for stability.

5 Results

The three prediction models described above were trained and evaluated independently using trip data generated from the different s-t selection models with 1000 agents simulated over a 24-hour virtual period. Table 1 summarizes the size of each dataset, each training instance consists of an input sequence of 10 consecutive road network nodes and the corresponding next node as the output. Table 2 reports the validation loss, as well as the final the validation and test accuracies achieved by each model on each dataset.

Table 1. Dataset sizes used for training, evaluation and testing

Dataset	Train. Size (60%)	Val. Size (20%)	Test Size (20%)
Random	348885	116295	116294
Activity	330516	110171	110171
Zone	252849	84282	84282
Gravity	259719	86573	86573
Hub-and-Spoke	342494	114164	114164

Table 2. Validation accuracy, train loss, and final validation loss for each model on five s-t selection datasets with 1000 agents over 24-hour simulation.

Dataset	Model	Train Loss	Val. Loss	Val. Accuracy (%)
Random	LSTM	1.4619	1.626	92.77
Random	Transformer	1.5143	1.6871	92.23
Random	GNN	1.5247	1.6635	92.36
Activity-Based	LSTM	1.4187	1.5357	93.97
Activity-Based	Transformer	1.4549	1.5745	93.59
Activity-Based	GNN	1.4673	1.562	93.71
Zone-Based	LSTM	1.4746	1.7235	90.88
Zone-Based	Transformer	0.1626	0.613	90.52
Zone-Based	GNN	0.8477	0.9382	85.94
Gravity	LSTM	1.4806	1.7403	90.74
Gravity	Transformer	1.512	1.7933	90.35
Gravity	GNN	1.5209	1.749	90.66
Hub-and-Spoke	LSTM	1.4559	1.6092	93.04
Hub-and-Spoke	Transformer	1.5023	1.6666	92.55
Hub-and-Spoke	GNN	1.513	1.6439	92.67

We observe that the Activity and Hub-and-Spoke datasets yield the highest validation accuracies and lower validation losses across all models, indicating that these datasets are the least challenging for the learning task. Specifically, validation accuracies for these datasets consistently reach around 94% across LSTM, Transformer, and GNN architectures, demonstrating that the models effectively learn the $s-t$ selection patterns under these conditions. In contrast, the Zone dataset presents the greatest challenge for all models, particularly for the GNN, which achieves a notably lower validation accuracy of 85.94%. This performance degradation suggests that the spatial aggregation mechanisms employed by GNNs may not fully capture the relevant patterns required for next node prediction in the Zone dataset structure. Additionally, it is notable that the Zone dataset is the smallest among the evaluated datasets, and the lower accuracy observed, especially for GNNs. Interestingly, larger dataset sizes do not guarantee higher model performance in our experiments. For example, the Random dataset, which has the largest sample size, yields an average validation accuracy of approximately 92%, whereas the smaller Activity dataset consistently achieves higher accuracies around 94%. This finding suggests that dataset complexity plays a more significant role than dataset size, since there is no strong linear relationship between dataset size and validation accuracy or loss across the evaluated models and datasets. Finally, we observe that the training loss is consistently lower than the validation loss across all models and datasets, indicating a mild degree of overfitting. However, the gap between training and validation losses is not severe, suggesting that the models retain generalization capability under the current training pipeline and data configurations.

6 Conclusion and Future Work

For the next-node prediction task, Activity and Hub-and-Spoke datasets are easiest, achieving around 94% accuracy across models. The Zone dataset is hardest, especially for GNNs, with accuracy dropping to 85.94%, likely due to limited data and complex spatial patterns. Dataset size alone does not determine performance, as smaller Activity data outperforms the larger Random dataset. Mild overfitting is observed but models generalize well. Future work will incorporate richer node features, spatial and temporal context, and traffic data to improve accuracy. Additionally, further evaluation using K-Fold cross-validation and model averaging is needed, but was not feasible here due to the long training times required for these preliminary results.

References

1. Hannah Bast, Daniel Delling, Andrew Goldberg, Matthias Mueller-Hannemann, Thomas Pajor, Peter Sanders, Dorothea Wagner, and Renato F Werneck. Route planning in transportation networks. *Algorithm Engineering*, pages 19–80, 2015.
2. Ye Hong, Henry Martin, and Martin Raubal. How do you go where?: Improving next location prediction by learning travel mode information using transformers. In *Proceedings of the 30th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pages 61–70, 2022.
3. Guangtong Hu and Jun Zhang. Origin-destination prediction from road average speed data using graphreslstm model. *PeerJ Computer Science*, 11:e2709, 2025.
4. Sven Maerivoet and Bart De Moor. Transportation planning and traffic flow models, 2005.
5. Luis M. Martínez, José Manuel Viegas, and Elisabete A. Silva. A traffic analysis zone definition: a new methodology and algorithm. *Transportation*, 36(5):581–599, 2009.
6. Brussels Mobility. Enquête mobilité bruxelloise (ovg/ecd) 2021–2022. `https://mobilite-mobiliteit.brussels`. Accessed: 2025-07-10.
7. Yufei Shi, Haiyan Tao, and Li Zhuo. Combining the spatiotemporal mobility patterns and mmc for next location prediction of fake base stations. *Computational Urban Science*, 4:20, 2024.
8. Sandor M Veres and Moustafa Moussa. Deep learning for intelligent transportation systems: A survey of emerging trends. *IEEE Transactions on Intelligent Transportation Systems*, 21(8):3133–3146, 2020.

9. Xiao-Yong Yan, Tao Zhou, Haijun Li, Jingyuan Wang, Weiqing Song, Yong Liu, Kunfeng Wang, Yingqiu Zhang, Song Zhang, Wei Wei, Xiao-Kai Hu, Guang Wu, Ruijia Zhang, Liang-Cheng Su, Weixiao Song, Chaoming Song, and Manuel Cebrian. Destination choice game: A spatial interaction theory on human mobility. *Scientific Reports*, 9(1):7662, 2019.
10. Jian Yu, Lucas Guo, Jiayu Zhang, and Guiling Wang. A survey on graph neural network-based next poi recommendation for smart cities. *Journal of Reliable Intelligent Environments*, 10:299–318, 2024.