Proceedings of the 4th International Conference on Mathematical Physics and Computational Simulation
DOI: 10.54254/2753-8818/2026.34028

Spatiotemporal Modeling and AI Integration for Traffic Signal Countdown Prediction

Wenchang Gu

Reading Academy, Nanjing University of Information Science and Technology, Nanjing, China
202383100044@nuist.edu.cn

Abstract. Urban transport relies on efficient distribution of fuel and minimization of greenhouse gas emissions. The incorporation of accurate Signal Phase and Timing (SPaT) predictions into navigation systems along with Green Light Optimized Speed Advisory (GLOSA) technology will improve the predictive mobility aspect of urban transport. This paper provides a comprehensive study of the evolution of such technologies from statistical and shallow machine learning to RNNs and GNNs as deep learning technologies, and offers a comparison of the various alternatives of deep learning techniques, in particular, EVs of the Transformers such as Informer and FE-Transformer, for the prediction of long-range spatiotemporal dependencies. Despite the reductions in prediction accuracy provided by deep learning, for commercial applications, the limitations of the latency in the end-to-end prediction and the adaptive control logic remain a barrier. This paper reviews the most recent models and examples of industry applications from Google and Baidu, and provides a comprehensive guide for the use of predictive SPaT in the design of Green Light Optimized Speed Advisory (GLOSA) technology. The tools and techniques describe the elements of a carbon neutral Intelligent Transportation System and provide a valuable foundation for developing predictive, dynamically adjusting, and integrated technologies for next-generation traffic control systems.

Keywords: Spatiotemporal Modelin, AI Integration, Traffic Signal Countdown Prediction

1. Introduction

Based Integrated Traffic Light Phases and Timing (SPaT) information with Vehicle Navigation Systems (VNS) created the ability to predict travel times and recommend optimal vehicle speeds to maintain green wave progression for a traffic light. Green wave progression improves operational efficiency of the road network and minimizes idling and the starting cycles of a vehicle. These cycles contribute additional fuel consumption and emissions.

The advancement of vehicle sensors and Vehicle to Everything (V2X) technology allows the transition from historical methods to predict traffic light cycles to the most current methods of using Transformative models for long term predictions. In 2024 and 2025, the utilization of large Floating Car data (FCD) from V2X technology and sensors will permit the ability to fine-grain predict at a large numbers of intersections with a small amount of infrastructure. This paper will outline the evolution of the forecasting duration predictions of VNS and provide a comprehensive study of

DOI: 10.54254/2753-8818/2026.34028

advanced Artificial Intelligence technologies and the performance forecasting of FCD data for long term traffic light cycles using Informer, PatchTST, and FE-iTransformer models [1, 2].

2. Basic big data methods for traffic light duration prediction

2.1. Statistical modeling and fixed time analysis

In the early stages of intelligent transportation development, the prediction of traffic light countdowns was based on the statistics of historical traffic data. Traditional traffic light control systems mostly adopt a fixed time-of-day schedule, that is, based on different times such as morning peak, evening peak and off-peak, a fixed signal cycle and phase ratio are preset.

In this context, early researchers often used signal processing methods such as Fast Fourier Transform (FFT) to periodically analyze the starting and stopping trajectories of vehicles at intersections. By extracting the frequency domain characteristics of vehicle passage times, the most likely traffic light operation cycle is identified [3]. In addition, the Autoregressive Moving Average (ARIMA) model and the Markov chain model are widely used to simulate the temporal fluctuations of traffic flow, and the traffic light switching point of the next cycle is estimated by using historical mean and autocorrelation [4]. These methods have been widely used in early traffic guidance systems due to their simple computational logic and strong interpretability.

2.2. Introduction of shallow machine learning methods

With the increasing abundance of data acquisition methods, machine learning methods, represented by gradient boosting decision trees (such as XGBoost), have begun to be applied to the SPaT prediction field. These methods can handle nonlinear features and perform multi-dimensional regression analysis on signal cycles by inputting vehicle probe data, road segment counter data, and time features (such as day of the week and whether it is a holiday) [5]. In some experimental cases, the average error of the XGBoost-based model in predicting the cycle length of adaptive traffic lights can be controlled at around 0.56 seconds, while the prediction error for the duration of red lights is about 7.2 seconds [5]. Although these indicators have been significantly improved compared to pure statistical methods, shallow learning models still have difficulty capturing long-term time features when faced with extremely complex non-stationary traffic sequences.

2.3. Limitations and challenges of traditional methods

Although traditional big data-based methods have the advantages of low implementation cost and fast training speed, they face the following bottlenecks in practical applications (Table 1).

Table 1. Conclusion of limitations and challenges of the traditional method

Limitation categories	Specific manifestations	Impact on prediction accuracy
Inductive/Adaptive Control Disturbance	The controller dynamically extends or shortens the phase based on real-time traffic flow [1].	This causes the model to be unable to accurately predict the duration shift of the next cycle using historical means.
Poor robustness to sudden interference	It lacks sensitivity to random events such as traffic accidents, road construction, and extreme weather [6].	The predicted values will deviate significantly from the actual phase, causing navigation suggestions to fail.

DOI: 10.54254/2753-8818/2026.34028

Table 1. (continued)

Data sparsity problem	Floating cars have insufficient penetration in non-busy road sections, and GPS drift is severe [3].	This results in the loss of a reference point for inverting the signal state, leading to significant random errors.
Spatial coupling missing	The logical connections between adjacent intersections in arterial coordination (green wave) are ignored.	Unable to achieve regional-level route optimization and speed guidance.

In other words, traditional statistical and shallow machine learning methods are inadequate for handling highly dynamic and uncertain urban traffic environments, prompting researchers to seek solutions from deeper neural network models.

3. Deep learning prediction methods: from RNN to GNN

3.1. Recurrent neural networks (RNNs) and long short-term memory networks (LSTMs)

Since the state of traffic lights is essentially a sequence that evolves over time, RNNs and their variants (such as LSTM and GRU) became the mainstream methods in the early stages of trying deep learning methods. LSTM effectively alleviates the gradient vanishing problem of traditional RNNs when dealing with long sequences by introducing forget gate and memory gate mechanisms [7].

In the SPaT prediction task, LSTM can comprehensively consider the phase sequence of the current cycle and multiple previous cycles to capture the small correction trend of the traffic light duration. Research shows that after combining the internal logic parameters of the signal controller with the multimodal input of the vehicle trajectory, the LSTM model can control the phase duration MAE within 9 seconds in the prediction of the next 6 signal cycles [8]. The GRU model has fewer parameters and faster convergence, which makes it more capable of real-time inference in resource-constrained vehicle terminals [9].

3.2. Convolutional neural networks (CNNs) and spatial feature extraction

Although RNNs excel at handling temporal dependencies, they still have shortcomings in capturing spatial correlations between intersections. CNNs, by performing convolution operations on the traffic density matrix, can identify spatial distribution patterns in local road networks.

Researchers have proposed a CNN-LSTM hybrid architecture, which uses CNNs to extract spatial features from the road network topology and then incorporates them into an LSTM model for time step extrapolation [5]. This type of model has stronger generalization performance than a single temporal model when dealing with arterial coordination intersections with high spatial correlations, and can more effectively identify downstream green light extensions caused by large traffic flows from upstream intersections.

3.3. Graph neural networks (GNNs) and spatiotemporal topology modeling

Considering the natural graph structure properties of urban roads (intersections are nodes, road segments are edges), graph neural networks (GNNs) have emerged in the field of traffic prediction in recent years. Spatiotemporal graph convolutional networks (STGCNs) simulate the propagation of traffic flow in physical road networks by performing spatial convolution on graph structures [10].

DOI: 10.54254/2753-8818/2026.34028

In their study in 2025, Assolie et al. proposed an enhancement framework combining XGBoost and ST-GNN, which uses particle swarm optimization (PSO) to dynamically adjust signal timing parameters. Experiments showed that this method accurately predicted the countdown timer while successfully reducing the average delay of the road network by 29.6% [11]. The advantage of GNNs is that they can handle the correlation of the road network, enabling the model to not only predict individual intersections but also to collaboratively estimate the signal light linkage patterns of the entire area.

4. Comparative analysis of the transformer and its variants in long-term prediction

Although LSTM and GNN have solved the spatiotemporal modeling problem to some extent, they still have the disadvantages of insufficient long-range dependency capture and low parallel computing efficiency when facing long-term time series predictions of several hours or even days. Since 2023, the Transformer architecture and its variants optimized for traffic prediction have become the advanced approach in the field of traffic prediction [2].

4.1. A revolution in core mechanisms: self-attention and parallelization

Transformer achieves direct calculation of weights at any position in the sequence through self-attention, breaking the serial dependency of recursive models, which has a natural advantage in predicting traffic data with strong periodicity and long-term trends [4]. In response to the long input and multivariate characteristics of SPaT prediction, a number of excellent variant models have emerged in academia.

4.2. Comparative technical analysis of mainstream transformer variants

Table 2 provides an in-depth comparison of current mainstream models in terms of core innovations, prediction mechanisms, and performance in traffic scenarios:

Table 2. Comparison of mainstream Transformer variants

Model Architecture	Latest scientific research contributions in 2024-2025	Advantages of long-term time series forecasting	Limitations in traffic signal prediction
Informer	ProbSparse attention and self-attention distillation mechanisms were proposed [12]	The computational complexity was reduced from $O(L^2)$ to $O(L\log\mathrm{L})$, which greatly extended the input window [13].	When processing high-frequency sampled data, sampling strategies may lead to the loss of detailed features.
Autoformer	Introducing Series Decomposition and Autocorrelation Mechanisms [14]	It can capture trends based on periodic similarity rather than point-to-point similarity, making it particularly suitable for regular signal cycles [13].	Low sensitivity to non-periodic disturbances caused by adaptive control
Patch TST	Patch-based segmentation and channel independence modeling are adopted [15]	It preserves local temporal semantics and can effectively filter out instantaneous noise generated by GPS positioning [16].	By default, the interaction between different traffic flow directions (variables) is ignored.
iTransformer	An architecture flip was implemented, treating each intersection/variable as an independent token [17]	The multivariate correlation between traffic lights at different intersections was explored in depth, making it suitable for regional-level prediction [17].	Relatively weak for dynamic capture of extremely long input time series.

Table 2. (continued)

GATTF	The Geographic Aware Transformer introduces mutual information (MI) to measure geographic relationships [18]	Without the need for a predefined adjacency matrix, spatial linkages between intersections can be automatically discovered [18].	The computational overhead of mutual information during training is relatively large.

4.3. In-depth insights into the SOTA model in 2025

In the experiments performed in 2025, the FE-iTransformer (Feature-Enhanced iTransformer) showed the best results. This model uses a feature enhancement module (FEM) to extract global context vectors of the spatiotemporal dynamics, periodic patterns, and temporal context and uses these to enhance the sequence being input [17]. Tests run on an urban road network data set confirm that this phased "extraction-enhancing" strategy improves the prediction accuracy for a duration of up to 120 minutes, significantly.

On the other hand, to meet the real-time objectives, DTLight (Lightweight Decision Transformer) integrates offline reinforcement learning with the Transformer, and by way of an adapter module, online fine-tuning is reduced. This model, post pre-training, is shown to be capable of being adapted by requiring little online input to control change strategies at new intersections, and improves by 40.7% [19]. This shows that the Transformer has moved from being purely a predictive solution to being an intelligent agent that supports decision-making.

5. Architecture and commercialization of AI traffic light navigation system

The value of an AI-powered traffic light navigation system lies in transforming complex algorithms into intuitive speed guidance for users. A complete system typically comprises four core levels: perception, prediction, decision-making, and interaction.

5.1. System core technology architecture

The core technology architecture contains three main parts, which are Multi-source perception, Cloud AI prediction and Navigation guidance interaction.

The first one is Multi-source perception. The real-time status of the signal controller is directly obtained through the roadside unit (RSU), or the trajectory of hundreds of millions of floating cars is accessed through the API of Baidu, Gaode and other platforms. The trend in 2025 is to integrate ultra-high-definition cameras, LiDAR and environmental meteorological sensors to build a global perception field [6].

The second one is Cloud AI prediction. This is the brain of the system, running the aforementioned FE-iTransformer or PatchTST model. This layer is responsible for handling concurrent data from millions of intersections and outputting standardized SPaT messages.

The last one is Navigation guidance interaction: The mobile terminal App or the in-vehicle infotainment system (IVI) receives SPaT prediction data and, in combination with the vehicle's current position, acceleration and target path, calculates and displays the countdown and suggested speed.

5.2. Global typical application case analysis

The first one is the Baidu Maps Real-time SPaT Estimation. As a leading case in mainland China, the system utilizes 15 million floating car data points daily to cover 2 million intersections in mainland China. Its robust scheme of combining FFT with the KS test has successfully kept 75% of the prediction error within 5 seconds, and its commercial coverage is extremely wide [3]. The next one is the Google Green Light. Google uses Google Maps' massive driving trend dataset to provide optimization suggestions to urban traffic engineers through AI models. Initial deployment results show that the system has successfully reduced the number of stops at intersections by 30% and reduced greenhouse gas emissions by 10% [20]. The last one is the Connected Traffic Pilot by the University of Michigan (U-M). Researchers developed an algorithm that only requires 5% network vehicle penetration to run and achieved a 20%-30% reduction in the number of stops at 34 intersections in Birmingham, Michigan, through a purely software strategy, demonstrating extremely high economic benefits [21].

6. Key points and challenges in technical

Although the model's accuracy has reached the second level, there is still some distance to go before achieving perfect AI leadership, and the following challenges still need to be overcome.

The first one is End-to-end real-time performance and low latency. Currently, there is an obvious delay from data collection to App display in the industry. In highly dynamic scenarios, even a delay of 5 seconds may cause navigation suggestions to be out of phase with the actual situation. In the future, research will tend to strengthen edge computing capabilities so that key predictions can be completed at the intersection.

Next comes the Analysis of complex control logic. There are a large number of sensor-activated traffic lights in cities that contain complex triggering instructions (such as pedestrian sensing and bus priority). Existing data-driven models still need to further enhance their logical reasoning capabilities when dealing with these "long-tail scenarios".

Finally is the Model generalization and adaptability. Traffic control system manufacturers vary across different cities and their protocols differ. The key to large-scale deployment is to utilize transfer learning or foundation models to enable the prediction system to quickly adapt to new regional and road network structures [2].

7. Future outlook

From early historical mean statistics to spatiotemporal modeling based on RNNs and GNNs, and then to the Transformer long-term architecture, breakthroughs have been made in the prediction accuracy and prediction window depth of traffic light durations. AI traffic light navigation systems, by combining SPaT prediction with the GLOSA algorithm, are profoundly changing the way people travel.

Looking to the future, the field of traffic light duration prediction will present the following three major trends: First, large-scale spatiotemporal modeling, based on pre-trained time series basic models, can achieve plug-and-play prediction of hundreds of millions of intersections worldwide through Prompt Tuning technology; Second, end-to-end security reinforcement, the defense mechanism will evolve from a simple filter to an adversarial training architecture with self-awareness, ensuring the integrity of data in the autonomous driving environment; Finally, deep empowerment of green and low-carbon, the prediction information will be deeply coupled with the powertrain management of electric vehicles, and the carbon neutrality goal will be achieved by planning energy recovery strategies in advance [1]. Through the deep integration of deep learning, V2X communication and network security, intelligent transportation systems will surely move towards a new era of greater accuracy, safety and efficiency.

8. Conclusion

In conclusion, this report systematically traces the evolution of traffic signal countdown prediction from early historical statistics to state-of-the-art Transformer architectures. While early statistical methods and shallow machine learning models established the foundation for SPaT estimation, they often struggled to capture long-range dependencies and the high non-linearity of modern urban traffic. The subsequent shift toward deep learning—utilizing RNNs for temporal sequences and GNNs for spatial topology—significantly improved modeling accuracy by accounting for intersection correlations.

Recent advancements in Transformer variants, such as FE-iTransformer and PatchTST, have further pushed the boundaries of long-horizon forecasting through self-attention mechanisms and multi-variable interaction. Real-world deployments by platforms like Baidu and Google highlight the commercial viability of these models in reducing congestion and emissions. However, challenges regarding end-to-end latency and the generalization of complex, adaptive control logic remain. Moving forward, the field is trending toward spatiotemporal foundation models and robust adversarial defense frameworks. Ultimately, integrating high-precision SPaT predictions with green energy management will be a decisive step toward achieving smarter, more efficient, and carbon-neutral urban mobility.

References

[1] Zenith Urban Light. (2025). 2025 best traffic light innovations for smart cities. https://www.zenithurbanlight.com/blog/2025-best-traffic-light-innovations-for-smart-cities/
[2] Okamoto, Y. (2026). Transformers for time-series forecasting: A comprehensive survey through 2024. Nonlinear Theory and Its Applications, IEICE, 17(1), 260-278.
[3] Liao, M., Feng, Z., Fan, M., Xu, S., & Xiong, H. (2025). Traffic Signal Phase and Timing Estimation with Large-Scale Floating Car Data. arXiv preprint arXiv: 2507.14190.
[4] Transformer-based short-term traffic forecasting model considering traffic spatiotemporal correlation – Frontiers
[5] Ugirumurera, J., Severino, J., Bensen, E. A., Wang, Q., & Macfarlane, J. (2023). A Machine Learning Method for Predicting Traffic Signal Timing from Probe Vehicle Data. arXiv preprint arXiv: 2308.02370.
[6] We Build Value. (2025, September 3). New AI traffic lights take over American roads. https://www.webuildvalue.com/en/megatrends/ai-traffic-lights.html
[7] Eteifa, S. O. (2024). Traffic Signal Phase and Timing Prediction: A Machine Learning and Controller Logic Hybrid Approach.
[8] Islam, Z., Abdel-Aty, M., & Ugan, J. (2024). Signal phasing and timing prediction using connected vehicle data. Transportation research record, 2678(1), 662-673.
[9] Mohammadi, A., Ahmari, R., Hemmati, V., Owusu-Ambrose, F., Mahmoud, M. N., Kebria, P., ... & Saif, M. (2025). Gps spoofing attack detection in autonomous vehicles using adaptive dbscan. arXiv preprint arXiv: 2510.10766.
[10] Samieiyan, B., & Awasthi, A. (2026). A Custom Transformer-Based Framework for Joint Traffic Flow and Speed Prediction in Autonomous Driving Contexts. Future Transportation, 6(1), 15.
[11] Assolie, A. A., Khliefat, I., Alqatawna, A., AL-kharabsheh, E. M., & Kasassbeh, S. A. (2025). Machine learning-based prediction of traffic signal timing for optimized intersection management. Innovative Infrastructure Solutions, 10(9), 400.
[12] Caetano, R., Oliveira, J. M., & Ramos, P. (2025). Transformer-based models for probabilistic time series forecasting with explanatory variables. Mathematics, 13(5), 814.
[13] Xu, J. (2025). Transformer-based Time Series Forecasting.

DOI: 10.54254/2753-8818/2026.34028

[14] Mkrtchian, G., & Gorodnichev, M. (2025). Patch-Based Transformer–Graph Framework (PTSTG) for Traffic Forecasting in Transportation Systems. Applied Sciences, 15(19), 10468.
[15] Nie, Y., Nguyen, N. H., Sinthong, P., & Kalagnanam, J. (2022). A time series is worth 64 words: Long-term forecasting with transformers. arXiv preprint arXiv: 2211.14730. https: //doi.org/10.48550/arXiv.2211.14730
[16] Ma, X., Ge, S., Yang, F., Li, X., Chen, Y., Ma, M., ... & Liu, Z. (2025). TimeExpert: Boosting Long Time Series Forecasting with Temporal Mix of Experts. arXiv preprint arXiv: 2509.23145.
[17] Duan, Y., Zhang, Y., Wang, X., Xue, Y., Wang, Z., & Wu, D. (2026). Feature-enhanced iTransformer: A two-stage framework for high-accuracy long-horizon traffic flow forecasting. PLoS One, 21(1), e0340389.
[18] Kušić, K., Cahill, V., & Dusparic, I. (2026). Geographically-aware Transformer-based Traffic Forecasting for Urban Motorway Digital Twins. arXiv preprint arXiv: 2602.05983.
[19] Huang, X., Wu, D., & Boulet, B. (2026). Traffic signal control using lightweight transformers: An offline-to-online rl approach. IEEE Open Journal of Intelligent Transportation Systems.
[20] Google Research. (n.d.). Green Light. https: //sites.research.google/gr/greenlight/
[21] University of Michigan Transportation Research Institute. (2025, November 5). New traffic signal reduces unnecessary stops. UMTRI. https: //www.umtri.umich.edu/2025/11/05/new-traffic-signal-reduces-unnecessary-stops/