Dynamic Route-Aware Graph Neural Networks for Accurate ETA Prediction

Guy Tordjman, and Nadav Voloch

***Abstract*—Accurate Estimated Time of Arrival (ETA) prediction is crucial for modern navigation systems and intelligent transportation applications. While existing approaches rely on static routing algorithms or machine learning models that infer likely paths between origin and destination, they fail to leverage the user's actual intended route and the dynamic nature of traffic interactions. This paper introduces a novel Dynamic Graph Neural Network (DGNN) architecture that explicitly incorporates user route intent and temporal memory to achieve superior ETA prediction accuracy.**

Our approach represents the transportation network as a dynamic multi-relational graph where junctions and vehicles are nodes, connected by typed edges capturing road segments, vehicle traversals, traffic interactions, and route intentions. The model employs a mixture-of-experts architecture with specialized experts and a memory layer that records historical snapshots to capture temporal dependencies. Unlike existing benchmarks that lack explicit route representation, our dataset unifies junction states, vehicle dynamics, and pre-planned route information, spanning ˜927,000 trips with durations ranging from 3 to 78 minutes and 80,640 graph snapshots collected over four simulated weeks.

Experimental results on simulated traffic data demonstrate the effectiveness of our approach, achieving a best validation error of MAE = 24.6 seconds for short trips, 39.6 seconds for medium trips, and 80.2 seconds for long trips (overall MAE $\approx$ 47.8 seconds). Progressive ablation studies highlight the contribution of each modeling component: transitioning from a structural baseline to interaction-aware, demand-aware, and finally the full intent-aware model yields consistent improvements, with long trips remaining the most challenging. These findings confirm that explicit route representation, dynamic node modeling, and expressive edge features are the primary factors enabling accurate ETA prediction, while the mixture-of-experts framework further enhances robustness under heterogeneous traffic regimes and dynamic conditions.

***Index Terms*—ETA prediction, graph neural networks, spatio-temporal graphs, user intent, traffic forecasting**

I. Introduction

Estimated Time of Arrival (ETA) prediction is a fundamental component of modern navigation systems and intelligent transportation applications. Accurate ETA enables commuters to make informed travel decisions, supports fleet management and logistics operations, and reduces congestion by distributing demand across the road network. With the ubiquity of GPS-enabled devices and connected vehicles, navigation platforms such as Google Maps and Waze [1]–[3] have transformed how travelers plan and adapt their journeys. Despite this progress, ETA estimation remains challenging due to the dynamic and stochastic nature of urban traffic.

G. Tordjman is with The Open University, Israel.
N. Voloch is with Ruppin Academic Center, Israel.

Traditional approaches rely on shortest-path algorithms such as Dijkstra's algorithm [4], which efficiently compute routes under static conditions but fail to capture evolving congestion or vehicle interactions. As a result, travel times obtained from such methods can diverge substantially from reality when conditions change during a trip.

With the growth of urban mobility datasets, machine-learning models have been applied to ETA prediction. Tree-based methods such as XGBoost [5] leverage handcrafted features and perform well on aggregated trip records, including NYC Taxi and Porto Taxi [6], [7]. Deep learning approaches improve accuracy by modeling temporal patterns and spatio-temporal context along a route. For example, DeepTTE learns ETA from raw GPS traces [8], TADNM incorporates transportation-mode awareness [9], MetaTTE applies meta-learning for cross-city generalization [10], and STAD corrects routing-engine outputs using spatio-temporal adjustments [11].

In parallel, graph-based spatio-temporal learning has advanced traffic prediction by representing road networks as graphs and capturing dynamics through diffusion or convolution operators (e.g., DCRNN, ST-GCN, Graph WaveNet) [12]–[14]. However, existing benchmarks such as NYC, Porto, Chengdu/DiDi, and Geolife [6], [7], [15], [16] lack explicit representation of pre-planned routes. Models must therefore infer likely paths between origin and destination, introducing ambiguity and reducing accuracy. Prior work has proposed incorporating future-traffic–aware route selection to improve ETA estimation [17]. We extend this direction by explicitly modeling pre-computed source–destination routes within a dynamic spatio-temporal graph, combined with a graph-based mixture-of-experts architecture.

The main contributions of this work are as follows:

- **Dataset:** We construct a dynamic graph-based dataset that unifies junction states, vehicle dynamics, and explicit pre-planned route information.
- **Model:** We propose a Dynamic Graph Neural Network (DGNN) architecture that integrates graph attention, temporal encoding, and mixture-of-experts specialization for accurate ETA prediction.
- **Evaluation:** We conduct comprehensive experiments on simulated traffic data, including ablation studies, demonstrating that explicit route representation yields substantial improvements compared to route-agnostic variants.

II. Related Work

A. Classical pathfinding and routing

On static graphs $G = (V, E)$ with non-negative edge costs, shortest paths are classically computed via Dijkstra's algorithm [4]. Alternatives such as Bellman–Ford [18] and bidirectional search [19] offer different trade-offs between generality and efficiency. To enable web-scale applications, numerous speed-up techniques have been developed. These include A*/ALT landmarks [20], which guide search with precomputed heuristics; Contraction Hierarchies [21], which accelerate queries by shortcutting less important edges; highway-dimension theory [22], which formalizes why sparse "highway" structures allow fast queries; and Customizable Route Planning [23], which separates preprocessing from query time to support rapid updates. Extensions such as multi-criteria routing and k-shortest path algorithms [24] broaden the option set by considering cost trade-offs or alternatives. While these methods are efficient, travel-time estimates are typically obtained from static or exogenous edge-time models. As a result, they cannot account for real-time congestion, stochastic traffic dynamics, or interactions between vehicles, limiting their predictive accuracy.

B. Learning-based ETA and traffic forecasting

The proliferation of large-scale urban mobility datasets has motivated learning-based approaches to ETA prediction. Classical machine-learning models such as gradient boosting [5] leverage handcrafted features (e.g., trip distance, time-of-day, origin-destination statistics) and have shown strong performance on aggregated trip datasets, notably NYC Taxi and Porto Taxi [6], [7]. These methods are simple and interpretable, but their reliance on aggregate features prevents them from capturing fine-grained spatio-temporal dependencies.

Deep learning approaches address this gap by directly modeling temporal patterns and route context. DeepTTE learns ETA from raw GPS traces with recurrent and convolutional layers [8]. TADNM incorporates transportation-mode awareness to adapt predictions across different mobility contexts [9]. MetaTTE applies meta-learning techniques for cross-city generalization, allowing a model trained in one region to adapt quickly to another [10]. STAD refines routing-engine outputs by applying spatio-temporal corrections to the estimated times [11]. These models demonstrate that neural networks can effectively capture sequential and contextual features beyond static trip attributes.

In parallel, spatio-temporal graph neural networks (STGNNs) have advanced traffic prediction by explicitly representing road networks as graphs. Diffusion Convolutional Recurrent Neural Networks (DCRNN) [12], Spatio-Temporal Graph Convolutional Networks (ST-GCN) [13], and Graph WaveNet [14] propagate information across road graphs to capture local correlations and temporal evolution. These models achieve state-of-the-art results in short-term traffic forecasting, such as predicting traffic flow or speed on sensor-equipped road segments. However, widely used benchmarks (NYC Taxi, Porto Taxi, Chengdu/DiDi, Geolife) [6], [7], [15], [16] lack explicit representation of pre-planned routes. Models trained on these datasets must infer the likely path between an origin and destination, which introduces ambiguity and reduces ETA prediction accuracy. Moreover, benchmark datasets are often limited to aggregated trip records or fixed sensor locations, restricting their ability to model vehicle-level interactions.

C. Intent- and route-aware prediction

Accurately modeling route intent is critical for ETA prediction, as travel time depends strongly on the chosen path. Prior work has attempted to address this in several ways. Voloch et al. [17] incorporate future-traffic–aware route selection, showing that improved routing decisions can reduce ETA errors. In the trajectory-prediction literature, destination-aware models leverage partial trajectories and contextual cues to infer intent, yielding better predictions of future positions. Similarly, intent modeling has been applied in ride-hailing and fleet management scenarios, where anticipated destinations influence dispatch and rebalancing strategies. These approaches demonstrate that explicitly reasoning about intent can improve both accuracy and downstream decision-making.

Nevertheless, most intent-aware works either focus on trajectory forecasting or optimize route selection strategies, rather than integrating intent into the ETA prediction model itself. Furthermore, they typically rely on datasets that lack explicit representations of the complete pre-planned route. In contrast, our work directly integrates pre-computed source–destination routes into a dynamic spatio-temporal graph framework. By representing both junctions and vehicles as nodes, and including typed edges for road segments, vehicle interactions, and route associations, we unify structural, dynamic, and intent-related features in a single model. This design bridges the gap between routing, trajectory forecasting, and ETA prediction, enabling fine-grained spatio-temporal learning conditioned on actual paths rather than inferred ones.

III. Methodology

A. Dynamic Graph Representation

We represent the transportation system at snapshot time t as a directed multi-relational graph

$$G_t = \big(V_t, E_t, \{A_t^{(\rho)}, W_t^{(\rho)}\}_{\rho\in\mathcal{R}}\big),$$

where $V_t = V^j \cup V_t^v$ consists of static junction nodes V^j and dynamic vehicle nodes V_t^v. Each relation $\rho \in \mathcal{R}$ corresponds to a typed edge set with optional weights. We define four relation types:

1) **Road edges ($\rho =$ road):** For adjacent junctions (u, v) with legal direction $u \to v$, we add $(u, v) \in E_t^{(\text{road})}$, encoding static road topology.
2) **Traversal edges ($\rho =$ trav):** If vehicle v_i^t occupies segment (a, b), we add (a, v_i^t) and (v_i^t, b), linking the vehicle to its upstream and downstream junctions.
3) **Interaction edges ($\rho =$ inter):** For vehicles v_i^t, v_j^t on the same segment with spacing $d_{ij}(t) \le \varepsilon$ and aligned headings, we add (v_i^t, v_j^t) weighted by $\omega(d_{ij}(t)) = e^{-d_{ij}(t)/\lambda}$.
4) **Intent edges ($\rho =$ intent):** Given a pre-computed route $\mathcal{R}_i = (e_1, \dots, e_K)$, vehicle v_i^t is connected to upcoming junctions with decaying weights α_k, encoding future intent.

Each node and edge carries feature vectors capturing static attributes (e.g., number of lanes, road type), dynamic states (e.g., vehicle speed, occupancy), and route-related information.

B. Temporal Windowing

ETA prediction requires reasoning over temporal dynamics. We construct a window of H consecutive snapshots:

$$\mathcal{G}_{t-H+1:t} = \{G_\tau\}_{\tau=t-H+1}^{t},$$

where $H = 30$ in our experiments (30-second interval). The prediction target is the ETA of vehicles present in the final snapshot G_t.

C. Model Architecture

Our Dynamic Graph Neural Network (DGNN) integrates spatial, temporal, and route information. The workflow is illustrated in Fig. 1.

- **Graph Encoder:** Each snapshot G_t is processed by a multi-layer GATv2-based encoder with residual connections and edge features, producing embeddings for junction and vehicle nodes.
- **Temporal Aggregator:** Encoded snapshots are fed to a temporal module. In our current implementation, prediction is made at the final snapshot (t), but the architecture supports GRU/Transformer-based sequence aggregation.
- **Vehicle Selection:** Only vehicle node embeddings at t are retained for prediction.
- **Route Encoder:** For the full intent-aware variant, each vehicle's remaining route is embedded via an edge-ID embedding with mean pooling, then concatenated with its vehicle embedding.
- **Fusion and MoE Head:** Vehicle embeddings are fused through a feed-forward network and routed to a sparse Top-k Mixture-of-Experts (MoE) head, where specialized experts capture heterogeneous traffic regimes. The output is the ETA prediction for each vehicle.

D. Loss Functions and Ablation Variants

We train the model with supervised regression on ETA targets using mean absolute error (MAE) in seconds:

$$\mathcal{L}_{\text{MAE}} = \frac{1}{N}\sum_{i=1}^{N} \left|\hat{y}_i - y_i\right|,$$

where $\hat{y}_i$ and y_i are the predicted and true ETAs. Additional metrics include RMSE, weighted absolute percentage error (WAPE), and percentile errors (P50, P90, P95).

To assess the contribution of each component, we design four ablation variants:

- **None (structural baseline):** Road edges only; static features.
- **Weak (interaction-aware):** Adds dynamic traversal and interaction edges.
- **Medium (demand-aware):** Adds aggregate route-related features (e.g., edge route counts).
- **Full (intent-aware):** Adds explicit route encoder over remaining vehicle paths.

All variants share the same training protocol and data splits, enabling a controlled comparison of structural, dynamic, demand, and intent features.

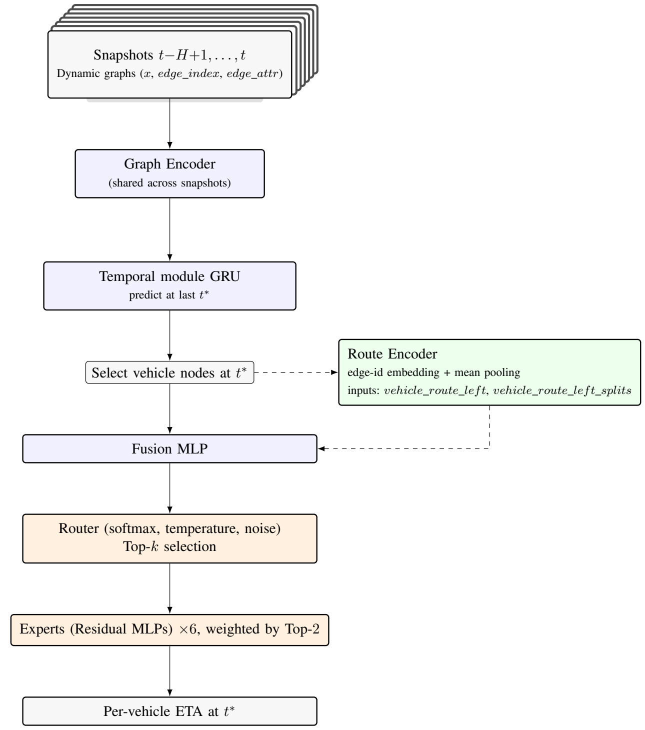

Fig. 1. Temporal MoE ETA model. A window of T=30 dynamic graph snapshots (shown as an overlapped deck) is processed by a *shared* Graph Encoder; prediction is made at the last snapshot t^*. In the Full variant, a Route Encoder summarizes each vehicle's remaining path before Fusion and a Top-k MoE head produces per-vehicle ETA.

IV. Experiments

We evaluate on simulated urban traffic with explicit route plans per trip. We report MAE (seconds) and relative improvements vs. baselines. Baselines include: (i) static route-sum ETA; (ii) XGBoost with handcrafted features [5]; (iii) deep sequence models (e.g., DeepTTE [8]); and (iv) graph models (DCRNN, ST-GCN, Graph WaveNet) [12]–[14]. We ablate intent edges and the MoE router.

Implementation uses PyTorch Geometric [25]. Training details: optimizer, learning rate, window H, batch size, and early stopping. We stratify results by trip duration bins to analyze regimes.

V. Results

Our DGNN achieves lower MAE than route-unaware baselines across all duration bins. Intent edges consistently reduce errors for longer trips. The MoE router exhibits meaningful specialization (measured via entropy and gate usage), contributing to accuracy gains.

VI. Discussion

Route intent reduces path ambiguity and aligns ETA modeling with user plans. Limitations include simulation-to-reality gaps and potential sensor/route noise. Future work: real-world datasets with planned routes, robust intent inference, and uncertainty-aware ETAs.

VII. Conclusion

We presented a dynamic, intent-aware graph approach to ETA prediction, combining spatio-temporal encoding with a mixture-of-experts. Experiments indicate that explicit route intent and expert specialization improve accuracy. Future work will validate on real-world data and explore policy feedback between ETA and route selection.

Acknowledgments

We thank collaborators and institutions supporting this research.

References

[1] A. Derrow-Pinion, J. She, D. Wong, O. Lange, T. Hester, L. Perez, M. Nunkesser, S. Lee, X. Guo, B. Wiltshire, P. W. Battaglia, V. Gupta, A. Li, Z. Xu, A. Sanchez-Gonzalez, Y. Li, and P. Veličković, "Eta prediction with graph neural networks in google maps," in *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, 2021, pp. e.g., 4034–4042. [Online]. Available: https://arxiv.org/abs/2108.11482

[2] N. Hoseinzadeh, Y. Liu, L. D. Han, C. Brakewood, and A. Mohammadnazar, "Quality of location-based crowdsourced speed data on surface streets: A case study of waze and bluetooth speed data in sevierville, tn," *Computers, Environment and Urban Systems*, vol. 83, p. 101518, 2020.

[3] M. Amin-Naseri, P. Chakraborty, A. Sharma, S. B. Gilbert, and M. Hong, "Evaluating the reliability, coverage, and added value of crowdsourced traffic incident reports from waze," *Transportation Research Record*, vol. 2672, no. 43, pp. 34–43, 2018.

[4] E. W. Dijkstra, "A note on two problems in connexion with graphs," *Numerische Mathematik*, vol. 1, no. 1, pp. 269–271, 1959.

[5] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2016, pp. 785–794.

[6] NYC TLC, "New york city taxi and limousine commission trip record data," https://www.nyc.gov/site/tlc/about/tlc-trip-record-data.page, 2013–, accessed: 2025-08-26.

[7] L. Moreira-Matias, J. Gama, M. Ferreira, J. Mendes-Moreira, and L. Damas, "Predicting taxi–passenger demand using streaming data," in *2013 IEEE 16th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2013, pp. 140–145.

[8] J. Wang, Y. Fu, and Z. Zhang, "When will you arrive? estimating travel time based on deep neural networks," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2018, pp. 2500–2507.

[9] W. Xu, Z. Lin, Y. Zhao, T. Zhang, and B. Yang, "Tadnm: A transportation-mode aware deep neural model for travel time estimation," *Applied Sciences*, vol. 10, no. 21, p. 7599, 2020.

[10] X. Wang, J. Li, and N. J. Yuan, "Metatte: A meta-learning framework for travel time estimation," in *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. ACM, 2022, pp. 4080–4088.

[11] S. Abbar, A. Anagnostopoulos, S. Bhagat, P. Cudre-Mauroux, and A. Kumar, "Stad: Spatio-temporal adjustment for improving travel-time estimation," in *Proceedings of the Web Conference 2020*. ACM, 2020, pp. 2839–2845.

[12] Y. Li, R. Yu, C. Shahabi, and Y. Liu, "Diffusion convolutional recurrent neural network: Data-driven traffic forecasting," in *International Conference on Learning Representations*, 2018.

[13] B. Yu, H. Yin, and Z. Zhu, "Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting," in *Proceedings of the 27th International Joint Conference on Artificial Intelligence*, 2018, pp. 3634–3640.

[14] Z. Wu, S. Pan, G. Long, J. Jiang, and C. Zhang, "Graph wavenet for deep spatial-temporal graph modeling," in *Proceedings of the 28th International Joint Conference on Artificial Intelligence*, 2019, pp. 1907–1913.

[15] DiDi Chuxing Research, "Di-tech challenge 2016," https://outreach.didichuxing.com/research/opendata/en/, 2016, accessed: 2025-08-26.

[16] Y. Zheng, L. Zhang, Z. Ma, X. Xie, and W.-Y. Ma, "Mining interesting locations and travel sequences from gps trajectories," in *Proceedings of the 18th international conference on World wide web*. ACM, 2009, pp. 791–800.

[17] N. Voloch and N. Voloch-Bloch, "Finding the fastest navigation route by real-time future traffic estimations," *2021 IEEE International Conference on Microwaves, Communications, Antennas and Electronic Systems (COMCAS)*, 2021.

[18] R. Bellman, "On a routing problem," *Quarterly of Applied Mathematics*, vol. 16, pp. 87–90, 1958.

[19] I. Pohl, "Bi-directional search," *Machine Intelligence*, vol. 6, pp. 127–140, 1971.

[20] A. V. Goldberg and C. Harrelson, "Computing the shortest path: A* search meets graph theory," in *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2005, pp. 156–165.

[21] R. Geisberger, P. Sanders, D. Schultes, and D. Delling, "Contraction hierarchies: Faster and simpler hierarchical routing in road networks," in *International Workshop on Experimental and Efficient Algorithms (WEA)*. Springer, 2008, pp. 319–333.

[22] I. Abraham, A. Fiat, A. V. Goldberg, and R. F. Werneck, "Highway dimension, shortest paths, and provably efficient algorithms," in *Proceedings of the 21st Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2010, pp. 782–793.

[23] D. Delling, A. V. Goldberg, T. Pajor, and R. F. Werneck, "Customizable route planning," in *International Symposium on Experimental Algorithms (SEA)*. Springer, 2011, pp. 376–387.

[24] J. Y. Yen, "Finding the k shortest loopless paths in a network," in *Management Science*, vol. 17, no. 11, 1971, pp. 712–716.

[25] M. Fey and J. E. Lenssen, "Pytorch geometric: A library for graph deep learning," in *ICLR Workshop on Representation Learning on Graphs and Manifolds*, 2019.