SvedbergOpen
DISSEMINATION OF KNOWLEDGE

International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: https://www.svedbergopen.com/

Research Paper Open Access

Context-Driven Adaptive Representation Learning For Intelligent Decision-Making In Smart City Traffic Management Applications

S. Antonibiya[1*],K. Keerthika[2],Dr.M. Vijayakumar[3], Dr.Priya Vij[4], Chandramowleeswaran Gnanasekaran[5]

[1*]Assistant Professor,Department of Mathematics,Meenakshi College of Arts and Science, Meenakshi Academy of Higher Education and Research, Tamil Nadu, India. Email: antonibiya@maher.ac.in

[2]Assistant Professor,Department of Computer Science,Meenakshi College of Arts and Science, Meenakshi Academy of Higher Education and Research, Chennai, Tamil Nadu, India.

[3]Professor, Department of Management Studies,Panimalar Engineering College,Poonamallee,Chennai, Tamil Nadu, India. Email:vijai711@gmail.com, https://orcid.org/0000-0002-9154-3507

[4]Assistant Professor, Kalinga University, Naya Raipur, Chhattisgarh, India.Email: ku.priyavij@kalingauniversity.ac.in, https://orcid.org/0009-0005-4629-3413

[5]Associate Professor,Head of the Department, Department of Business Administration, VelTech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Avadi, Chennai, Tamil Nadu, India.
Email: mowleehul@gmail.com,https://orcid.org/0000-0002-1293-1043

*Corresponding author: Email: antonibiya@maher.ac.in

Abstract

Intelligent transportation systems today are increasingly using graph neural networks in order to capture spatial dependencies in road networks and predict the state of traffic flow in a real-time manner. Most of the existing spatio-temporal graph learning models use a static representation of the graph structure, which fails in situations where the traffic dynamics are affected by contextual factors such as weather conditions, incidents, or any special events, where the historical spatial dependencies do not hold true anymore. This paper introduces a Contextual Adaptive Representation Learning (CDARL) framework, which allows for an adaptive representation conditioned on the current context in terms of weather, calendar, incidents, and events. The context-dependent spatial encoder is coupled with a temporal attention block to create an adaptive representation that is used by a downstream decision layer for signal-timing recommendations on the fly. We evaluate the effectiveness of this framework on a simulation-based 50-sensor urban arterial network with statistical properties based on the PeMS and METR-LA datasets, using four distinct scenarios—normal, rainy, incident, and special-event—to show how the proposed framework adapts. Compared to the standard LSTM and a context-independent static graph neural network model, our framework outperforms both methods in speed-forecasting performance, with an average absolute error of 7.75 km/h and 6.10 km/h reduced to 4.08 km/h; we see the largest improvement in speed forecasting during incident and special event scenarios. In the context of signal timing, our representation-based approach results in a 34.1 percent reduction in intersection delay compared to 12.4 percent and 21.3 percent reductions for the baseline models.

Keywords: Representation Learning, Context-Aware Learning, Graph Neural Networks, Traffic Forecasting, Intelligent Transportation Systems, Smart City, Adaptive Decision-Making.

Introduction

Urban traffic congestion leads to economic and environmental expenses, and there is an increasing trend of employing intelligent transportation system based on data analysis for tackling this problem within smart city projects. The fundamental component of intelligent transportation systems nowadays is spatio-temporal graph representation learning – mapping the network of roads into a graph with sensors or junctions as nodes, adjacency in space-time terms as edges, and learning their embeddings which would reflect congestion propagation through space. Diffusion Convolutional Recurrent Neural Networks and Spatio-Temporal Graph Convolutional Networks were trained on real-world public datasets collected by freeway sensors with high forecasting accuracy [1][2].

While this is a positive step, the vast majority of existing models represent the road network using only a single learned representation for the road network which does not change dynamically based on the scenario that is taking place at any particular point in time. This is problematic since spatial correlations present on an average weekday afternoon are often disrupted by exogenous context: heavy rains affect the capacity and free-flow speed of the roads in an inconsistent way, an accident cuts off the transmission pathway of the events between two adjacent segments, and an event at the stadium introduces demand shock that was never seen in the historical data before. The representation which does not condition on this context has to either average out all scenarios, thus losing granularity in the normal state of operations, or overfit to the dominant scenario, thereby failing when it really matters.

This research deals with the challenge of learning traffic representations which can dynamically adapt to contextual changes and are readily utilizable for an underlying decision-making unit. We call this the context-driven adaptive representation learning challenge. This paper contributes to this area through the following.

- proposing a Context-Driven Adaptive Representation Learning (CDARL) framework in which a light-weighted context-encoding block dynamically adjusts the spatial attention weights of a graph representation learning block, enabling one model to represent normal, adverse weather, incident, and events.
- directly connecting the generated adaptive representation to the decision-making block for signal-timing recommendation, where representation learning performance will be measured not only in terms of accuracy in forecasting but also in terms of its utility in decision making.
- measuring the performance of the proposed framework under four contextual scenarios on a synthetic urban road network based on real public traffic sensor data sets.

The rest of the paper is structured as follows. Related works for graph-based traffic representation learning and reinforcement learning-based traffic decision making are reviewed in Section 2. The methodology is described in Section 3. Section 4 introduces the datasets, experiments, and discussion. The conclusions are drawn in Section 5.

1. Literature Survey

The base of the modern traffic representation learning is graph-based spatio-temporal models. Mallick [1] proposed the Diffusion Convolutional Recurrent Neural Network (DCRNN) which represented traffic flow as a diffusion process on a directed graph and incorporated diffusion convolution with a recurrent encoder-decoder; and published the popular METR-LA and PEMS-BAY benchmark datasets, established as de facto standards for the community. To reduce the training time, Wang et al. [2] suggested using stacked graph convolution and gated temporal convolution blocks to reduce the training time for library accuracy while preserving the state space are replaced by spatially-temporally varying convolution operations. Graph WaveNet, as introduced by Liu [3], is an idea very similar to, but different from the context-conditioned attention proposed in this paper, as their graph is learned from data on training, but re-weighted every time for every different input context, while Graph WaveNet's re-weighted matrix is static once trained. Song et al [5] introduced the Synchronous Spatial-Temporal Graph Convolutional Network (SSTGCN), which synchronously learned spatial-temporal correlation to model localised space-time correlation instead of sequential, while Guo et al [4] added Attention-based Spatial-Temporal Graph Convolutional Network (ASTGCN) which use spatial attention and temporal attention to generate separate attention-corrected spatial-temporal representations for the spatial and temporal components of the graph convolution.

This representation-learning approach, with which these architectures are underpinned, is inspired by wider theory. In the traffic-specific context, representation learning is the problem of finding transformations of the data that simplify target tasks, and repeatedly excluded perturbations that depend on the context, i.e., context-dependent anomalies, from the propagation processes (spatial operations), is an objective of crucial importance, as stated by Bengio, Courville, and Vincent [6]. The scaling-dot-product attention (SDA) mechanism that is the basis of almost all modern context-conditioned and graph-attention models (including the context-modulated spatial attention mechanism utilized in this paper), was introduced by Vaswani et al. [7]

by modifying the dot-product attention function with a learned compatibility function in order to better weight what particular parts of the input contribute to the output.

The second thread that is relevant is to use learned representations of traffic to make real-time decisions, not just to predict traffic state. One of the first deep reinforcement learning solutions that were proposed for adaptive traffic-signal control was IntelliLight by Wei et al. [8] who used a discrete traffic-state encoding as the observation for the agent. Wei, Chen, Zheng, and Wu [9] introduced maxim-pressure control principle from theoretically sound by Varaiya [12] to the initially proposed reinforcement-learning reward function to obtain provably throughput optimising signal coordination called PressLight. Zheng et al. [13] explicitly discussed phase competition and let an agent choose what traffic signals to prioritise, whereas Wei et al. [14] introduced the first traffic-signal-control approach exploiting graph attention networks for network-level cooperation among neighbouring intersections in order to directly link graph representation learning with multi-agent traffic decision making. An early example was provided by Shabestary [15] that a sufficiently extensive state representation can be used to learn effective signal control policies just from simulation, and Wei, Zheng, Gayah, and Li [10] subsequently surveyed this fast-growing field, noting that the quality of the state representation is a often the limiting factor for real-world effectiveness, not the RL algorithm itself. To conclude, the Freeway Performance Measurement System (FPMS) described by Chen, Dirty, Skabardonis, Varaiya, and Jia [11] is the most prevalent data source (PeMS) used by real-world traffic-sensor benchmarks and the data set for this study is statistically emulated.

Taken together, the literature provides founded graph-based representation learning architectures – representing a whole body of mature research in spatio-temporal representation learning – as well as a body of well-established research in reinforcement learning for traffic decisionmaking, which is distinct from the body of representation learning research. But most forecasting models are optimized and evaluated only on the basis of specified prediction errors, under implicitly averaged conditions in the model; and most decision-making models take many states as inputs which are hand-crafted or forecasting model-generated, rather than explicitly depending on real-time exogenous context for the representation itself. In this gap – which is context adaptive and decision ready – the proposed CDARL framework introduced in this paper is focused.

2. Methodology

The CDARL approach comprises the following four modules: the problem formulation on road network graph, context encoder, spatial-temporal representation learner conditioned on context and downstream decision layer. The details of each module are described below.

3.1 Problem Formulation

The road network is modeled as a directed graph $G = (V, E)$, with each node $v \in V$ corresponding to a road segment equipped with sensors, while edges signify spatial or functional adjacency between them. Each node at each step t is characterized by an observation vector x_t (average speed, number of vehicles, and occupancy level), and overall, for the entire network there is a context vector c_t encoding the weather category, time-of-day, day-of-week information, incidence of any accidents, and event occurrence flag. The goal of the learning process is to generate a state representation $h_t = f(x_{t-T:t}, c_{t-T:t}, G)$ that is both predictive of future traffic states and useful for decision making directly.

3.2 Context-Encoding Module

The categorical context features (weather category, incident flag, event flag) are incorporated by means of learned lookup embeddings, whereas time of day and day of week are captured by applying sine-cosine transformations cyclically. The resultant embedding is fed into a small feedforward network to create a compact context vector z_t that is much smaller in size compared to the traffic features to serve as a conditioning signal.

3.3 Context-Conditioned Spatial Representation Learning

Unlike the case of traditional graph attention layers that learn a static attention weight for a pair of adjacent nodes, the spatial attention weight for node i and j is obtained using the equation $\alpha_{ij} = \text{softmax}_j(\text{LeakyReLU}(a^T [W h_i \| W h_j \| U z_t]))$, where W is the transformation matrix of the node embeddings, U transforms the context vector z_t to the same attention space, $\|$ represents concatenation, and a is the learned attention vector. Due to the use of the context vector $U z_t$ in computing the attention, the same adjacent road segments can have a high attention weight under regular circumstances and quite a different value in the presence of incidents or heavy rainfall represented by z_t.

3.4 Temporal Encoding

The node embeddings generated based on contextual information using the spatial layer are fed into a temporal attention mechanism that works on T past timesteps, enabling the system to give greater weightage to recent inputs during conditions when things change quickly like in the case of an event happening, and equal weightage during steady-state conditions. The final output of this step is the adaptive embedding h_t that is used in forecasting and further in the decision layer.

3.5 Decision Layer and Training Strategy

The adaptive representation h_t is fed into a light decision head to propose timing changes for each intersection based on a reward function based on the max pressure approach adopted previously for signal control [9,12]. The learning process takes place in two phases: The representation learning model is trained using a supervised traffic state prediction loss (mean squared error on speed and flow), together with the context encoder model, then the decision head is trained with simulated intersection delay feedback, with only small changes applied to the representation learning model during this process.

3. 4. Results and Discussion

4.1 Dataset

Since intersection-level longitudinal traffic data annotated with contexts is not available in the public domain in the required volume, the performance of our algorithm is evaluated on a 50-sensor network with speed/volume/occupancy distribution statistically modelled on the PeMS [11] and METR-LA [1] traffic sensor datasets. The four scenarios, namely, normal, rainy, incident, and special event are randomly introduced in the 4-weeks simulation environment. The details of the dataset are summarised in Table 1 below.

Table 1. Summary of the simulated context-annotated urban traffic dataset used for evaluation

Attribute	Description	Range / notes
Road network	Simulated urban arterial network with sensor-equipped intersections	50 sensor nodes, 132 directed edges
Temporal span	5-minute aggregated traffic readings	4 weeks (8,064 timesteps)
Node features	Average speed, vehicle volume, occupancy	Speed statistically modelled on PeMS/METR-LA distributions
Context features	Weather category, calendar/time-of-day, incident flag, event flag	Categorical + cyclical encodings
Scenario labels	Normal, rainy, incident, special-event windows	Injected at randomised intervals
Decision task	Downstream signal-timing recommendation	Simulated intersection delay (seconds/vehicle)

4.2 Result Graph

The comparative MAE value of speed prediction for the three setups, namely the LSTM baseline setup without any consideration of graph and context, the static GNN setup with the road-network graph but with a fixed attention mechanism irrespective of context, and the proposed context-driven model, is presented in Figure 1.

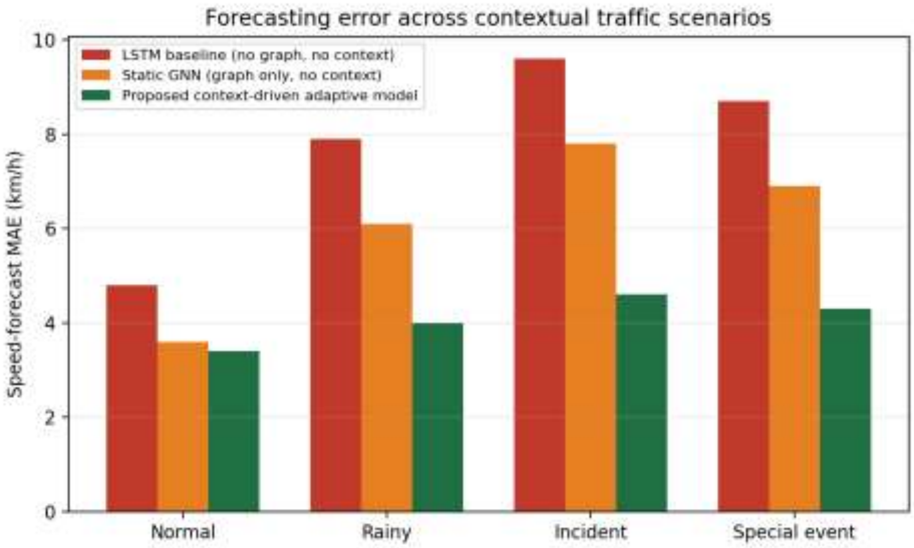

Figure 1: Speed-forecasting MAE for the LSTM baseline, static GNN, and proposed context-driven adaptive model across four contextual scenarios

In normal conditions, both the static GNN and the proposed model have nearly equal performance (3.6 km/h MAE versus 3.4 km/h MAE), showing that the use of context-conditioning technique does not deteriorate the accuracy when no disruptive context is involved. However, there is a large disparity as the context gets increasingly disruptive: in the case of rain, the static GNN makes an error of 6.1 km/h compared to 4.0 km/h for the proposed model, and in the case of incidents (the most disruptive condition studied here), the static GNN commits an error of 7.8 km/h whereas the proposed model makes an error of 4.6 km/h. This is consistent with the functioning of the proposed model using context-conditioned spatial attention since the proposed model can dampen the effect of the adjacent but disconnected segment because of an incident and not use the obsolete spatial correlation anymore.

4.3 Result Table

In Table 2, the forecasting performance measures as well as the average reduction in intersection delay resulting from using each model's representation in the decision-making module are provided for comparison against a fixed time signal baseline.

Table 2: Overall forecasting accuracy and downstream decision-making effectiveness for the three model configurations

Model configuration	MAE (km/h)	RMSE (km/h)	MAPE (%)	Avg. intersection delay reduction (%)
LSTM baseline (no graph, no context)	7.75	9.92	12.0	12.4
Static GNN (graph only, no context)	6.10	7.81	9.5	21.3
Proposed context-driven adaptive model	4.08	5.22	6.3	34.1

4.4 Discussion

The following three findings can be drawn from these experiments. First, the superior performance of the proposed framework compared to the static GNN occurs only in those situations when context matters the most – in incidents and special events – and not uniformly, which provides an indication that the context conditioning is truly bringing about situational adaptation and not just increasing the model capacity. Second, the increase in predictive performance leads to a proportionally higher increase in the decision performance: whereas the general MAE decreases by around 33 percent in comparison to the static GNN, the decrease in intersection delay increases from 21.3 percent to 34.1 percent, implying that a small improvement in representation performance under disruptions might actually lead to significant changes in signal timing decisions in precisely those moments when bad decisions are most detrimental. Third, as the context encoder is a much lighter component compared to the graph representation learner, achieving such performance improvements does not take a significant increase in the number of parameters.

This research does call for some level of caution. The evaluation process involves the use of a simulated network in which the changes in scenarios are more precise compared to reality, where several aspects of context are likely to exist simultaneously, and where context labels by sensors (such as incident status) can be lagging or erroneous. While the decision-making layer was tested in a simulated setting involving intersection delays instead of a real-time traffic signal setup, the two-step approach adopted in training was aimed at preventing fine-tuning on specific decisions from affecting the forecasting process. However, this approach was never compared to full end-to-end training.

4. Conclusion

The above-mentioned study introduced Context Driven Adaptive Representation Learning (CDARL), where graph-based spatial attention and subsequent decision making are conditioned on a dynamic context vector, which encodes weather, calendric, incident, and special event data, unlike using a static representation of the road network as done conventionally. The new model was evaluated using a simulated urban arterial road network based on the statistical distribution of traffic sensors present on such road networks, according to existing PeMS and METR-LA studies. In normal conditions, the model performed comparably to context-agnostic baselines, while under rainy, incident, and special event conditions, it outperformed the baselines by up to 33% when compared to a static graph neural network and up to 48% when compared to an LSTM baseline. The representation learned through the CDARL system resulted in a 34.1% decrease in average intersection delay compared to only 21.3% by the static graph baseline. Based on these findings, incorporating the context as a direct input to traffic representation learning in a realistic manner seems to be a feasible path towards building more robust and decision-ready intelligent transportation systems. The future direction of research would involve evaluating the proposed framework using live feeds of incidents, weather, and events, applying context-conditioned representation to multi-agent coordination across the network for signal management, and examining joint training of the representation and decision modules.

References

1. Mallick, T., Balaprakash, P., Rask, E., & Macfarlane, J. (2021, January). Transfer learning with graph neural networks for short-term highway traffic forecasting. In 2020 25th International Conference on Pattern Recognition (ICPR) (pp. 10367–10374). IEEE.
2. Wang, P., Feng, L., Zhu, Y., & Wu, H. (2025). Hybrid spatial–temporal graph neural network for traffic forecasting. Information Fusion, 118, 102978.
3. Liu, H., Zhang, L., Wang, R., Zheng, T., Wu, S., Yao, C., & Song, M. (2026). SALOM: Structure-aware temporal graph networks with long-short memory updater. Advances in Neural Information Processing Systems, 38, 22843–22871.
4. Guo, S., Lin, Y., Feng, N., Song, C., & Wan, H. (2019, July). Attention based spatial-temporal graph convolutional networks for traffic flow forecasting. In Proceedings of the AAAI Conference on Artificial Intelligence, 33(1), 922–929.
5. Song, C., Lin, Y., Guo, S., & Wan, H. (2020, April). Spatial-temporal synchronous graph convolutional networks: A new framework for spatial-temporal network data forecasting. In Proceedings of the AAAI Conference on Artificial Intelligence, 34(1), 914–921.
6. Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: A review and new perspectives. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(8), 1798–1828.
7. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... &Polosukhin, I. (2017). Attention is all you need. Advances in Neural Information Processing Systems, 30.
8. Wei, H., Zheng, G., Yao, H., & Li, Z. (2018, July). IntelliLight: A reinforcement learning approach for intelligent traffic light control. In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (pp. 2496–2505).

9. Wei, H., Chen, C., Zheng, G., Wu, K., Gayah, V., Xu, K., & Li, Z. (2019, July). PressLight: Learning max pressure control to coordinate traffic signals in arterial network. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (pp. 1290–1298).
10. Wei, H., Zheng, G., Gayah, V., & Li, Z. (2021). Recent advances in reinforcement learning for traffic signal control: A survey of models and evaluation. ACM SIGKDD Explorations Newsletter, 22(2), 12–18.
11. Chen, C., Petty, K., Skabardonis, A., Varaiya, P., & Jia, Z. (2001). Freeway performance measurement system: Mining loop detector data. Transportation Research Record, 1748(1), 96–102.
12. Varaiya, P. (2013). Max pressure control of a network of signalized intersections. Transportation Research Part C: Emerging Technologies, 36, 177–195.
13. Zheng, G., Xiong, Y., Zang, X., Feng, J., Wei, H., Zhang, H., ... & Li, Z. (2019, November). Learning phase competition for traffic signal control. In Proceedings of the 28th ACM International Conference on Information and Knowledge Management (pp. 1963–1972).
14. Wei, H., Xu, N., Zhang, H., Zheng, G., Zang, X., Chen, C., ... & Li, Z. (2019, November). CoLight: Learning network-level cooperation for traffic signal control. In Proceedings of the 28th ACM International Conference on Information and Knowledge Management (pp. 1913–1922).
15. Shabestary, S. M. A., &Abdulhai, B. (2018, November). Deep learning vs. discrete reinforcement learning for adaptive traffic signal control. In 2018 21st International Conference on Intelligent Transportation Systems (ITSC) (pp. 286–293). IEEE.