Comparative Study on Traffic Prediction Using Different Models

Zhuofan Jiao
Lanzhou Oriental Secondary School, Lanzhou, Gansu, China

Abstract. Traffic flow prediction (TFP) is a complex and critical field that is of great significance for urban planning, management, and resource allocation. This paper discusses the development history and optimization strategies of TFP models. This paper first introduces the importance of TFP and outlines the basic concepts and characteristics of traditional and modern TFP models. Subsequently, through comparative analysis, the prediction accuracy and applicability of the two models were discussed in depth. On this basis, the key factors affecting TFP accuracy are further analyzed, and the corresponding model optimization strategy is proposed. This paper proposes an improved method for fusion prediction by combining multiple data sources. Through these optimizations, the model is better able to respond to sudden traffic changes and improve the robustness and real-time prediction of the forecast. Finally, the research results are summarized and the future research directions are prospected. Through the systematic study of TFP models, this paper provides theoretical support and practical guidance for traffic management and planning, which has important academic value and application prospects.

1 Introduction

Transportation, as an important component of modern society, plays a crucial role in economic development, residents' quality of life, and urban planning due to its efficient operation. With the acceleration of urbanization and the continuous growth of transportation demand, problems such as traffic congestion and frequent accidents are becoming increasingly prominent. To effectively solve these problems and improve the operational efficiency and safety of the transportation system, traffic flow prediction (TFP) is particularly important.

The prediction of traffic flow (TF) is one of the important methods for gradually increasing traffic volume. It has the functions of rational allocation of police resources for traffic police, strengthening and improving traffic monitoring, and enhancing the level of safe travel. In addition, TFP can effectively help navigation achieve optimal path guidance, thereby reducing traffic congestion, saving time for police and travelers, conserving resources, and reducing pollution. It can not only provide a decision-making basis for traffic management departments, and allocate traffic resources reasonably, but also provide accurate travel information for travelers, and optimize travel routes and mode choices.

Corresponding author: outlook a23f77f739b91279@outlook.com

Accurate TFP is a key prerequisite for achieving traffic optimization management and planning.

In today's complex and ever-changing traffic environment, the application and performance of different models in TFP have become a research keynote.

Deep learning methods can now effectively establish high-dimensional time and space data models for urban TP. For example, Yin's CNN and multi-scale feature fusion urban TFP model [1]. Wang's GA, RBF, and improved Cao methods for air TFP, and Xu's gated convolution time and space TFP model [2, 3]. These models are of great help in predicting TF. However, with the rapid development of information technology and the continuous enrichment of data collection methods, more and more advanced models and algorithms are being introduced into the field of TP. These models have significant differences in theoretical basis, data processing methods, prediction accuracy, and other aspects.

The traffic characteristics and demands vary in different regions and traffic scenarios, which puts higher demands on the adaptability of TFP models. Therefore, conducting comparative research on predicting traffic using different models has important theoretical value and practical significance. By analyzing the advantages and disadvantages of various models in depth, it can better select and apply suitable models, improve the accuracy and reliability of TP, and provide strong support for traffic planning, management, and control.

In summary, this study aims to reveal the inherent laws and applicability of different TFP models through comparative analysis, provide useful references and inspirations for the development of the transportation field, integrate new and old prediction methods, promote continuous innovation and progress of TFP technology, and adapt to increasingly complex traffic demands and challenges.

2 Overview of traffic model prediction

2.1 Traditional TFP models

Traditional TFP models play an important role in the development of the transportation field. These models are based on long-term practice and research accumulation, providing a certain theoretical basis and practical methods for transportation planning and management.

In traditional TFP models, the four-stage method is commonly used. This method divides traffic demand forecasting into four stages: travel generation, travel distribution, mode division, and traffic allocation. The travel generation stage mainly predicts the travel generation and attraction of various transportation communities through analysis of factors such as population and socio-economic conditions. The travel distribution stage determines the travel exchange volume between each community based on the generated travel results. The stage of mode division considers the characteristics of different modes of transportation and people's travel preferences and allocates travel volume to different modes of transportation. In the transportation allocation stage, the travel volume of each transportation mode is allocated to specific road networks to evaluate the service level of transportation facilities [4].

High precision TFP plays an important role in traffic management and smart travel in large cities, and mining the dynamic time and space correlation of TF is the key to improving prediction accuracy To address the issue of insufficient consideration of the high similarity of TF at different time scales and the similarity of TF changes between non adjacent nodes in similar functional areas in existing research, a dynamic graph convolutional neural network considering time and space similarity is constructed [5]. In addition, regression analysis models are also an important component of traditional TFP models. It predicts indicators such as TF and speed by establishing linear or nonlinear relationships between traffic variables and related influencing factors. For example, by analyzing the relationship between historical traffic data and factors such as land use and population density, a regression equation can be constructed for prediction.

However, traditional TFP models also have certain limitations. They often assume relatively simple conditions, making it difficult to fully consider the complexity and uncertainty of the transportation system. For example, the four-stage method may not accurately reflect the actual situation when dealing with dynamic traffic changes and emergencies. Regression analysis models may be affected by data quality and model fit, leading to bias in prediction results.

Despite these shortcomings, traditional TFP models still have certain application value in specific scenarios and conditions. In situations where data is relatively limited and transportation systems are relatively stable, they can provide preliminary decision-making basis for transportation planning and management. At the same time, the research and application of traditional models have laid the foundation for the development of modern TFP models, prompting researchers to continuously explore and innovate to improve the accuracy and reliability of TP.

2.2 Modern TFP model

Modern TFP models have emerged in the context of continuous development and innovation in the field of transportation. With the rapid advancement of information technology and the increasing richness of data collection methods, these models have demonstrated stronger predictive capabilities and adaptability.

Modern TFP models make full use of advanced technologies such as big data and artificial intelligence. Among them, deep learning based models such as Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) perform well in processing time and space data. Deep learning methods can effectively establish high-dimensional time and space data models for urban TP. Among them, RNN was first applied in TFP, but RNN has problems such as gradient explosion and gradient disappearance in dealing with long-term dependency problems [6]. CNN can effectively extract spatial features from traffic data, such as the topological structure of road networks and the distribution patterns of TF. RNNs and their variants, such as Long Short Term Memory Networks (LSTM) and Gated Recurrent Units (GRU), excel at capturing long-term dependencies in time series and are of great significance for predicting dynamic changes in TF. In addition, ensemble learning models also play an important role in modern TP. Random Forest and Gradient Boosting Decision Tree methods can effectively reduce prediction errors, and improve model stability and generalization ability by combining multiple weak learners.

There is also a type of model based on Graph Neural Network, which views the transportation network as a graph structure, with nodes representing road intersections or segments and edges representing their connection relationships. By conducting message passing and feature learning on the graph, it is possible to better capture the complex topology and dynamic characteristics of the transportation system.

Modern TFP models also focus on the fusion of multi-source data. In addition to traditional TF data, it also integrates external factors such as weather information, holiday arrangements, and major events. By comprehensively analyzing these multidimensional data, the model can gain a more comprehensive understanding of the operating rules of the transportation system, thereby providing more accurate prediction results.

However, modern TFP models also face some challenges in practical applications. For example, data quality and missing data issues may affect the training effectiveness of the model; The complexity of the model leads to high computational costs, making it difficult to respond quickly in real-time scenarios; The adaptability of the model still needs to be improved for newly emerging transportation modes and special situations.

In summary, modern TFP models, with advanced technology and innovative methods, have brought new opportunities and breakthroughs to the prediction work in the field of transportation. But at the same time, continuous improvement and perfection are also needed to address various issues in practical applications and better serve transportation planning and management.

3 Comparative analysis of models

3.1 Comparison of prediction accuracy

Traditional TFP models can to some extent capture the basic trends of TF. However, they are often limited by data quality and model complexity and have poor adaptability to complex traffic conditions and emergencies. Modern TFP models, with their powerful data analysis and pattern recognition capabilities, have significantly improved prediction accuracy. For example, neural network models can handle a large number of nonlinear relationships to more accurately predict changes in TF.

To objectively compare the prediction accuracy of different models, it is necessary to establish a comprehensive and scientific evaluation index system. Common indicators include mean square error, mean absolute error, coefficient of determination, etc. Through these indicators, the performance of different models in different scenarios can be quantitatively evaluated.

In practical applications, the prediction accuracy of different models varies in different types of traffic scenarios. For high-density traffic areas in urban centers, due to the complex changes in TF and the influence of multiple factors, certain complex modern models may perform better. In areas where TF is relatively stable and changes are more obvious, traditional models may also be able to provide more accurate predictions.

In addition, the quality and characteristics of data also have a significant impact on the prediction accuracy of the model. Rich, accurate, and representative data can significantly enhance the predictive ability of the model. Meanwhile, the parameter adjustment and training methods of the model will also have an impact on the prediction accuracy.

3.2 Model applicability analysis

In the field of transportation, different TFP models have their characteristics and applicable scenarios. Model applicability analysis is crucial for selecting appropriate prediction tools to achieve accurate and effective traffic planning and management.

Traditional TFP models still have certain applicability under specific conditions. For example, statistical models based on historical data may perform well in areas with stable data and relatively fixed traffic patterns. However, when traffic conditions are affected by unexpected factors such as large-scale events, road construction, etc., their predictive ability may be limited.

Modern TFP models have shown advantages in handling complex and dynamic traffic data. They can capture nonlinear relationships and space-time features, but require high data quality and quantity, and may have shortcomings in interpretability.

In the central areas of cities, where TF is high and changes frequently, modern models such as neural network models may be more suitable as they can quickly adapt to new traffic patterns and changes. In some small and medium-sized cities or suburbs, the transportation mode is relatively simple, and traditional regression models may be sufficient to meet the predicted demand [7].

For short-term traffic forecasting, real-time requirements are high, and some online learning-based models can quickly respond to the latest data and provide timely prediction results. For long-term planning, more macro factors need to be considered, such as urban development planning, policy changes, etc. At this time, a mixed model that comprehensively considers multiple factors may be more appropriate.

In short, the applicability of the model is not absolute but requires comprehensive consideration based on various factors such as specific traffic scenarios, data characteristics, predicted demand, and time ranges. Only by selecting the most suitable model can accurate and reliable decision support be provided for transportation planning and management, thereby improving the operational efficiency and service quality of the transportation system.

4 Influencing factors and model optimization

4.1 Factors affecting traffic forecasting

TFP is an important research direction in the field of transportation, and accurate TFP is of crucial importance for traffic planning, management, and operation. Before exploring optimization strategies for TFP models, it is necessary to conduct in-depth analysis of various factors that affect TP.

Data quality is a key factor. The accuracy, completeness, and timeliness of traffic data directly affect the prediction results. Inaccurate or incomplete data may lead to a bias in the model's understanding of traffic conditions, thereby reducing prediction accuracy. For example, sensor malfunctions, data collection errors, or delayed data updates can all cause deviations in the input model data.

The complexity and dynamism of transportation systems cannot be ignored. TF is influenced by various factors, such as holidays, weather conditions, and special events. The changes in these factors have uncertainty and randomness, which increases the difficulty of prediction. For example, adverse weather conditions may lead to a decrease in road capacity, thereby altering TF patterns; Large-scale events or unexpected events may lead to a sudden increase in transportation demand in local areas.

The assumptions and parameter settings of the model can also have an impact on the prediction. Different models are based on different assumptions and theoretical frameworks, and their ability to describe and explain traffic phenomena varies [8]. If the selection and adjustment of parameters are not reasonable, it may lead to the model being unable to accurately capture the characteristics of the transportation system.

Geographic and regional characteristics are also important influencing factors. The transportation modes and characteristics in different regions may vary significantly, with different transportation needs and behaviors between urban and rural areas, and developed and underdeveloped areas. Therefore, when making traffic forecasts, it is necessary to fully consider factors such as the local geographical environment, population distribution, and economic development level. Technological development and innovation are also constantly influencing traffic forecasting. The emergence of new data sources and technological means, such as satellite positioning data, social media data, etc., provides more information and possibilities for TP. However, how to effectively integrate and utilize these new data is also a problem that needs to be solved.

Overall, there are numerous and complex factors that affect traffic forecasting, including data quality, complexity of transportation systems, model assumptions and parameter settings, geographical and regional characteristics, and technological development. In practical TFP work, it is necessary to comprehensively consider these factors to improve the accuracy and reliability of predictions and provide strong support for traffic decision-making.

4.2 Model optimization strategy

In the field of TP, the optimization strategy of the model is crucial. To improve the accuracy and reliability of traffic forecasting, it is necessary to comprehensively consider multiple factors and take effective optimization measures.

The improvement of data quality is the foundation of optimizing models. Ensure that the collected traffic data is accurate, complete, and timely, providing more reliable input for the model. By adopting advanced data collection technologies such as sensor networks, satellite positioning systems, etc., richer and more accurate data can be obtained. At the same time, preprocessing of the data, including cleaning, denoising, normalization, and other operations, is carried out to reduce the impact of data errors on the model.

The improvement of the model structure is key. For traditional TFP models, the idea of deep learning can be introduced to increase the complexity and expressive power of the model. For example, in time series-based models, combining convolutional neural networks or recurrent neural networks can better capture spatiotemporal features in traffic data. For modern TFP models, the network structure can be further optimized and parameter settings can be adjusted to improve the model's generalization ability and robustness.

The optimization of algorithms is also an aspect that cannot be ignored. Choosing appropriate optimization algorithms, such as stochastic gradient descent, Adagrad, Adadelta, etc., can accelerate the training speed and convergence performance of the model. At the same time, regularization techniques such as L1 and L2 regularization are used to prevent overfitting of the model and improve its stability [9]. The advantage of integrating multiple models is also an effective optimization strategy. Different types of TFP models can be combined, such as combining statistical-based models with machine learning-based models, to fully leverage their respective advantages and improve prediction performance. At the same time, consider introducing multiple sources of data, such as weather data, holiday information, etc., to enrich the input of the model and improve its adaptability to complex traffic conditions.

Real-time updates and dynamic adjustments of the model are also necessary. Traffic conditions are constantly changing, and models need to be updated and adjusted in real time based on new data and actual conditions to maintain good predictive performance. By using online learning and incremental learning methods, the model can quickly adapt to changes in the traffic environment.

In summary, strategies such as improving data quality, improving model structure, optimizing algorithms, integrating multiple models, and real-time updating and adjusting can effectively optimize TFP models, providing more accurate and reliable decision support for traffic management and planning.

5 Future prospects

In the field of TP, although current research has achieved certain results, there is still vast space waiting for future exploration. Future research directions can be explored from multiple levels.

In terms of data collection and processing, with the continuous advancement of technology, more high-precision and multi-dimensional data will be available for acquisition. For example, real-time satellite image data, vehicle sensor data, and individual travel behavior data. How to effectively integrate and utilize these massive and complex data to improve the accuracy and reliability of predictions will be an important research direction.

Moreover, the integration and innovation of models are also crucial. Traditional and modern models each have their own advantages, and in the future, it can explore the organic combination of them to form more powerful hybrid models. At the same time, new algorithms and technologies such as reinforcement learning in deep learning and generative adversarial networks are introduced to develop more adaptive and flexible TFP models [10].

Considering the complexity and dynamism of the transportation system, future research should pay more attention to the real-time updating and adaptive capabilities of the model. The traffic situation may change at any time, and the model needs to be able to quickly respond and adjust the prediction results to better serve traffic management and decision-making.

Cross disciplinary research collaboration will become increasingly important. Traffic forecasting not only involves engineering technology, but is also closely related to fields such as sociology and economics. Through interdisciplinary communication and collaboration, a more comprehensive understanding of traffic behavior and demand can be achieved, providing richer ideas for model improvement. At the application level, future research should focus on deeply integrating TFP models with intelligent transportation systems. For example, real-time optimization of traffic signals, intelligent recommendation of travel routes, and early warning of traffic congestion can be achieved to improve the operational efficiency and service quality of the entire transportation system.

In conclusion, the evaluation and validation methods for TFP models also need to be further improved. Establish a more scientific and comprehensive evaluation index system to accurately measure the performance and effectiveness of models, providing a strong basis for model selection and improvement. In short, future research on TFP is full of challenges and opportunities, requiring researchers to constantly innovate and explore to promote the sustainable development of this field and provide more effective solutions to increasingly complex traffic problems.

6 Conclusion

In this comparative study on predicting traffic using different models, this article has drawn a series of important conclusions. From the perspective of prediction accuracy, modern TFP models have demonstrated relatively high accuracy in handling complex traffic conditions and large amounts of data. This is due to its advanced algorithms and ability to fuse multi-source data. However, traditional TFP models can still provide prediction results with certain reference values in specific simple scenarios and limited data conditions. On the other hand, in terms of model applicability, modern TFP models are more suitable for large cities and areas with frequent changes in TF, and can quickly adapt to new traffic patterns and emergencies. Traditional TFP models have shown good applicability and stability in small cities or areas with relatively stable traffic patterns.

In addition, this article also found that data quality and quantity have a significant impact on the prediction results. High-quality, large-scale, and diverse data can significantly improve the predictive accuracy of models. Meanwhile, the parameter settings and training methods of the model also directly affect its performance. Overall, different TFP models have their own advantages and limitations. In practical applications, appropriate prediction models should be selected based on specific traffic scenarios and demands. For complex and ever-changing traffic conditions, modern TFP models should be given priority consideration; For relatively simple and stable traffic situations, traditional TFP models may be a more cost-effective choice.

References

1. Q. Yin, F. Ding, Y. Zhu, J. Li, Y. Sha, Urban traffic flow prediction model based on CNN and multi-scale feature fusion. Comput. Technol. Dev. **10**, 175-181 (2022).
2. L. Wang, Y. Zhao, Air traffic flow prediction method based on GA, RBF, and improved Cao method. Traffic Inf. Saf. 1, 115-123 (2023).
3. L. Xu, X. Fu, H. Li, Spatiotemporal traffic flow prediction model based on gated convolution. Comput. Appl. 43, 2760-2765 (2023).
4. A. Pan, J. Lai, Research on traffic accident prediction model based on improved grey prediction model. Comput. Age 33-38 (2022).
5. Y. Bengio, P. Y. Simard, P. Frasconi, et al., Learning long-term dependencies with gradient descent is difficult. IEEE Trans. Neural Netw. 5, 157-166 (1994).
6. L. Hu, R. Zhang, X. Zhao, et al., Research on the identification model of motor vehicle conflict risk areas at road roundabouts. Transp. Syst. Eng. Inf. 1-12 (2022).
7. D. Yang, J. Liu, B. Zheng, et al., Vehicle lane changing trajectory prediction model based on time-domain convolutional network and attention mechanism. Transp. Syst. Eng. Inf. 1-14 (2022).
8. S. Zhu, B. Ye, W. Wu, Review and simulation research on short-term traffic flow prediction methods based on deep learning. Softw. Guide 23, 182-193 (2024).
9. L. Zhong, A. Zhang, X. Tian, Traffic flow prediction model based on periodic graph convolution and multi-head attention GRU combination. Comput. Appl. Res. 1-7 (2022).
10. J. Chen, G. Yang, Z. Zhang, et al., Urban regional traffic flow prediction based on spatiotemporal multi-view attention residual network. J. Syst. Simul. 1-16 (2022).