ADVANCES IN TRANSPORTATION STUDIES
An International Journal

Guest Editor: Y. Wu

2024 Special Issue, Vol. 1

Contents

aracne

Advances in Transportation Studies an international Journal 2024 Special Issue, Vol. 1

ISBN 979-12-218-1230-5 – ISSN 1824-5463 – DOI 10.53136/97912218123051 – pag. 3-14

A layout planning method for urban agglomeration transportation lines from the perspective of carbon reduction policies

Y. Li[1] Q.L. Zhan[1*] F. Gao[2]

[1]*Hunan International Engineering Consulting Center Co., LTD, Changsha 410016, China*
[2]*Lyceum of the Philippines University-Batangas, Batangas City, The Philippines*
**corresponding author; email: 25333244@qq.com*

Abstract

To enhance the accessibility of transportation systems in urban agglomerations and mitigate carbon emissions from traffic, this paper introduces a comprehensive planning approach for transportation routes. Initially, it calculates the carbon emissions generated by urban mass transit, and set minimizing these emissions as a key objective. Subsequently, for road traffic within urban agglomerations, a layout model is established that aims to minimize both travel time and carbon emissions. By incorporating the Pareto optimal concept optimization algorithm, the study determines the optimal solution. In terms of rail transit, the paper establishes the optimal planning outcome by employing the probability density function of rail transit as a constraint, after determining the appropriate scale of the system. Experimental results demonstrate that this approach effectively controls the overall carbon emissions of the urban agglomeration transportation system, with the layout coordination coefficient peaking at 0.926.

Keywords - urban agglomeration transportation system, carbon reduction, transportation lines, layout planning

1. Introduction

With the acceleration of global urbanization, urban agglomerations, as an important form of urban development, have gradually become an important engine for economic development in various countries. In the formation and development process of urban agglomerations, transportation route layout planning plays an important role in the connection between cities and regional economic development [8, 14].

Urban agglomeration transportation route layout planning refers to the scientific and reasonable planning and layout of transportation lines within the urban agglomeration based on factors such as regional economic development, resident travel needs, and transportation conditions. This plays an important role in promoting transportation connections and economic development within the urban agglomeration.

Liu and Tong [10] designed a Dyna framework to improve the efficiency of traffic route planning and vehicle safety, enhancing the computational efficiency of reinforcement learning algorithms. Then, the Sarsa algorithm is introduced to optimize the Dyna framework, with constraints on convergence speed and number of vehicle collisions, to plan traffic lines. In Wang and He [15], in order to achieve the planning of urban road traffic, based on the analysis of the functional level and road network layout form of urban roads, guided by the concept of collaborative planning, the four stage method is used to layout the road lines, and the planning results are optimized and adjusted according to the importance of hub nodes and traffic location theory. In order to achieve the planning of urban transportation lines, Zhou et al. [20] analyzed the current operational characteristics and traffic flow characteristics of urban roads. Then, based on the peak section of traffic flow, the scale of traffic operation was determined. With the idea of developing outward along the central city, a "fast line+ordinary line" composite channel was planned, and the transportation coordination planning in different urban areas was optimized by combining corridor fast lines and ordinary lines.

However, in practical applications, it has been found that the traditional methods mentioned above do not fully consider the characteristics of traffic flow, nor do they fully consider the connections and transfers between different modes of transportation, which may lead to a mismatch between planned lines and actual needs, reducing the coordination of route layout. In addition, the current layout of transportation lines in urban agglomerations in China is mainly influenced by factors such as urban spatial structure, transportation demand, and transportation supply [5, 6]. However, traditional transportation route layout planning usually focuses on meeting transportation needs as the main objective, with less consideration given to the impact of carbon emissions on the environment. Therefore, how to reduce carbon emissions by optimizing transportation route layout planning while meeting the development needs of urban agglomerations is an urgent problem that needs to be solved [4]. Therefore, guided by carbon reduction policies, this study designed a layout planning method for urban agglomeration transportation lines.

2. Implementation of layout planning for urban agglomeration transportation lines based on carbon reduction policies

From the perspective of carbon reduction policies, the layout planning of transportation lines in urban agglomerations is crucial for reducing carbon emissions [16]. Based on the statistics of carbon emissions over a period of time, this study aims to minimize carbon emissions and conducts research from the perspectives of urban agglomeration highway lines and rail transit lines, which helps to explore methods for achieving optimal planning.

2.1. Design of the target function for carbon emissions from transportation in minimum urban agglomeration

This study sets up multiple monitoring points in the existing urban agglomeration transportation line system, and uses a multi-sensor array to collect vehicle carbon emission information. The structure of multi-sensor array for collecting vehicle carbon emission information is shown in Figure 1.

Due to the strong uncertainty of urban carbon emission monitoring data [9], the Kalman filter multi-source information fusion method is adopted to achieve the fusion of low-level, real-time, and dynamic multi-source information. Calculate the Euclidean distance between each sample in the collected information and the fusion center, using the formula:

$$d(x_k, o_l) = \sqrt{\int_{l=1}^{k} (x_k - o_l)^2} \tag{1}$$

where x_k represents the k-th collected sample; o_l represents the l-th fusion center.

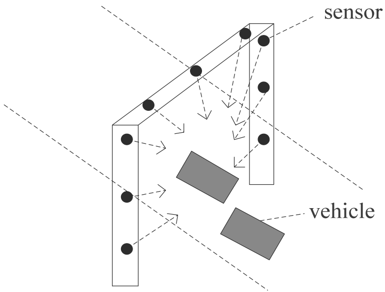

Fig. 1 - Multi sensor array collects vehicle carbon emission information

Firstly, divide it into the class with the closest Euclidean distance, and then take the mean of the sample points under each class to obtain a new set of multi-source datasets. Repeat this process until the standard deviation function reaches its minimum value. Draw a circle with this distance as the radius and integrate all collected carbon emission information into it [7, 19].

The carbon emissions of urban transportation vehicles refer to the carbon dioxide emissions generated by fuel combustion, which are determined by the fuel carbon dioxide emission factor and fuel activity level. The fuel carbon dioxide emission factor is calculated using the formula:

$$A_i = C_i \times O_i \times \frac{44}{12} \tag{2}$$

where C_i represents the carbon content per heat of the i-th type of transportation fuel; O_i represents the carbon oxidation rate of the i-th type of automotive fuel; $\frac{44}{12}$ represents the relative molecular weight ratio of carbon dioxide to carbon.

Then calculate the fuel activity level using the following formula:

$$\varphi_i = Q_i \times W_i \tag{3}$$

where Q_i represents the heat generated by the fuel; W_i represents fuel consumption. Based on the above two parameters, calculate the carbon emissions of urban mass transit using the formula:

$$C = \sum A_i \times \varphi_i \tag{4}$$

Therefore, the objective function for designing the minimum carbon emissions from urban agglomeration transportation is as follows:

$$\min\ C = \frac{d(x_k, o_l) \times C}{k \times l} \times \varphi_i \tag{5}$$

2.2. Layout planning method for urban agglomeration transportation lines

With the minimum carbon emissions of urban agglomeration transportation as the goal, carry out the layout planning of urban agglomeration highway lines and rail transit lines.

2.2.1. Layout planning and design of urban agglomeration highway transit lines

Firstly, obtain the spatial shape and structural characteristics of the urban agglomeration public route network, with the minimum travel time and minimum carbon emissions $\min\ C$ as objectives, establish an integrated layout optimization combination objective model, and then apply the Pareto optimal concept optimization algorithm to solve the model, obtaining the optimal layout plan. The specific steps are as follows:

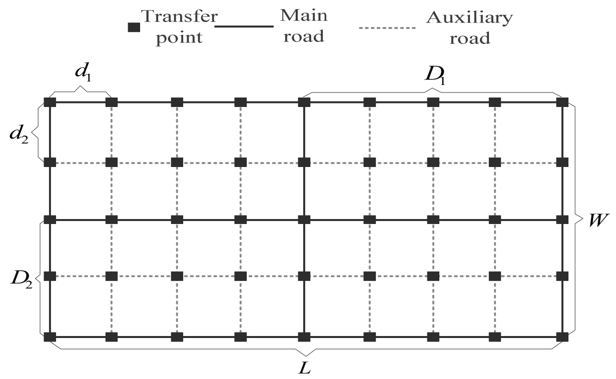

Fig. 2 - Urban highway network and characteristic parameters

Step 1: Extract the features of the urban agglomeration public route network. Simplify the urban highway network in the form of a rectangular grid [3, 18]. The urban highway network and characteristic parameters are shown in Figure 2.

In Figure 2, L represents the length of the transportation community, W represents the width of the transportation community, d_1 and d_2 represent the horizontal and vertical station spacing, respectively, and D_1 and D_2 represent the road spacing.

Starting from Figure 2, obtain several key geometric parameters, and then calculate the proportion of horizontal and vertical lines in the entire transportation area network as follows:

$$\begin{cases} a = \frac{D_2}{D_1+D_2} \\ b = \frac{D_1}{D_1+D_2} \end{cases} \tag{6}$$

where a represents the proportion of horizontal lines in the highway network of the transportation area; b represents the proportion of vertical lines.

Step 2: Build an optimized layout objective model. Starting from the perspectives of decision-makers and travelers, propose three layout optimization objectives and develop the most reasonable planning scheme through multi-level planning.

The first goal is to minimize the travel time for travelers. Simply put, it means that most travelers can find the path with the lowest travel cost [2]. To this end, a stochastic user balance model is introduced to construct a layout optimization mathematical model.

The second goal is to reduce construction expenses. The layout expenditure consists of two parts: the total travel expenses of travelers and the daily construction expenses converted within the service life.

The third goal is the minimum carbon emissions $\min\ F$.

After integrating three optimization objectives using a combination objective model, the following objective model is obtained:

$$\begin{cases} \min\ T = -\int_{r_1,r_2}^{n} \lambda_{r_1 r_2} \times \left(\min_k C_k^{r_1 r_2}\right) + \sum_u^U x_u t_u - \sum_u^U t_u \\ \min\ P = \sum_u^U x_u t_u + \tau \sum_j^J \sigma_i p_i \\ \min\ C = \frac{d(x_k, o_l) \times C}{k \times l} \times \varphi_i \end{cases} \tag{7}$$

where T represents the traveler's mental traffic time; x represents the traffic flow; n represents the number of stations within the transportation area; r_1 and r_2 represent any two sites; λ represents the travel volume; k represents the path between two sites; U represents the total number of road sections; t represents traffic time; P represents the overall cost expenditure; τ represents the depreciation rate of highway investment per hour; J represents the total number of interchanges; σ represents the investment decision variable; p represents the expected investment amount.

Step 3: Define the constraints of the layout optimization model. Among them, the constraint conditions corresponding to $\min\ T$ are:

$$\begin{cases} C_k^{r_1 r_2} = \sum_u^U B_u \psi_{uk}^{r_1 r_2} \\ \psi_{uk}^{r_1 r_2} = \begin{cases} 1, u\ is\ between\ r_1\ and\ r_2 \\ 0, \text{others} \end{cases} \end{cases} \tag{8}$$

where B represents the time consumed by the traveler; ψ represents a relational variable.

The constraint conditions corresponding to $\min\ P$ are:

$$\begin{cases} \sum_j^J \sigma_i p_i \le A \\ \varphi_u \le x_u \le y_u \end{cases} \tag{9}$$

where A represents the upper limit of the budget investment amount; φ_u, y_u represents the minimum and maximum traffic flow of the road section, respectively.

Step 4: Generate the optimal integrated layout optimization plan for overpasses. When solving the optimization objective model of highway route layout by combining multiple constraints, the concept of Pareto optimality is introduced, and a genetic algorithm is used to conduct a global search in the solution space to obtain the sub objective function values of all dominant decision vectors. After comparison, the final optimization solution is obtained. The Pareto optimal solution and decision diagram are shown in Figure 3.

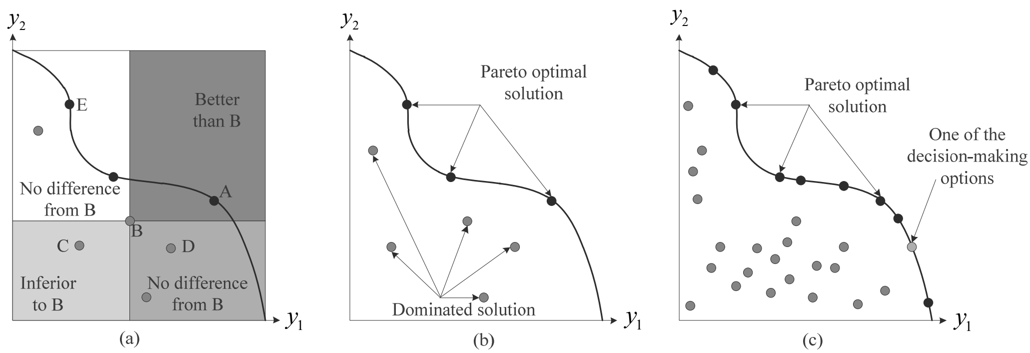

Fig. 3 - Pareto optimal solution and decision diagram

When executing the Pareto optimal solution search method, it needs to rely on genetic algorithms [11]. After genetic mutation, the elimination of individuals in the population needs to be determined based on the Hamming distance. For members who are excluded after mutation and the next generation members, gene values are analyzed separately. The Hamming distance between the two is calculated as follows:

$$\|z_\vartheta - z_\epsilon\| = \sqrt{\sum_1^\epsilon (z_{\vartheta\epsilon} - z_{\epsilon\epsilon})^2} \tag{10}$$

where z_ϑ represents the ϑ-th excluded member; z_ϵ represents the ϵ-th next-generation member individual; ϵ represents the number of gene values that make up an individual.

After deleting some individuals according to the Hamming distance, the remaining member individuals are sorted in descending fitness order to obtain the evolved genetic population [1]. Repeat the above purification process until the optimal solution that meets the accuracy requirements of the optimization variables is output. Reverse infer the layout optimization plan corresponding to this solution, and use it as the final result of the urban agglomeration highway route layout optimization.

2.2.2. Layout planning and design of urban agglomeration rail transit lines

Rail transit lines play an important role in urban agglomeration transportation, and their optimized layout is of positive significance in alleviating urban traffic pressure and promoting green travel [13]. The steps for planning the layout of urban rail transit lines are as follows:

Step 1: Obtain the variation characteristics of urban rail transit flow speed. In order to improve the economic efficiency of urban rail transit operation, a rail transit lane model is constructed using feature point tracking method as follows:

$$S^* = \frac{(i,j)+Y(i,j)}{D_{(n,j)}} + C_x \tag{11}$$

where (i,j) represents the row and column coordinates of the center of the rail transit network; $Y(i,j)$ represents the transfer point in the i-th row and j -th column of the rail transit network; C_x represents the differential sum of the line; $D_{(n,j)}$ represents the correlation gradient between urban rail transit lines and highway transportation lines.

When obtaining the characteristics of changes in the speed of rail transit flow, the Gaussian background model is combined to extract the information of running vehicles as follows:

$$A_y = \frac{Z_s+h_r}{f\times\delta(i,j)} \times \theta_{ab} \tag{12}$$

Among them, Z_s represents the frame with the highest disappearance frequency in the extraction of running vehicles; h_r represents the distance between the vehicle information acquisition camera and the ground; θ_{ab} represents the angle between two intersecting rail transit lines; f represents the frequency of occurrence of two intersecting rail transit lines; $\delta(i,j)$ represents the location where the vehicle disappeared. Based on A_y, obtain the driving state characteristics of the specified vehicle as follows:

$$S_z = \frac{S_L\times t_g}{N_d} + B(i,j) \tag{13}$$

where S_L represents the length of the rail transit line where the vehicle is located; t_g represents the time when the vehicle enters the transfer interval; N_d represents the number of vehicles; $B(i,j)$ represents the position of the vehicle. Then, use the following equation to establish the membership function of the urban rail transit flow state:

$$g = \frac{\bar{v}+t_g}{\sigma_p} \times S_z \tag{14}$$

where $\bar{v}$ represents the average speed of the specified vehicle; σ_p represents the offset angle of the center of mass during vehicle operation. Based on this, the variation characteristics of the speed of rail transit flow are extracted as follows:

$$D = S_z \times \frac{Z_x}{L_g \times g} \tag{15}$$

where Z_x represents the normal operating state of the rail vehicle, and L_g represents the flow rate of the rail transit line.

By utilizing the state characteristics of vehicles traveling on urban rail transit lines, a membership function for the flow state of urban rail transit is established, and the variation characteristics of the flow velocity of urban rail transit are extracted.

Step 2: Determine the layout scale of urban rail transit lines. Based on the characteristics of changes in the speed of urban rail transit flow, establish a scale index for line layout, and calculate the total length L, traffic density ρ, and transportation capacity E of urban rail transit lines:

$$\rho = \frac{N_{all}}{S_s}$$
$$E = \sum_{i=}^{n} E_i \tag{16}$$

where N_{all} represents the population of the planned area; S_s represents the planned area of urban agglomeration rail transit lines; E_i represents the passenger flow capacity of the i-th rail transit line.

According to the total length L of urban rail transit lines, the total amount of urban rail transit is calculated through regression as follows:

$$\beta = L \times \eta \times \alpha \tag{17}$$

where η represents the average travel rate of the urban population; α represents the expected population size of the urban agglomeration. According to β, the passenger flow of urban rail transit lines can be calculated as follows:

$$W = \beta \times \gamma \tag{18}$$

where γ represents the proportion of rail transit in urban agglomeration transportation [12].

Step 3: Calculate the optimal solution for urban rail transit line planning. For the urban rail transit network, with a vehicle arrival rate of δ, the probability density function of urban rail transit volume is constructed as follows:

$$\rho_R = \frac{\delta m}{P_{XY} \times v_t} + \frac{t_{jc} \times \kappa}{H^* \times (H_1 + H_2)} \tag{19}$$

where P_{XY} represents the probability of a station reaching m vehicles within the $X \sim Y$ time period; v_t represents the driving speed of the rail vehicle at time t; t_{jc} represents the duration of traffic volume statistics; κ represents information on rail transit flow; H^* represents a discrete interval; H_1 and H_2 represent the lower and upper limits of the discrete interval.

The optimal planning of urban rail transit lines can quickly evacuate the passenger flow of various transportation hubs. The following equation is used to construct an evacuation path model for the population flow of rail transit stations:

$$W_Q = \frac{H_{max} - H_{min}}{n_u \times n_{zy}} \times t_w \times \rho_R \tag{20}$$

where H_{max} and H_{min} represent the upper and lower limits of population flow; n_u represents the population flow evacuated on the u-th urban rail transit line; n_{zy} represents the number of railway hub stations; t_w represents the evacuation time. Then, multiple population genetic algorithms are introduced to record the migration path of ants in the population flow evacuation path model, namely:

$$L_A = \frac{W_Q \times m_A}{\varsigma_0} + I_R \times \varsigma_R \mp q_f \times I_f \tag{21}$$

where m_A represents the number of ants; ς_0 represents the initial pheromone concentration of the line; ς_R represents the pheromone concentration of the current line; I_R represents the importance of ς_R to the urban rail transit network; q_f represents heuristic information; I_f represents the importance of q_f to the urban rail transit network. If all m_A ants can independently search for an urban rail transit line, update the pheromone concentration of ants passing through each urban rail transit line to obtain the optimal solution for urban rail transit line planning:

$$R_k = \frac{\Delta\varsigma \times \varsigma_\forall}{t_{min}} \times L_A \tag{22}$$

where $\varsigma_\forall$ represents the pheromone evaporation coefficient, $\Delta\varsigma$ represents the amount of pheromone transformation during ant search, and t_{min} represents the shortest time for ant optimization.

3. Experiments and result analysis

3.1. Experimental design

The experiment was simulated in a MATLAB environment with the following parameters: computer operating system: Windows 10; 200 iterations; The penalty coefficient is 1000; The smoothness weight is 0.5; The spatial dimension is 5; The weight of the transportation route is 0.8; The inertia weight is 0.6897. Building an experimental platform for urban agglomeration transportation route layout planning with the support of the above experimental parameters. In the platform, load the transportation network of a certain urban agglomeration as the experimental object, where there are already 5 transportation lines. The specific simulated transportation network is shown in Figure 4.

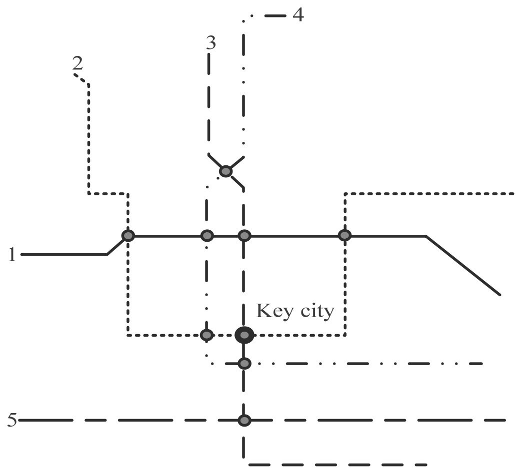

Fig. 4 - Simulated transportation network

In Figure 4, both Line 1 and Line 5 are rail transit systems that run through the eastern and western parts of the urban agglomeration. Line 3 and Line 4 are rail transit systems that run through the southern, northern, and eastern parts of the city. Line 2 passes through the central city and can form a hub for road traffic with Line 1, Line 3, and Line 4.

Import the transportation network shown in Figure 4 into the experimental platform, and use method of this paper, method of Wang and He [15], and method of Zhou et al. [20] to optimize the layout planning of urban agglomeration transportation lines.

3.2. Experimental indicators

In the experiment, the total carbon emissions and layout coordination of the urban agglomeration transportation system were used as indicators. Among them:

- The total carbon emissions of the urban agglomeration transportation system refer to the total amount of carbon emissions generated by all transportation activities within the urban agglomeration. This includes various modes of transportation such as road transportation, public transportation, and rail transit. The specific calculation method for this indicator can be found in section 2.1 above.
- The coordination of layout refers to the degree to which the transportation route planning and layout of various cities within the urban agglomeration are coordinated and coordinated with each other. From the perspective of carbon reduction policies, this indicator measures whether the layout of transportation lines in urban agglomerations can effectively support and promote the goal of carbon reduction. The specific calculation method for this indicator is as follows:

  $$h_{\varphi} = \frac{C_{\varepsilon}}{C} + \frac{N_{\beta}+N_{\sigma}}{N_{\vartheta}+N_{\gamma}} \tag{23}$$

  where h_{φ} represents the coordination index of the layout; C_{ε} represents the low-carbon emissions, which refers to the carbon emissions reduced by adopting low-carbon transportation modes; N_{β} represents reducing traffic congestion, which refers to the amount of traffic congestion reduced after optimizing transportation lines and hubs; N_{σ} represents the reduction in traffic accidents, which refers to the number of accidents reduced after the reasonable design of traffic lines and hubs; C represents the total carbon emissions; N_{ϑ} represents the total traffic congestion volume; N_{γ} represents the total number of traffic accidents.

3.3. Results display

Verify the application performance of method of this paper, method of Wang and He [15], and method of Zhou et al. [20] using the total carbon emissions of the urban agglomeration transportation system as an indicator. The total carbon emissions of the urban agglomeration transportation system after applying different methods are shown in Table 1.

By observing Table 1, it can be seen that with the increase of testing time, the cumulative total carbon emissions of the urban agglomeration transportation system increase. After applying the method of Wang and He [15], the total carbon emissions have exceeded 2 million tons after 12 hours. After applying the method of Zhou et al. [20], the total carbon emissions reached nearly 1 million tons at 6 hours. After applying the method of this paper, although the total carbon emissions continued to increase, the increase was relatively small, reaching 1.123 million tons in 12 hours.

Tab. 1 - Comparison of total carbon emissions of urban agglomeration transportation systems after applying different methods (10000 tons)

Test time/h	Method of this paper	Method of Wang and He [15]	Method of Zhou et al. [20]
2	17.6	26.3	27.8
4	32.8	58.4	57.0
6	56.9	72.6	90.5
8	66.1	139.7	120.6
10	89.2	180.4	141.1
12	112.3	229.8	178.1

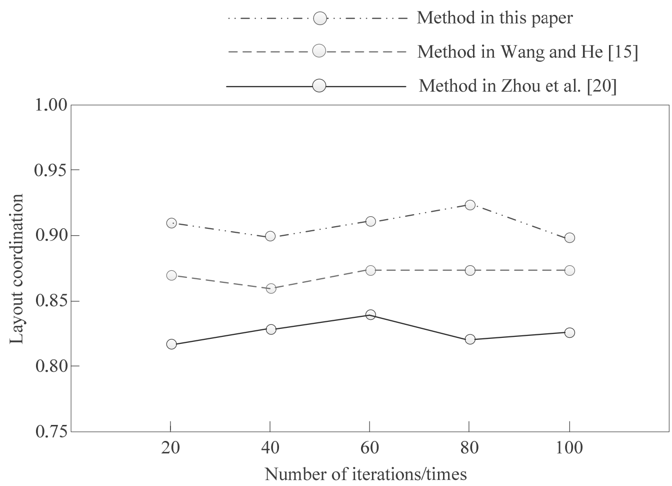

Fig. 5 - Comparison of coordination of layouts using different methods

By comparing and analyzing the above results, it can be concluded that the application of the route layout planning method designed in this article can effectively control the total carbon emissions of the urban agglomeration transportation system and meet the requirements of carbon reduction policies.

To further demonstrate the feasibility of the method of this paper, based on equation (23), the coordination of the layout of method of this paper, method of Wang and He [15], and method of Zhou et al. [20] is verified. The coordination of layouts using different methods is shown in Figure 5.

By observing Figure 5, it can be seen that as the number of experimental iterations changes, the coordination of the layout of transportation lines in urban agglomerations is also constantly changing after applying the three methods. Among them, after applying method of Wang and He [15], the coordination of the layout of urban agglomeration transportation lines can reach a maximum of 0.875, the method of Zhou et al. [20] is 0.821, and the method proposed in this paper is 0.926. In contrast, after applying the method of this paper, the coordination of the layout of urban agglomeration transportation lines is higher, indicating that the layout planning effect of the method of this paper is better.

4. Conclusion

This study introduces an innovative carbon-reduction-oriented planning approach for transportation lines in urban agglomerations. By collecting vehicle emission data, it establishes an optimization model to minimize travel time and carbon emissions, while enhancing the flow of rail transit. This comprehensive strategy takes into account the complexity of urban transportation, aiming to improve accessibility and coordination while significantly reducing carbon emissions. This approach supports sustainable urban development, promoting green and low-carbon travel modes, and contributes to creating a more livable and environmentally friendly urban environment.

References

1. Afandizadeh, Z.S., Memarnejad, A., Mirzahossein, H. (2021) Hourly Origin–Destination Matrix Estimation Using Intelligent Transportation Systems Data and Deep Learning. Sensors, 21(21), PP. 7080-7089.
2. BAO, C., LI, X.J. (2022) Low-carbon Transportation Planning of Ribbon Group City. Planners, 38(S1), PP. 71-75.
3. HU, L.W., ZHAO, X.T., YANG, J.Q., ZHANG, S.H., FAN, Z.J., YIN, X.F., GUO, Z. (2021) Urban Traffic Congestion Mitigation and Optimization Study Based on Two-level Programming. Journal of Transportation Systems Engineering and Information Technology, 21(05), PP. 222-227+234.
4. Jiang, F., Ma, L., Broyd, T., Chen, W.Y., Luo, H.B. (2022) Digital twin enabled sustainable urban road planning. Sustainable Cities and Society, 78(03), PP. 16-22.
5. Jiang, F., Ma, L., Broyd, T., Li, J.D., Jia, J., Luo, H.B. (2023) Systematic framework for sustainable urban road alignment planning. Transportation Research Part D, 120(16), PP. 311-318.
6. LV, Y. (2021) Research on Evaluation Index of Integrated Rail Transit Network Layout in Metropolitan Area. Railway Standard Design, 65(04), PP. 25-30+35.
7. Jiang, F.L., Li, L., Tang, Y.M., Zhang, H.L., Liu, X.P. (2023) A Facility Layout Algorithm for Logistics Scenarios Driven by Transport Lines. Applied Sciences, 13(12), PP. 1556-1564.
8. LI, T.T., YANG, J.W. (2020) Optimization Model of Highway Passenger Transport Networkin Urban Agglomeration Coordinating with Intercity Rail Transit. Journal of Chongqing Jiaotong University (Natural Science), 39(06), PP. 1-7+52.
9. LIU, H., ZHANG, N., ZHANG, P. (2020) A Method for Selecting Transportation Facilities in Urban Agglomeration Based on Weighted Grid Overlay. Statistics & Decision, 36(05), PP. 181-184.
10. LIU, S.J., TONG, X.R. (2021) Urban transportation path planning based on reinforcement learning. Journal of Computer Applications, 41(01), PP. 185-190.
11. LIU, X.J., LI, Z.M., LU, L.L. (2022) Optimal model of rail transit feeder bus lines based on bi-level programming. Journal of Shanghai Maritime University, 43(03), PP. 50-55.
12. Sai, W., Wang, H.Z. (2021) Optimal design of urban transportation planning based on big data. Environmental Technology Innovation, 23(2), PP. 129-135.
13. Wang, C.H.J., Hess, D.B. (2021) Role of Urban Big Data in Travel Behavior Research. Transportation Research Record, 2675(4), PP. 222-233.
14. WANG, L.B., JIN, J.G. (2021) Designing Urban Rail Transit Network Expansion for Optimal Origin-Destination Demand Capture. Journal of Hunan University of Science and Technology(Natural Sciednce Edition), 36(02), PP. 37-45.
15. WANG, Y., HE, T.H. (2020) New City (District) Road Traffic Planning Methods and Optimization Strategies. Highway, 65(08), PP. 301-305
16. WEN, H.Y., LIN, Y.F., WU, H.S., JIANG, H., WU, J.B. (2021) Extended Co-evolutionary Algorithm for Path Planning Based on the Urban Traffic Environment Evolution. Journal of South China University of Technology(Natural Science Edition), 49(10), PP. 1-10.
17. WU, J.Y. (2023) Research on Urban Traffic Management Planning and Design Technology under the Background of Green Ecology. Environmental Engineering, 41(02), PP. 317-328.

18. Yang, B.Y., Tian, Y., Wang, J., Hu, X.W., An, S. (2022) How to improve urban transportation planning in big data era? A practice in the study of traffic analysis zone delineation. Transport Policy, 127(1), PP. 1-14.
19. Zhong, Q.K., Liu, Q.H. (2021) Automatic Generation of Urban Road Planning Network Under Deep Learning. Journal of Physics: Conference Series, 2074(1), PP. 187-196.
20. ZHOU, J., XU, X.H., ZHOU, J.N., DENG, X.Q. (2022) Planning of Urban Rail Transit Lines Radiating from Urban Central Area: A Case Study of Shenzhen. Urban Transport of China, 20(02), PP. 82-89.

Advances in Transportation Studies an international Journal 2024 Special Issue, Vol. 1

ISBN 979-12-218-1230-5 – ISSN 1824-5463 – DOI 10.53136/97912218123052 – pag. 15-24

Urban road traffic congestion prediction based on knowledge graph

J.C. Long[1] L.J. Xie[1*] H.Y. Xie[2]

[1]*School of Electronics and Information Engineering, Wuhan Donghu University, Wuhan 430212, China*
[2]*Hengyang branch of China Telecom Co., Ltd, Hengyang 421000, China*
**corresponding author; email: xielijuan19882023@163.com*

Abstract

Urban road traffic congestion prediction holds significant research importance as it enables timely response to traffic flow patterns, optimizes urban planning, and enhances transportation efficiency. In order to improve the accuracy and recall of predicting urban road traffic congestion, a knowledge graph-based intelligent prediction approach is provided in this paper. This method first analyzes the measurement indicators of urban road traffic congestion status, and uses vehicle speed as the evaluation standard to develop a measurement standard for congestion degree. Secondly, a spatiotemporal knowledge graph of urban transportation networks was constructed using multi-source spatiotemporal data. Finally, based on the constructed spatiotemporal knowledge graph and the correlation of dynamic traffic, a novel prediction model for urban road traffic congestion was constructed using graph convolutional neural networks to obtain traffic congestion prediction results. The experimental results validate the effectiveness of the proposed approach, and indicate its significant potential for practical applications.

Keywords - knowledge graph, urban roads, traffic congestion, intelligent prediction

1. Introduction

With the acceleration of urbanization, road traffic congestion has become a common problem facing the world. Traffic congestion has had many negative impacts on the lives, work, and economic activities of urban residents, such as extended commuting time, environmental pollution, and resource waste. Solving traffic congestion requires accurate prediction and analysis of traffic flow to develop corresponding traffic management and planning measures. Traditional traffic prediction methods are mainly based on historical data and statistical models, but they are difficult to adapt to the complexity and real-time requirements of urban transportation systems. With the development of new generation information technology, intelligent transportation systems are gradually emerging. Collecting a large amount of traffic data through new technologies such as sensors, GPS positioning, and intelligent devices, and using algorithms such as artificial intelligence and machine learning for data analysis and modeling, to achieve intelligent prediction of urban road traffic conditions [7]. The intelligent traffic prediction system can monitor and analyze traffic data in real-time, predict future traffic flow, detect congestion in advance, and provide corresponding traffic optimization solutions. This has important guidance and decision-making significance for urban traffic management departments and traffic participants.

At present, many scholars have conducted research on intelligent prediction methods for urban road traffic congestion and have achieved certain results. Li et al. [9] proposes a predicting urban traffic congestion method based on deep learning. This method first uses convolutional neural networks to extract spatial features of urban area traffic data based on grid partitioning, and then enhances the model's expressive power through fully connected neural networks. Next, the similarity location encoding mechanism is used to incorporate location information into traffic data, and finally, the Transformer network is used to capture the time-dependent characteristics of traffic data, in order to achieve urban traffic congestion prediction. However, the literature points out that this method has the problem of poor real-time data acquisition and cannot accurately predict urban traffic congestion, resulting in poor application effectiveness. Ming et al. [14] proposed an improved method to address the issue of incomplete analysis of velocity characteristics. This method is based on floating bus data, combined with speed time correlation and spatial correlation analysis, introducing two characteristics of bus flow and time occupancy. An improved particle swarm optimization radial basis function neural network speed prediction model considering spatiotemporal and bus flow characteristics is proposed. The prediction results are compared with speed thresholds to obtain the traffic congestion situation on urban main roads. However, in the application process, this method did not fully consider the data of non bus route sections, resulting in inaccurate prediction of urban traffic congestion and limited application effectiveness. Liu et al. [12] proposed a short-term traffic flow prediction model based on spatiotemporal feature fusion. The model first uses point mutual information algorithm to analyze the correlation of monitoring stations, determine the stations with high correlation, and then process their traffic data into periodic and neighboring sequences. Next, introduce a long short-term memory network to extract temporal features and construct relevant models to complete the fusion of temporal and spatial features. Introduce absolute error sequence analysis to optimize the model and obtain the final prediction result. However, in the process of spatiotemporal data processing, this method has certain shortcomings, resulting in poor application effectiveness and inability to be widely applied.

To solve the problems in the above methods, this study introduces knowledge graph technology to complete the design of an intelligent prediction method for urban road traffic congestion.

2. Overview of measuring urban road traffic congestion status

The degree of traffic congestion in the urban road network is the basis for the urban traffic management department to obtain the traffic status of the urban road network, as well as the basis for urban traffic management and control. Therefore, it is necessary to evaluate the degree of traffic congestion in the urban road network. Vehicle speed is an important indicator reflecting road traffic conditions. On urban roads, circular detectors can be used to obtain vehicle speed information. On this basis, this article takes vehicle speed as the evaluation indicator, selects two values (30 and 50) as the dividing points of urban road traffic conditions, and completes the design of congestion measurement standards, as shown in Table 1.

Based on this, carry out subsequent research on predicting road traffic congestion status.

Tab. 1 - Measurement standards for congestion level

Speed value	Traffic situation	Traffic status description
[0,10]	Severe	Average speed is very low and the road traffic is in poor condition
[10,20]	Congestion	Low average speed and poor road traffic conditions
[20,30]	Slight	Low average speed and poor road traffic conditions
[30,40]	Normal	Average speed is average, and the road traffic is in good condition
[40,+∞]	Unblocked	High average speed and good road traffic conditions

3. Construction of spatiotemporal knowledge graph for urban transportation network

In order to better understand and analyze the structure, state, and trend of urban transportation systems, it is necessary to first construct a spatiotemporal knowledge graph of urban transportation networks. This process requires collecting and integrating spatiotemporal traffic data from different data sources such as traffic monitoring systems, GPS trajectory data, mobile device data, etc., and conducting conceptual modeling and entity extraction to determine the entity types and association relationships required to construct a knowledge graph [11, 1, 3, 8]. Meanwhile, considering the spatiotemporal characteristics of traffic data, corresponding spatiotemporal attributes are added to each entity and relationship to reflect the dynamic changes of the traffic network. Finally, store the constructed spatiotemporal knowledge graph in an appropriate database, and use query languages or APIs to query and analyze the knowledge graph to obtain information about the structure and features of the transportation network. The entire process is as follows:

1) Obtain multi-source spatiotemporal data. This time, a certain city's road network information and urban taxi trajectory data were used as initial data. Among them, the road network data is stored in shp file format, which includes basic information of roads in all administrative areas under a certain city and the connection relationships between roads [4]. In terms of trajectory data, continuous GPS trajectory points of 2000 taxis in the city were selected as sample data. The trajectory data of these taxis includes detailed information on their driving paths and stopping positions, and was recorded at a sampling frequency of approximately 25 seconds.

2) Preprocess multi-source spatiotemporal data to ensure accuracy and consistency. Let P represent the trajectory point data sequence, which can be represented as $P = (t, x, y)$, t represents timestamp, and x represents longitude; y represents latitude, and the formula for calculating its velocity is as follows:

$$V = \frac{(x_{i-1}-x_i)^2+(y_{i-1}-y_i)^2}{t_{i-1}-t_i} \tag{1}$$

After calculating its speed according to equation (1), set the threshold T and compare it with the calculated value. If $V < T$, it is judged as a normal value, otherwise it is considered an outlier, and the outlier is removed [5].

3) Based on the preprocessed data, conduct conceptual modeling and entity extraction to determine the entity types and association relationships that need to be included in the spatiotemporal knowledge graph. The entity model includes roads and vehicles. Based on the road information in the road network data, each road has unique attributes such as ID, road name, road type, length, and road segment [15, 13].

4) Add spatiotemporal attributes, add corresponding spatiotemporal features for each entity and relationship, and then store the constructed spatiotemporal knowledge graph in the database.

Following the above steps, construct a spatiotemporal knowledge graph of urban transportation network. The specific spatiotemporal knowledge graph of urban transportation network is shown in Figure 1. As shown in Figure 1, the traffic status of each time period in the city forms a graph, and its dynamic graph is defined as $G_t = (Q, E, A_t)$. Here, G_t represents the urban traffic status map of time period t, Q represents the set of congestion status nodes in each section of the urban road network, E represents the geographical connections between roads in the road network, and A_t represents the adjacency matrix of the urban traffic map in time period t [2].

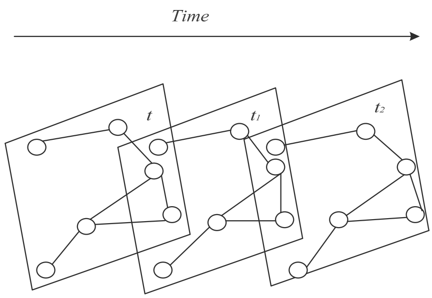

Fig. 1 - Spatial and temporal knowledge graph of urban transportation network

In summary, the construction of a spatiotemporal knowledge graph for urban transportation networks has been completed.

4. Design of urban road traffic congestion prediction

Based on the spatiotemporal knowledge graph of the urban transportation network constructed above, fully considering its dynamic traffic correlation, a graph convolutional neural network is used to construct a prediction model for urban road traffic congestion status. The expression is as follows:

$$O_k = f_o(G_n, O_n) \tag{2}$$

where G_n represents the set of dynamic traffic maps for a given historical period n; O_n represents a set of observed values of historical urban traffic congestion status, which can be judged based on the criteria in Table 1; O_k represents the predicted traffic congestion status for the future k period; n represents the time interval of observed historical urban traffic data; k represents the time interval for predicting urban traffic congestion; f_o represents the urban traffic congestion prediction function [16]. The prediction model for urban road traffic congestion status is shown in Figure 2. As shown in Figure 2, the model mainly consists of dynamic adjacency matrix, LSTM network, stepwise multi-layer graph convolutional network, and spatiotemporal graph convolutional network.

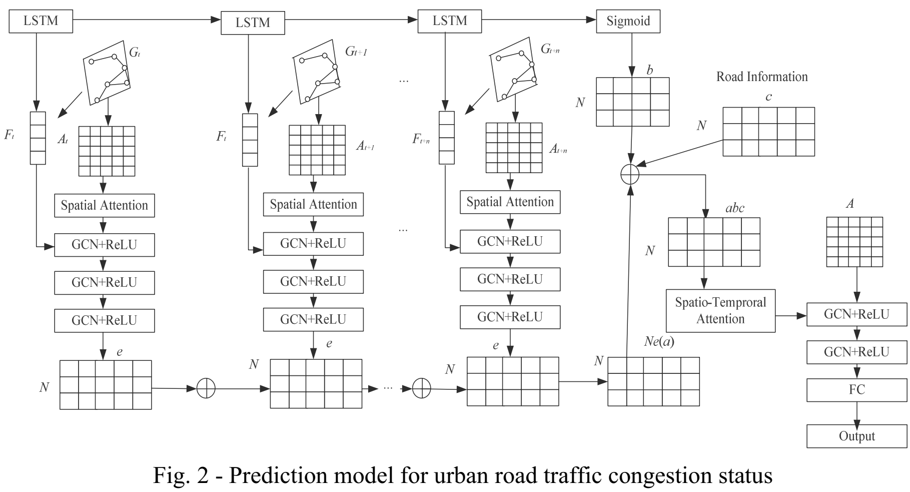

Fig. 2 - Prediction model for urban road traffic congestion status

Its design is as follows:

1) Dynamic adjacency matrix. This matrix is mainly used to describe the dynamic spatial correlation of the initial transportation network and the fixed geographical connections of the urban road network, which can be represented as:

$$A_t = S^t + A \tag{3}$$

where S^t represents the dynamic correlation matrix of the urban transportation network, and A represents the fixed adjacency matrix of the fixed geographical connection of the urban transportation network, where:

$$S^t = N^t \tag{4}$$

$$A = \begin{cases} 1, \ if \ \ v_i \ \ and \ \ v_j \ \ is \ \ connected \\ \qquad \qquad 0, if \ \ not \end{cases} \tag{5}$$

where N^t represents the number of vehicles transferring from section i to j during period t; v_i and v_j represents the speed of vehicles transferring from section i to j. Through the above, the establishment of a dynamic adjacency matrix is completed, which is dynamic and can not only represent the fixed geographical connections of urban transportation networks, but also the spatial correlation connections of urban transportation dynamics.

2) LSTM network. This time, the network is mainly used to learn long-term dependency sequence relationship data, in order to capture the evolution patterns of dynamic transportation networks [6]. For period t, the LSTM unit takes the current input vector X_t and the previous state vector h_{t-1} as inputs, and then outputs the state vector h_t in the current period. The process is as follows:

$$i_t = \sigma(W_x^i x_t + W_x^i h_{t-1} + b^i) \tag{6}$$

$$f_t = \sigma\left(W_x^f x_t + W_h^f h_{t-1} + b^f\right) \tag{7}$$

$$o_t = \sigma(W_x^o x_t + W_h^o h_{t-1} + b^o) \tag{8}$$

$$s_t = f_t \odot s_{t-1} + i_t \odot \widetilde{s_t} \tag{9}$$

$$\widetilde{s_t} = \sigma(W_x^s x_t + W_h^s h_{t-1} + b^s) \tag{10}$$

$$h_t = o_t \odot tanh(s_t) \tag{11}$$

where i_t, f_t, o_t and s_t represent input gates, forget gates, output gates, and storage units, respectively; W_x, W_h, b is the learnable parameter of the corresponding unit, and $\sigma(.)$ represents the activation function; $\odot$ represents the Hadamard matrix multiplication of the corresponding elements. Through the above, capture the dynamic temporal correlation of urban road traffic [19, 17].

3) Step wise multi-layer graph convolutional network. The stepwise multi-layer graph convolutional network based on attention and dynamic adjacency matrix designed in this project is used to capture dynamic spatial correlations in urban traffic data. Assuming there are N nodes in the graph with M dimensional features, the topology and node properties of the graph can be represented by the adjacency matrix A and the feature matrix F. This time, the adjacency matrix is used to obtain its corresponding Laplace matrix, which can be expressed as:

$$\mathrm{L} = \mathrm{D} - \mathrm{A} \tag{12}$$

where D represents the degree matrix, which can represent the degree of each node in the graph.

The multi-step multi-layer graph convolution module designed in this project can perform multi-layer graph convolution operations at each historical time period to capture the spatial characteristics of the current historical urban road network. The graph convolution process in the multi-step multi-layer graph convolution module is as follows:

$$f\left(H_t^i, A_t\right) = ReLu\left(D^{-1/2}\dot{A}_t D^{-1/2} H_t^{i-1} W_t^i\right) \tag{13}$$

where $\dot{A}_t$ represents the normalized value of the adjacency matrix A_t; i represents the convolutional layer of the i layer graph; f represents an iterative function used to summarize the feature information of neighboring nodes and oneself; H_t^i represents the output of the convolutional layer of the i layer graph in the t period; W_t^i represents the learnable weight matrix of the i layer graph convolutional layer [10].

In order to adaptively capture important dynamic spatial correlations between nodes and achieve higher accuracy in predicting future urban traffic congestion, a new spatial attention algorithm p^t was utilized in the first graph convolutional layer of each time period. Considering the strong temporal correlation of urban transportation, the attention mechanism of this article is as follows:

$$p^t = Q_t^s \sigma\left((F^{t-1} W_t^s) W_t^p (W_t^\propto A) + b_t^s\right) \tag{14}$$

where p^t represents the spatial attention torque of time period t, and F^{t-1} represents the feature matrix of the previous time period; A represents the fixed adjacency matrix of the graph. Q_t^s, W_t^s, $W_t^\propto$, and b_t^s are learnable parameter matrices. $\sigma(.)$ is the activation function. Then, this time, the softmax function is used to ensure that the sum of attention weights for each node is 1. Each element value in P represents the semantic strength of the correlation between node i and node j. This module has a total of n different spatial attention matrices p^t. When performing the first graph convolution operation, the adjacency matrix A and the spatial attention matrix p^t will be combined to dynamically adjust the influence weights between nodes.

Through the above calculation, the urban traffic spatial feature vector representation $H_1, H_2, \ldots, H_n$ for n time periods can be obtained. Among them, e represents the number of hidden units in the last graph convolutional layer. This time, the spatial feature vector representations of all historical time periods are merged as the dynamic spatial features of the urban road network in n historical time periods, which can be expressed as:

$$H_S = H_1 \odot H_2 \odot, \ldots, \odot H_n \tag{15}$$

Through the above, achieve dynamic spatial correlation capture in urban transportation data.

4) Spatiotemporal graph convolutional network. It is used to simultaneously integrate the dynamic spatiotemporal dependencies and road information features obtained from the stepwise multi-layer graph convolutional network and LSTM network mentioned above, in order to achieve urban traffic congestion prediction. According to urban road traffic information, the characteristics of road information include the number of lanes, whether the road is one-way or two-way, and the type of road. This time, the unique hot encoding method was used to feature encode three types of information, represented as F_l, F_d, F_y. The vectors of the three road features were merged as the overall information features of the road, which can be represented as:

$$F_r = F_l \oplus F_d \oplus F_y \tag{16}$$

Assuming that the spatial features generated by the attention based graph convolutional network module are a dimensional F_s, the time dependence generated by the LSTM network is b dimensional F_t, and the road information feature is c dimensional F_r, the spatiotemporal feature vector and the road information feature vector are merged as the complete feature vector of the road:

$$F_o = F_s \oplus F_t \oplus F_r \tag{17}$$

To adaptively learn the impact of each feature on future urban traffic congestion status, this paper uses a fully connected layer to implement spatiotemporal attention mechanism R_o:

$$R_o = w_o F_s + b_o \tag{18}$$

where w_o, b_o represent learnable parameters.