<u>Original Research Article</u> Road Network Pre-partitioning Method with Priority for Congestion Control

ABSTRACT

Urban traffic congestion seriously affects the traffic efficiency, causing travel delays and resources wasted directly. In this paper, a road network pre-partitioning method with priority for congestion control is proposed. Traffic status of Intersection is pre-estimated through CNN-SVM model. Subarea congestion coefficient and intersection merger coefficient are defined to expand the control area of congestion coordination. Spectral clustering is used for non-congested intersection partitioning. The results show that the proposed method could reduce the number of congested intersections faster, and reduce road network congestion coefficient more quickly in the later stage of simulation, which is an effective subarea partition method.

Keywords: Road Network Pre-partition; Congestion Control; CNN-SVR; Spectral Clustering

1. INTRODUCTION

1.1 Background and significance

Traffic congestion is a problem with high frequency, great impact, and long duration. Through traffic coordination control, the traffic efficiency of the road network can be optimized and vehicle waiting time can be reduced. The traffic network is divided into several control areas by dividing adjacent intersections in the road network. The areas are controlled by a specific strategy, and coordinated with each other, which can effectively improve the traffic control effect ^[1]. Traffic state estimation can monitor the state of the transportation system, make full use of the road capacity, and then provide guidance to traffic managers ^[2].

1.2 Related works and paper outline

The traffic control subarea partitioning method mainly includes cluster analysis method and road network traversal analysis method. The cluster analysis method mainly combines the intersections of similar intersections through the traffic attributes of the intersections^[3]. The road network traversal analysis method traverses the road network according to certain rules and completes the subarea partition^[4].

Traffic control subarea partition has always been the key issue of traffic coordination control. Gong divided the influencing factors of subarea into static factors and dynamic factors, and established intersection correlation model from cycle length, row length and traffic flow, and then developed dynamic partition algorithm based on the model^[5]. Liu proposed a dynamic traffic area partition scheme based on game theory. Each part of the road network was used as a rational game participant to seek the most similar payment between itself and the traffic area^[6]. Meshkat optimized partitioning through the Multi-Objective Evolutionary Algorithm, including the Fast and Elite Non-Dominant Sorting Genetic Algorithm and the Pareto Archive

Evolution Strategy^[7]. Cao used fuzzy control theory to determine the coordination coefficient of adjacent intersections. The coefficients considered five factors: cross distance, signal period, traffic flow, traffic flow impact factor and queue length^[8]. Shen proposed a fuzzy computing method based on hierarchical structure to estimate the correlation, and based on the correlation, the control area of the urban main road was divided into subareas^[9]. Hu proposed a new vehicle detector layout method for continuously collecting and storing high resolution traffic data that identified key turning points in traffic state changes. Therefore, a correlation model of quantitative correlation between adjacent intersections on the main road based on traffic condition data was established^[10]. Tang divided the peak traffic network into supersaturated regions and their related regions, and used gray correlation analysis and spectral clustering methods to divide the relevant regions. The traffic coordination control model of urban oversaturated zone and its related areas was proposed^[11].

The current subarea partitioning method optimizes the traffic efficiency to a large extent, but most of the research focus on the low-saturation and medium-saturation, aiming at reducing traffic delays, lacking of research on over-saturation. The subarea partitioning method proposed in this paper starts from the perspective of congestion control, optimizes the efficiency of the road network based on the priority control of the congestion area, and combines the estimated state of the road network to realize the pre-partition of the control subarea. The rest of this article is organized as follows. In Section 2, we use the CNN-SVM model to estimate the state of the intersection. In Section 3, we use subarea congestion coefficient and intersection merger coefficient to complete the subarea partitioning of congestion areas. In Section 4, we use spectral clustering to complete the subarea partitioning of congested areas. In Section 5, we validate our algorithm by simulation. In Section 6, we summarize the paper.

2. ESTIMATION OF TRAFFIC STATUS AT INTERSECTION BASED ON CNN-SVM MODEL

In the process of congestion control, the use of the current information of the road network for coordinated control often has hysteresis. Therefore, accurate traffic state estimation is conducive to faster control measures and improved congestion control effects.

Traditional studies have achieved estimates of traffic conditions to varying degrees, but have higher requirements for monitor accuracy^[12]. The traffic system is a nonlinear system with strong uncertainty. Many phenomena cannot be studied by deterministic analysis methods, and methods of uncertainty analysis need to be considered. Combining the advantages of CNN and SVM, this paper uses multiple monitoring points data to conduct congestion prediction, and then improves traffic state estimation accuracy.

2.1 Traffic Flow Feature Extraction Based on CNN Model

Convolutional neural network is a kind of neural network specially used to process similar network structure data. With the development of deep learning, it has performed well in many fields, including image classification, object monitoring, semantic segmentation, etc^[13]. Structurally, CNN is mainly composed of a convolutional layer, a pooled layer, and a fully connected layer.

To predict the traffic state of monitoring point P at time t, it can be considered from both time and space. In time, the traffic state will not be abrupt in the short term, and the traffic state at the next moment can be considered as the continuation of the traffic state at the previous times. In space, the traffic state of monitoring point P is bound to be affected by

the state of the surrounding intersection. The original traffic state data is transformed into a two-dimensional feature matrix containing time information and spatial information. The longitudinal axis direction of the characteristic matrix is sorted according to the spatial position of the detector, and in the horizontal axis direction, the data monitored by the detector is sorted in the time sequence, and the $x_{p,t}$ is the traffic state of the detector P at the time t, and the corresponding input matrix is:



Where U is the number of import road detectors, D is the number of exit road detectors, and T is the length of the interval.

In terms of congestion factors, this paper uses speed, flow, road occupancy, and large vehicle ratio as input from the data detected by the detector. Therefore, $x_{p,t}$ is a vector of length 4. The input of CNN is a three-dimensional matrix composed of congestion factors, space and time, and the output is the traffic state estimation result.

2.2 Traffic State Classification Based on CNN-SVM Model

The SVM algorithm is a machine learning method based on statistical learning theory. It is based on minimizing structural risk, making empirical risk closer to actual risk, and improving sample scalability. The nonlinear transformation is used to map the sample space to the high-dimensional space, and the optimal linear classification hyperplane is found in the high-dimensional space to minimize the risk and the generalization ability of the algorithm^[14].

According to the degree of congestion, this paper divides the traffic state into three categories: smooth, crowded and congested. The reference indicator is the PeMS(Performance Measurement System) traffic data delay item of the California Department of Transportation Performance Measurement: the delay term is 0, indicating that the vehicle has no delay and is in smooth state; the delay term is greater than 0 and less than 1, indicating that the vehicle has a slight delay and is in crowded state; the delay term is greater than 1, indicating a severe delay and is in congested state.

In terms of extracting features, the CNN model can be automated, avoiding the impact of manual extraction of features on the results. In the classification problem, the SVM learning hyperplane is the plane farthest from the sample points of each category, and the classification accuracy rate is more advantageous. To combine the advantages of CNN and SVM, the CNN-SVM classification model is used to estimate traffic status. CNN performs feature extraction on traffic data, and SVM uses the extracted features to classify traffic status. The specific structure is shown in figure 1.



Fig. 1. Structure of CNN-SVM model

The model uses the data of the first 30 minutes of the current time to estimate the traffic state. The traffic characteristics are sampled in the first 15 minutes, the first 20 minutes, the first 25 minutes, and the first 30 minutes, so the input length of the time dimension is 4; The spatial dimension information uses 3 intersection flows into the intersection and 3 intersection flows out of the intersection, so the input of the spatial dimension is 6, and the spatial dimension can be adjusted according to the number of adjacent intersections; the feature dimension is commonly sampled by 4 features. Therefore, the model is entered as a three-dimensional matrix.

The CNN model consists of three convolutional layers and a fully connected layer, and finally a traffic state is estimated by a Softmax. After the CNN training is finished, the fully connected output is taken as a feature and input into the SVM model for classification. The CNN-SVM model training is divided into two processes. The first step is to train the CNN model using traffic data. In the second step, the fully connected output is used as a feature, and the SVM model is trained using the features extracted by CNN.

The size of the CNN model convolution kernel is 2×2 , the number of three convolution kernels is 6, 12, and 24, the number of neurons in the fully connected layer is 128, the activation function uses the relu function, and the output uses softmax. The model uses the cross entropy loss function, and the training process is optimized by the Adam algorithm. The SVM model uses the hinge loss function and the RBF kernel function with a penalty factor of 0.8. The experiment was conducted using PeMS traffic data from September 20, 2017 to October 27, 2017. The accuracy of the estimated traffic state in the training reaches 94.68%, and the accuracy of the test set reaches 95.32%. The accuracy meets the estimated requirements.

3. PRIORITY PARTITION OF CONGESTION CONTROL SUBAREA

3.1 Subarea congestion coefficient and intersection merger coefficient

Most traditional subarea partition methods divide subareas according to the similarity and correlation of intersections, tending to divide the intersections with similar saturation into the same subarea. In this paper, the goal of accelerating congestion disappear is considered. A dynamic network partitioning method with priority for congestion control is proposed. The congested area and other non-congested intersections around it are divided into the same subarea for congestion coordination control to improve the congestion control effect.

In this paper, from the idea of balanced congestion, the subarea congestion coefficient(SCC) is defined by the average saturation rate (ASR) of all roads in the road network and the average delay time (ADT) of each vehicle in the road network:

$$SCC = ASR + \alpha \times ADT \tag{3.1}$$

ASR represents the ratio of the actual flow of the road to the loadable flow. ADT represents the ratio of the actual running time of the vehicle to the theoretical running time. α is the weighting factor. In this paper, α is 1.

In order to determine the intersections to be merged next in the congestion subarea, define the intersection merger coefficient (IMC) for all intersections adjacent to the congestion subarea:

$$IMC = ICC + \beta(1-D) \tag{3.2}$$

The ICC (intersection congestion coefficient) is the SCC corresponding to all the road constituent regions adjacent to the intersection; The distance between the intersection to be merged and the intersection with the highest congestion coefficient in the target sub-area is normalized to [0, 1], which is D. β is the weight coefficient, which is 0.6 in this paper.

3.2 Partition of congestion control subarea

When the subarea is initially divided, the CNN-SVM model is used to determine the congestion intersections in the road network, and the adjacent congestion intersections are merged into one control subarea. Then, the intersections adjacent to the congestion subarea and having the largest IMC are merged to form a new subarea, and the SCC is calculated. If the SCC is greater than the subarea congestion coefficient threshold (SCCT), the subarea is still congested. The subarea should continue merging intersections.

Subarea with the higher congestion coefficient is first merged. After each merge operation, the judgment is re-determined, and the subareas with the highest congestion coefficient are selected from all the congestion subareas to merge the adjacent intersections until the congestion coefficient of all the congestion subareas falls below the SCCT. For a given subarea, the intersections with larger coefficients to be merged are preferentially merged when selecting the intersections that need to be merged.

Algorithm 1 Congestion subarea partitioning algorithm based on subarea congestion coefficient and intersection merger coefficient

- 1: Input : Traffic network structure G(V,E) and traffic data
- 2: CNN-SVM model identify congestion intersections set as V_{Congestion}
- 3: for all intersections vi \in V_{Congestion} do
- 4: Merge adjacent congestion intersections vi to form congestion subareas set as S_{Congestion}
- 5: If($V_{Congestion} == V$) Output: $S_{Congestion}$, End Algorithm 1
- 6: If(∀s_i∈S_{congestion}, SCC(si)<SCCT) Output: S_{Congestion}, End Algorithm 1
- 7: Choose s_i has largest SCC as s_{target}
- 8: Choose vi has largest IMC as vtarget
- 9: s_i=s_{target}+{v_{target}}
- 10: $V_{Congestion} = V_{Congestion} + \{v_{target}\}$
- 11: return step 5.

4. NON-CONGESTION SUBAREA PARTITION BASED ON SPECTRAL CLUSTERING

4.1 Establishment of spectral clustering connection matrix

Spectral clustering is a clustering algorithm whose main idea is abstracting the data into points in space and connecting them by edges. The closer the points are, the higher the weight of the edges is. The different subgraphs are made by using the cut graph whose idea is making the sum of connection weights between subgraphs is the lowest, and the sum of inner edges of the subgraph has the highest weight. Compared with the K-means algorithm, spectral clustering is more adaptable to data distribution and easier to implement^[15].

For non-congested subareas, the spectral clustering method is used for partitioning. From the road network, the intersections of all the congested subareas and their adjacent roads are removed from the road network structure diagram, and then the spectral clustering algorithm is used. The key of the spectral clustering algorithm is the establishment of the connection matrix.

The subarea partition of the traffic network is generally based on the model of association and similarity. The association model uses the Whitson model in the 1987 US Traffic Control System Handbook ^[16]:

$$I_b = \frac{0.5}{1+t} [I_f - 1] \tag{4.1}$$

$$I_f = \frac{h \cdot q_m}{\sum_{k=1}^{h} q_k} \tag{4.2}$$

$$=\frac{l}{v} \tag{4.3}$$

Where: I_b is the association degree between the upstream and downstream intersections; I_f is the road flow imbalance coefficient; h is the number of traffic flow branches entering from the upstream intersection; q_m is the maximum traffic flow in the upstream direction of the upstream intersection; $\sum_{s=1}^{h} q_s$ is the downstream Total flow at the intersection; t is the average travel time on the road segment; l is the length of the road; v is the average speed;

The similarity mode uses the traffic level between different intersections^[17].

$$d_x^y = \sum_{r=1}^{\min\{h_x, h_y\}} |s_{xr} - s_{yr}|$$
(4.4)

$$R_x^y = \max\{d_x^y, 1 \le x, y \le n\} - d_x^y$$
(4.5)

 d_x^y represents the saturation dissimilarity of intersections x and y, and h_x and h_y represent the number of imported roads of intersections $_x$ and y, respectively. The calculation of the saturation dissimilarity is based on the small number of imported roads in the two intersections. s_{xr} and s_{yr} represent the saturation of intersections x and y, respectively, and R_x^y is the similarity of intersections x and y. Combining the association and similarity of the intersections, the connection weights between adjacent intersections of the road network are obtained, and α is 0.6:

$$w_{xy} = \alpha \cdot I + (1 - \alpha) \cdot R_x^y \tag{4.6}$$

For graph G, the set of points is V and the set of edges is E. The degree d_i of the point i is the sum of all the edge weights connected to the point i, and the degree matrix D is a matrix whose diagonal is d_i and the remaining values are 0. The adjacency matrix of the graph W is defined as a matrix consisting of W_{ij} which represent direct connection weights of point i and point j. Define the Laplacian matrix: L = D - W

4.2 Non-congestion subarea partition algorithm based on spectral clustering

Algorithm 2 Non-congestion subarea partition algorithm based on spectral clustering

- 1: Input: Road network structure and traffic informations
- 2: Construct a spectral cluster adjacency matrix W_{xy} , where the weight of the non-adjacent intersection is 0, and the weights of the adjacent intersections x and y are:

$$W_{yy} = \alpha \cdot I + (1 - \alpha) \cdot R_y^y$$

- 3: Construction degree matrix D, and Laplacian matrix L
- 4: Constructing a standardized Laplacian matrix $D^{-\frac{1}{2}}LD^{-\frac{1}{2}}$
- 5: Calculate the eigenvector f corresponding to the smallest k_1 eigenvalues of $D^{-\frac{1}{2}}LD^{-\frac{1}{2}}$
- 6: Normalize the feature vector f to form an k_1 dimensional feature matrix F
- 7: For each row in *F* as an k_1 dimensional sample, *n* samples are clustered, and the clustering dimension is k_2 .
- 8: The partition of the non-congestion subarea is obtained with the cluster partition $C(c_1,c_2,...c_{k_2})$.
- 9: End Algorithm 2

After the subarea is divided, the initial timing of each intersection is completed by the Webster timing method ^[18] without considering the influence of the subarea. The period of the longest intersection in the subarea is taken as the common period of all the intersections in the subarea, and the corresponding phase time is proportional to the change of the period on the basis of the original timing, to complete the initial timing of the entire road network. To get the final timing of the road network, "mountain climbing method" ^[19] is used to optimize

the basis of the initial timing. The flow diagram of road network pre-partitioning method with priority for congestion control is shown in figure 2.



Fig. 2. Flow diagram of road network pre-partitioning method with priority for congestion control

5. SIMULATION RESULTS AND ANALYSIS

5.1 Simulation Settings

In order to verify the effectiveness of the congestion priority network subarea partitioning algorithm, the United States Los Angele road network is abstracted to obtain the road network structure diagram. The road network includes 21 signalized intersections and 35 road sections. Set the start status of the road network, as shown in Table 1. Simulations were performed using VISSIM.

The experimental parameters are as follows: the road is 4 lanes, the saturated flow rate is 1600 pcu/h, and the two-phase signal is controlled. The loss time of each phase signal is 5.2 s, and the yellow light time is 4 s. The default parameters are used for the regional timing parameters in the "mountain climbing method", and the congestion coefficient threshold SCCT is 3.

Table 1. Length and initial flow of each road in the traffic network

Road number	Road length (m)	Initial flow (pcu/h)	Association	Similarity	Road weight
1	250	1000	0.51	1.39	0.86

2	400	1000	0.45	0.68	0.54
3	900	1000	0.24	0.57	0.37
4	800	1000	0.25	1.68	0.82
5	400	1400	0.36	0.51	0.42
6	150	1400	0.68	0.18	0.48
7	200	1000	0.6	1.05	0.78
8	150	1000	0.78	0.91	0.83
9	850	1500	0.25	0.68	0.42
10	850	1000	0.26	1.29	0.67
11	500	1000	0.3	1.00	0.58
12	400	1400	0.35	0.38	0.36
13	150	1800	0.78	1.13	0.92
14	150	2200	0.76	1.06	0.88
15	250	2200	0.54	1.59	0.96
16	250	1500	0.51	0.14	0.36
17	150	1000	0.79	0.89	0.83
18	900	1000	0.24	1.19	0.62
19	250	1000	0.5	1.35	0.84
20	300	1000	0.45	0.75	0.57
21	400	1000	0.38	0.86	0.57
22	450	1200	0.39	1.02	0.64
23	400	1200	0.36	0.94	0.59
24	300	1500	0.41	0.24	0.34
25	250	1400	0.51	0.49	0.5
26	200	900	0.62	0.67	0.64
27	150	900	0.77	0.57	0.69
28	250	1000	0.55	1.40	0.89
29	200	1400	0.6	0.38	0.51
30	250	1400	0.51	0.64	0.56
31	700	1400	0.29	0.99	0.57
32	900	1000	0.21	0.61	0.37
33	250	1200	0.52	0.67	0.58
34	250	1200	0.56	1.44	0.91
35	300	1200	0.46	1.39	0.83

5.2 Simulation Results

The subarea is divided into the road network, and the influence of different subarea partition modes on congestion control is compared. The subarea partitioning results obtained by directly using spectral clustering for the entire road network without considering the congestion priority are shown in Fig. 3. The results of the network partitioning method for congestion control priority is shown in Figure 4.



Fig. 3. Spectral clustering for partition Fig. 4. Congestion priority control for artition It can be seen from the figure 4 that the subarea partitioning with priority of congestion increases the range of the congestion control subarea. In this example, the three congested intersections are coordinated and controlled by the subareas composed of six intersections. The comparison of three partitioning method is shown in figure 5 and figure 6.



Fig. 5. Number of congested intersections Fig. 6. Congestion coefficient of road based on three partitioning method network based on three partitioning method

It can be seen from the results that the congestion priority control has great advantages in reducing the number of congested intersections, and plays a role in relieving road congestion. In terms of vehicle delay time, the spectral clustering for subarea partition algorithm is better in the early stage, because reducing congestion will sacrifice the performance of other aspects of the road network. However, with the reduction of congestion, the congestion priority control partition method can reduce more vehicle delay time.

6. CONCLUSION

In order to control traffic congestion and reduce the size of the congestion area, a subarea partitioning method with priority for congestion control is proposed. The traffic conditions at the intersections are estimated through the CNN-SVM model to determine the congestion intersections. The congestion coordination control area is expanded by the subarea congestion coefficient and the intersection merger coefficient, while spectral clustering algorithm for non-congested intersection partitioning. Results showed that the proposed method could reduce the number of congested intersections faster, and reduce road network congestion coefficient more quickly in the later stage of simulation, which is an effective subarea partitioning method.

REFERENCES

[1] Zhou Z, Lin S, Xi Y. A fast network partition method for large-scale urban traffic networks[J]. Journal of Control Theory and Applications, 2013, 11(3): 359-366.

[2] Backfrieder C, Ostermayer G, Mecklenbräuker C F. Increased traffic flow through nodebased bottleneck prediction and V2X communication[J]. IEEE Transactions on Intelligent Transportation Systems, 2017, 18(2): 349-363.

[3] Wu N N, Yan X P, Huang G H, et al. Urban environment-oriented traffic zoning based on spatial cluster analysis[J]. Journal of Environmental Informatics, 2015, 15(2): 111-119.

[4] Wang Z, Li L, Wu B. Managing Regional Congestion in Urban Road Networks: An Approach Based on Fuzzy Control Theory[M]//CICTP 2014: Safe, Smart, and Sustainable Multimodal Transportation Systems. 2014: 1966-1980.

[5] Gong B W, Bie Y M, Liu Z Y. A Dynamic Algorithm of Partitioning Urban Road Network into Traffic Control Subareas[C]//Applied Mechanics and Materials. Trans Tech Publications, 2013, 409: 1398-1402.

[6] Liu H, Yang L, Zhang Y, et al. A dynamic traffic zone partition scheme based on game theory[J]. JOURNAL OF INFORMATION & COMPUTATIONAL SCIENCE, 2013, 10(10): 2961-2969.

[7] Meshkat A, Vrancken J L M. Multi-objective road network partitioning[J]. Procedia-Social and Behavioral Science, 2014, 1.

[8] Cao J, Zhang L. Research on Subarea Partition of Regional Traffic by Fuzzy Theory[J]. JOURNAL OF INFORMATION & COMPUTATIONAL SCIENCE, 12(6): 2429-2437.

[9] Shen G, Yang Y. A dynamic signal coordination control method for urban arterial roads and its application[J]. Frontiers of Information Technology & Electronic Engineering, 2016, 17(9): 907-918.

[10] Hu Y, Wang Y, Zhang J, et al. Correlation degree analysis of arterial adjacent intersections for coordinated control subunit partition[J]. Advances in Mechanical Engineering, 2018, 10(1): 1687814017748748.

[11] Tang S, Zhu W, Liu X, et al. Urban Traffic Cooperative Control Based on Regional Partition[C]//2018 37th Chinese Control Conference (CCC). IEEE, 2018: 7754-7759.

[12] Yang Y, Xu Y, Han J, et al. Efficient traffic congestion estimation using multiple spatiotemporal properties[J]. Neurocomputing, 2017, 267: 344-353.

[13] Sharif Razavian A, Azizpour H, Sullivan J, et al. CNN features off-the-shelf: an astounding baseline for recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition workshops. 2014: 806-813.

[14] Li Z, Yuan R, Guan X. Accurate classification of the internet traffic based on the svm method[C]//2007 IEEE International Conference on Communications. IEEE, 2007: 1373-1378.

[15] Li Y, Nie F, Huang H, et al. Large-Scale Multi-View Spectral Clustering via Bipartite Graph[C]//AAAI. 2015: 2750-2756.

[16] Lu K, Xu J M, Zheng S J, et al. Research on fast dynamic division method of coordinated control subarea[J]. Acta Automatica Sinica, 2012, 38(2): 279-287.

[17] Xu J M , Yan X W , Jing B B , et al. Dynamic Network Partitioning Method Based on Intersections with Different Degree of Saturation[J]. Journal of Transportation Systems Engineering & Information Technology, 2017.

[18] Betz C L, Allmaras R R, Copeland S M, et al. Least limiting water range: traffic and long-term tillage influences in a Webster soil[J]. Soil Science Society of America Journal, 1998, 62(5): 1384-1393.

[19] Robertson D I. Research on the TRANSYT and SCOOT Methods of Signal Coordination[J]. ITE journal, 1986, 56(1): 36-40.

MATHRA