Predicting traffic flow during peak commuting hours

Z. Shen (✉ shenzhou.research@gmail.com)

Research Article

Keywords: Traffic flow, peak commuting hour

Posted Date: September 7th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-877585/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Predicting traffic flow during peak commuting hours

Zhou Shen*

Abstract

Traffic congestion during the peak commuting period is becoming more and more serious in many cities in China. Accurate prediction of traffic flow on urban roads during the peak commuting period is one of the key fundamental problems to reduce traffic congestion and build intelligent transportation. This paper proposes a low-cost and high-efficiency method to address the challenges of the existing ORIGIN-DESTINATION survey method, which is costly, inefficient and has limited accuracy of results. By using the city-wide traffic, social security, insurance, population and other digital city data, we group the residents' travel modes into four types: walking, public transportation, rail transit and self-driving based on the service radius and travel distance of transportation, and analyze and count the number of private car trips and rail transit passenger flow by using the rail transit priority principle and the shortest path algorithm, and predict the commuting peak period based on this method. The road traffic flow during the peak commuting period is also predicted. The validity and reliability of the method are demonstrated by using typical data of Wuhan city as an example.

1 Introduction

China's new urbanization faces many "urban diseases", such as traffic congestion and air pollution, especially in many large cities, traffic congestion during the peak commuting period is particularly prominent [1]. In China's large cities, such as Beijing, the average speed of the road network during the morning rush hour is 26.0 km/h, of which the average speed of expressways is 35.5 km/h and the average speed of main roads is 23.3 km/h; in Wuhan, the traffic flow of the Yangtze River Second Bridge has grown rapidly, from less than 60,000 vehicles per day in 2000 to more than 130,000 vehicles per day in 2008, with a peak of The average traffic volume per lane was 23,000 vehicles [2].

In view of the worsening traffic congestion in cities, accurate, reasonable and scientific prediction of traffic flow during the peak commuting period, improvement of related infrastructure, and inducement of people to optimize their

*The authors declare no competing interests.

travel patterns are the basic ways to alleviate traffic congestion during commuting [3]. Digital city (Cyber City) creates conditions for urban planning, intelligent transportation, grid-based management and services, location-based services, and urban security emergency response, and is an important tool for harmonious urban development in the information age [4]. In this paper, we use the data accumulated from the creation of digital city to predict the traffic flow during the peak period by adopting relevant algorithms to provide a new solution to alleviate urban road congestion.

2 Status of traffic forecasting

At present, domestic and foreign traffic flow forecasting is mainly about short-term traffic flow forecasting, short-term traffic flow forecasting methods are divided into qualitative and quantitative two categories, qualitative forecasting methods are related analogy method and Delphi method [5, 6]. Quantitative forecasting methods are roughly causal analysis forecasting, trend analysis forecasting and intelligent model forecasting three kinds. At present, the main applications for short-term traffic forecasting are historical mean method; regression analysis algorithm; neural network method; Kalman filter method [7, 8]. However, all these forecasting methods have their limitations: the historical mean method is simple, but it cannot study the dynamic and nonlinear characteristics of traffic flow, and it is difficult to solve the unconventional and unexpected traffic conditions. Although the regression analysis method can be applied to the nonlinear dynamic system of traffic flow, the use of this method requires a large amount of analytical data, the prediction speed is slow, and the determination of parameters is relatively cumbersome, based solely on the predetermined regression equation, the prediction from the measured factors affecting the traffic flow, only applicable to a specific section of a specific range of traffic, and can not be corrected for errors in a timely manner [9–11]. The Kalman filter method is a linear model, so the model performance deteriorates when predicting nonlinear, uncertain traffic flow, and the weights have to be adjusted in each calculation, therefore, the calculation is too large, and the results, predicted output values are sometimes delayed by several time periods, and the prediction results have a certain time lag problem. In view of the nonlinearity and complexity of traffic flow itself, knowledge-based model-free intelligent prediction algorithms are applied to the field of short-time traffic flow prediction, and because of the advantages of neural network model and its combination model based on neural network with strong learning ability [12, 13], strong robustness and fault tolerance, parallel structure and nonlinear mapping [14, 15], it is more suitable for the field of short-time traffic flow prediction, therefore, it is currently in the field of short-time traffic flow prediction. Therefore, it is widely used in the field of short-time traffic prediction. However, the neural network algorithm is a "black box" learning mode, which requires a large amount of raw data in the training process, and insufficient data will lead to bad prediction results. Therefore, the generalization ability is poor. At the same time, the training speed is slow and

it is easy to fall into the local minima.

However, the above methods are based on existing traffic flow data, and they are not feasible for situations where there is no traffic flow data. This calls for other dimensions, especially technological innovation, to be considered for accurate and fast traffic forecasting. Traffic flow prediction is no longer limited to the study of traffic flow data on the road, but also provides more ideas for traffic flow prediction from multi-source heterogeneous data such as population data (residence, workplace, private car) and public transportation data (bus stops, bus routes). In this paper, we make full use of the data gathered from the construction of digital city to analyze the potential correlation between multi-source heterogeneous data, so as to predict the traffic flow during the peak commuting period.

3 Methodology for predicting traffic flow during peak commuting hours

Traffic volume is the number of traffic entities passing through a road section in a selected period of time, referred to as traffic volume. According to the type of traffic, traffic volume is divided into motor vehicle traffic, non-motorized traffic and pedestrian traffic. Commuter trips account for the majority of urban traffic, especially in the "morning peak and evening peak". For the commuter peak period, there are many vehicles moving, but most of them are buses and private cars which travel to and from work. Therefore, in this paper, only buses and private cars are considered for traffic flow prediction.

3.1 Commuting distance calculation method

The commuting distance in this paper is calculated by the shortest path algorithm. The shortest path problem is not only a typical problem in theoretical studies such as graph theory and algorithm design, but also an important problem in many applied industries. The shortest path algorithm aims to find the shortest path between two nodes in a graph, and has a wide range of applications in many real-world industries. For example, in trip planning, logistics planning, traffic simulation, car navigation, biomedicine, social networks, and location-based services. In this paper, shortest path refers to the shortest distance in the general geographic sense. Based on the road network data and Arc GIS network analysis secondary development module, the optimized correlation data obtained from data pre-processing is loaded into the shortest path algorithm to calculate the commuting distance of residents.

3.2 Rail Transit Preference Method

The role and effectiveness of public transportation in relieving urban traffic congestion has been fully tested and has been more widely implemented as an

urban transportation development strategy. The urban transportation development strategy of alleviating traffic congestion and achieving sustainable development of cities and transportation by giving priority to public transportation is becoming a trend. As an important core element of the priority development of urban public transportation, rail transit is of great significance to alleviate traffic pressure and facilitate the travel of citizens. Therefore, in this paper, we adopt the principle of rail transit priority to predict the traffic flow during peak periods. According to the relevant literature, the reasonable attraction range of rail transit stations can be divided into reasonable walking range and reasonable traffic range.

1) Reasonable walking area: The acceptable area for passengers to walk to the rail transit station is called reasonable walking area. The reasonable walking area can be divided into the reasonable walking area in the central area and the reasonable walking area in the peripheral area of the city. The reasonable walking range in the central area is about 500–600 m, and the reasonable walking range in the peripheral area of the city is about 800–1,000 m.

2) Reasonable traffic range: The acceptable area for passengers to go to the rail transit station by non-rapid transit is called the reasonable traffic range. For the reasonable traffic range of the rail transit station, according to the relevant literature, the reasonable attraction traffic range of the rail transit station is about 3,000 m. Therefore, the attraction range of the rail transit station is not an absolute range, it has a flexible boundary. Thus, it can be seen that the attraction range of a rail transit station is lower as the origin and destination points are farther away from the station.

For the passengers who arrive or leave the rail transit station on foot, the reasonable walking distance D_{min} of the rail transit station is calculated by the direct rail transit algorithm; for the passengers who arrive or leave the rail transit station by conventional transportation, the reasonable transportation distance D_{max} of the rail transit station is calculated by the bus-rail transit interchange algorithm.

3.2.1 Direct rail transit algorithm under reasonable walking range

Assuming a reasonable walking area radius D_{min} , we make an outer envelope box with a radius of $2D_{min}$ for the rail station, traverse the box of each station, and find the residence Z and workplace point J that fall within these boxes. If

$$(Z, J) \in \text{Box}(2D_{min}),$$

it means that the resident's place of residence and workplace fall within the outer box of the station, he/she will prefer to travel by direct rail transit (including inter-rail transit transfer) without other transportation connections. If

$$Z \in \text{Box}(2D_{min}) \text{ and } J \notin \text{Box}(2D_{min})$$

or

$$J \in \text{Box}(2D_{min}) \text{ and } Z \notin \text{Box}(2D_{min})$$

or

$$(Z, J) \notin \text{Box}(2D_{min}),$$

it means that the resident's residence and workplace are neither within these buffer zones nor one of them, the resident will not travel by direct rail transit (including inter-rail transfers). The number of passengers travelling on rail is calculated by traversing each station according to the above steps.

3.2.2 Bus-rail interchange algorithm under reasonable traffic range

Assuming that the radius of the reasonable traffic zone is D_{max} , we make an outer envelope box for the rail station with radius $2D_{max}$, and traverse the box of each station to find the residence Z and workplace point J that fall within these boxes. If

$$(Z, J) \in \text{Box}(2D_{max}),$$

it means that the residence and workplace of this resident fall within the buffer zone, then rail travel will be preferred, but regular public transportation will be required for the connection. It should be noted that the connection may be to and from the rail station, or just to and from the rail station. For the above three cases that require regular bus connections, it is necessary to consider whether there is a regular bus connection to the arrival and departure rail transit stations, whether there is a regular bus connection to the arrival rail transit station, and whether there is a regular bus connection to the departure rail transit station. To determine whether there is a regular bus connection from the residence to the rail transit station, the basic principle is: first, determine whether there is a regular bus stop near the rail transit station, if there is a stop A, then determine whether there is a regular bus stop near the residence, if there is a stop B, see whether the same regular bus line is between stop A and stop B. If so, you can take the regular bus line. If so, you can take regular transportation to the rail station, which means that the person can choose rail transportation as a priority. The basic principle of determining whether there is a regular bus connection from the rail station to the workplace is the same as whether there is a regular bus connection from the place of residence to the rail station. By following these steps, we can calculate the number of passengers who travel by rail.

4 Test and Analysis

In order to verify the credibility of the predicted traffic flow method proposed in this paper, experiments were conducted using typical data of Wuhan city. According to the data, the majority of passengers who arrive at or leave the rail station by walking and the majority of passengers who arrive at or leave the rail station by regular bus in Wuhan are walking passengers, of which 70% are walking passengers and 13% are regular bus passengers. In addition, 85% of the passengers reach the rail station within 20 min. in terms of the time spent on the

transfer between the travel end and the destination end. Taking into account the actual situation in Wuhan, it can be expected that after the completion of all the lines of rail transit, 66% of the population in Wuhan will be able to reach a rail transit station within a walking distance of about 600 m, which will achieve the real goal of "60 min through the city and 30 min to the central city". Therefore, the direct rail transit algorithm is used for passengers who walk to the rail transit station, and the bus-rail transit interchange algorithm is used for passengers who reach the rail transit station by regular bus.

In this paper, the area around the Yellow Crane Tower in Wuhan City was chosen for the experiment. The Yellow Crane Tower is located at the Snake Hill in Wuchang, west of the Simenghuang Bridge, south of the Horse-Reading Stadium and the Red Building, and north of the Beijing-Guangzhou railroad line, with an average altitude of 85 m. The Yellow Crane Tower is a landmark in Wuhan. The data processed are: population data, insurance data, social security data, road data, and public transportation point and line data. Among them, population data, insurance data and social security data summarize the basic information of residents.

The simulation results show that, the traffic flow of main roads, secondary roads and branch roads is generally larger than that of internal roads and minor roads, which is in line with the actual situation. For example, Zhongshan Road around Huanghelou scenic spot is the main road of the city, connected with Wuhan City Railway Line 2, Wuchang Railway Station and Railway Line 4, and the predicted traffic flow is very large, which is in line with the actual situation.

5 Conclusion

To meet the major needs of urban planning and design, this paper proposes a prediction method based on population data to break through the conventional method of predicting traffic flow. Firstly, the data is imported into Arc GIS for data pre-processing to reduce redundancy and optimize data; secondly, the shortest path is calculated based on the shortest path secondary development function; based on the shortest path, the residents within the walking threshold are filtered out by combining the service range of transportation with the shortest path; then the eligible residents are filtered out according to the rail transit priority algorithm; and then the shortest path is obtained based on the information of private cars, Then, the shortest route algorithm is combined with the shortest route information to obtain the shortest route of these people; finally, the bus route data and travel schedule are combined to obtain the traffic flow of the specified road and rail traffic flow. Through the analysis and experimental validation, we break through the limitation of relying on historical traffic data and effectively extract potential information from the basic population data of Wuhan city, which provides a new research idea for predicting traffic flow.

References

- [1] X. Guan *et al.*, “Assessment on the urbanization strategy in china: Achievements, challenges and reflections,” *Habitat International*, vol. 71, pp. 97–109, 2018.
- [2] Y. Jia *et al.*, “Traffic speed prediction using deep learning method,” in *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2016, pp. 1217–1222.
- [3] H. Couclelis, “The construction of the digital city,” *Environment and Planning B: Planning and design*, vol. 31, no. 1, pp. 5–19, 2004.
- [4] Un-Habitat, *State of the World’s Cities 2008/9: Harmonious Cities*. Routledge, 2012.
- [5] J. J. Q. Yu *et al.*, “Autonomous vehicle logistic system: Joint routing and charging strategy,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 7, pp. 2175–2187, Jul. 2018.
- [6] J. Pill, “The delphi method: substance, context, a critique and an annotated bibliography,” *Socio-economic planning sciences*, vol. 5, no. 1, pp. 57–71, 1971.
- [7] E. I. Vlahogianni *et al.*, “Short-term traffic forecasting: Where we are and where we’re going,” *Transportation Research Part C: Emerging Technologies*, vol. 43, pp. 3–19, 2014.
- [8] Y. Liu *et al.*, “Privacy-preserving traffic flow prediction: a federated learning approach,” *IEEE Internet of Things Journal*, vol. 7, no. 8, pp. 7751–7763, Aug. 2020.
- [9] C. Moorthy *et al.*, “Short term traffic forecasting using time series methods,” *Transportation planning and technology*, vol. 12, no. 1, pp. 45–56, 1988.
- [10] J. J. Q. Yu *et al.*, “A social spider algorithm for global optimization,” *Applied Soft Computing*, vol. 30, pp. 614–627, May 2015.
- [11] F. Guo *et al.*, “Predictor fusion for short-term traffic forecasting,” *Transportation research part C: emerging technologies*, vol. 92, pp. 90–100, 2018.
- [12] P. Vythoulkas, “Alternative approaches to short term traffic forecasting for use in driver information systems,” *Transportation and traffic theory*, vol. 12, pp. 485–506, 1993.
- [13] J. J. Q. Yu *et al.*, “Intelligent fault detection scheme for microgrids with wavelet-based deep neural networks,” *IEEE Transactions on Smart Grid*, vol. 10, no. 2, pp. 1694–1703, Mar. 2019.
- [14] Y. Zhang *et al.*, “A graph deep learning method for short-term traffic forecasting on large road networks,” *Computer-Aided Civil and Infrastructure Engineering*, vol. 34, no. 10, pp. 877–896, 2019.
- [15] J. J. Q. Yu *et al.*, “Intelligent time-adaptive transient stability assessment system,” *IEEE Transactions on Power Systems*, vol. 33, no. 1, pp. 1049–1058, Jan. 2018.
- [16] B. Yu *et al.*, “Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting,” *arXiv preprint arXiv:1709.04875*, 2017.
- [17] J. J. Q. Yu *et al.*, “Evolutionary artificial neural network based on chemical reaction optimization,” in *Proc. IEEE Congress of Evolutionary Computation*. New Orleans, LA, US: IEEE, 2011, pp. 2083–2090.
- [18] Y. Li *et al.*, “Diffusion convolutional recurrent neural network: Data-driven traffic forecasting,” *arXiv preprint arXiv:1707.01926*, 2017.
- [19] J. J. Q. Yu *et al.*, “A social spider algorithm for solving the non-convex economic load dispatch problem,” *Neurocomputing*, vol. 171, pp. 955–965, Jan. 2016.
- [20] —, “Online vehicle routing with neural combinatorial optimization and deep reinforcement learning,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 10, pp. 3806–3817, Oct. 2019.
- [21] C. Zhang *et al.*, “Spatial-temporal graph attention networks: a deep learning approach for traffic forecasting,” *IEEE Access*, vol. 7, pp. 166 246–166 256, 2019.

- [22] I. Lana *et al.*, “Road traffic forecasting: Recent advances and new challenges,” *IEEE Intelligent Transportation Systems Magazine*, vol. 10, no. 2, pp. 93–109, 2018.
- [23] J. J. Q. Yu *et al.*, “Delay aware power system synchrophasor recovery and prediction framework,” *IEEE Transactions on Smart Grid*, vol. 10, no. 4, pp. 3732–3742, Jul. 2019.
- [24] R. Yu *et al.*, “Deep learning: A generic approach for extreme condition traffic forecasting,” in *Proceedings of the 2017 SIAM international Conference on Data Mining*. SIAM, 2017, pp. 777–785.
- [25] J. J. Q. Yu *et al.*, “Delay aware intelligent transient stability assessment system,” *IEEE Access*, vol. 5, pp. 17 230–17 239, 2017.
- [26] J. J. Q. Yu, “Two-stage request scheduling for autonomous vehicle logistic system,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 5, pp. 1917–1929, May 2019.
- [27] C. Zhang *et al.*, “Long-term mobile traffic forecasting using deep spatio-temporal neural networks,” in *Proceedings of the Eighteenth ACM International Symposium on Mobile Ad Hoc Networking and Computing*, 2018, pp. 231–240.
- [28] J. J. Q. Yu *et al.*, “Optimal V2G scheduling of electric vehicles and unit commitment using chemical reaction optimization,” in *Proc. IEEE Congress of Evolutionary Computation*. Cancun, Mexico: IEEE, 2013, pp. 392–399.
- [29] —, “Double auction-based pricing mechanism for autonomous vehicle public transportation system,” *IEEE Transactions on Intelligent Vehicles*, vol. 3, no. 2, pp. 151–162, Jun. 2018.
- [30] Y. Sun *et al.*, “Chemical reaction optimization for the optimal power flow problem,” in *Proc. IEEE Congress of Evolutionary Computation*. Brisbane, Australia: IEEE, 2012, pp. 1–8.
- [31] Y. Liu *et al.*, “PPGAN: privacy-preserving generative adversarial network,” in *Proc. IEEE International Conference on Parallel and Distributed Systems*. Tianjin, China: IEEE, 2019, pp. 985–989.
- [32] A. Y. S. Lam *et al.*, “Power-controlled cognitive radio spectrum allocation with chemical reaction optimization,” *IEEE Transactions on Wireless Communications*, vol. 12, no. 7, pp. 3180–3190, Jul. 2013.
- [33] Y. Yan *et al.*, “Multi-objective design optimization of combined cooling, heating and power system for cruise ship application,” *Journal of Cleaner Production*, vol. 233, pp. 264–279, Oct. 2019.