A Network Partition Approach for MFD-Based Urban Transportation Network Model

Haitao Xu[1][*], Weiguo Zhang[1], Zuozhang zhuo[1]
[1]School of Computer Science and Technology, Hangzhou Dianzi University
1158, No. 2 Street, Baiyang Street, Hangzhou Economic Development Zone, Hangzhou
[email: xuhaitao@hdu.edu.cn; 182050081@hdu.edu.cn]

Abstract

Recent findings identified the scatter and shape of MFD (macroscopic fundamental diagram) is heavily influenced by the spatial distribution of link density in a road network. This implies that the concept of MFD can be utilized to divide a heterogeneous road network with different degrees of congestion into multiple homogeneous subnetworks. Considering the actual traffic data is usually incomplete and inaccurate while most traffic partition algorithms rely on the completeness of the data, we proposed a three-step partitioned algorithm called Iso-MB (Isoperimetric algorithm - Merging - Boundary adjustment) permitting of incompletely input data in this paper. The proposed algorithm was implemented and verified in a simulated urban transportation network. The existence of well-defined MFD in each subnetwork was revealed and discussed and the selection of stop parameter in the isoperimetric algorithm was explained and dissected. The effectiveness of the approach to the missing input data was also demonstrated and elaborated.

1. Introduction

The phenomenon of urban traffic congestion has affected modern urban life and restricted urban development. Many large cities carry out various traffic coordination strategies to alleviate traffic congestion. However, the large and complex characteristics of urban road networks and the spatiotemporal distribution of traffic congestion varies make it impossible to coordinate the traffic of the entire road network. In this context, the concept of traffic network partition was proposed [1]. Dividing a heterogeneous road network into several homogeneous subnetworks helps to implement fitting routing tactics, border control, traffic flow control, and management for different regions.

The research on the traffic partition algorithm mainly lies in two aspects: calculating the correlation index and the network partition algorithm. The correlation index measures the degree of correlation between nodes and is the necessary step for the partition algorithm. In the study of the correlation index, Yagoda et al. [2] considered the flow and distance between the two adjacent intersections; Chang et al. [3] regarded traffic flow, upstream coordinated phase flow, and the number of lanes on the road as influencing factors; Whitson [4] model combined the length of the route with the flow of the trip; Lin et al. [5] utilized the degree of congestion and Bie et al. [6] considered the period and the queuing situation. Based on previous research, Geo et al. [7] proved the existence of MFD (macroscopic fundamental diagram) in the urban transportation network. Unlike others, by utilizing the properties of MFD, one can model road networks from a macro perspective without knowing detailed road information, and the required data is easy to be gathered. It was proved in [8] the spatial distribution of link density in road networks has a great influence on the scatter and the shape of an MFD. This means one can establish a road network correlation model and divide a heterogeneous road network with different levels of congestion into multiple homogeneous road networks according to the properties of MFD. The partition methodology proposed in this paper can obtain subregions with different congestion levels. Furthermore, this methodology demonstrated the superiority of effectiveness, time complexity, and robustness compared to other traffic network partition algorithms.

In terms of the traffic network partition algorithm, clustering algorithms are widely employed. Such methods utilize the spatial and traffic information to construct a network model and partition the entire road network into several homogeneous subnetworks. Y. Ji et al. [9] introduced the NMB algorithm based on Ncut to partition urban transportation networks. Y. Ma et al. [10] partition network using Laplacian eigenvectors. A. F. Lentzakis et al. [11] attempted to partition a time-dependent network model by means of k-means, k-harmonic means, and normalized spectral clustering. Z. Zhou et al. [12] introduced a dynamic network partition method based on community detection algorithm. B. Wang et al. [13] proposed Parallel Traffic Sub-Areas Division (PTSD) method based on the parallel k-means algorithm. Anwar T et al. [14] built a density peak graph model according to the traffic density and adjacent connectivity, then applied α-Cut to partition the graph. Wang Z et al. [15] developed a traffic network partitioning method based on fuzzy clustering analysis. These algorithms have a disadvantage: They establish a correlation model based on the premise that they have accurate data on all links in the network, without considering the incompleteness of the actual data. Nevertheless, the incompleteness of traffic data is a common problem with reasons for equipment malfunctions, unavailability of equipment [16, 17], etc. This will cause the actual partition results ha25ve big difference with the ideal results, and the desired effect cannot be achieved. Some scholars applied the data preprocessing technique to eliminate the negative impact of missing data. Clélia Lopez et al. [18] coarsen network and estimate speed through travel time

data to establish a traffic network model, and cluster data point to obtain subnetworks. Some heuristic algorithms were recommended to partition traffic networks, they enable partition road networks in the absence of partial data but at the price of high computational complexity. For instance, M. Saeedmanesh et al. [19,20] introduced the "snake" similarity and symmetric non-negative matrix factorization, Kang An et al. [21] combined lambda-connectedness with region growing technique to extract homogeneous parts of the network.

An algorithm with low computational complexity and superior robustness is what we wanted. The isoperimetric algorithm was introduced by Grady L et al. [22] for solving the graph partition problem. It belongs to the spectral clustering algorithm, but partition graph by solving linear equations, which suggests it has the property of low computation complexity. The algorithm was originally implemented in image segmentation and shows excellent segmentation results [23]. Grady L et al. also explored the robustness of the isoperimetric algorithm and Ncut by adding different kinds and levels of noise, the experimental outcome demonstrated the superior segmentation stability of the isoperimetric algorithm compared to Ncut. They also found that the isoperimetric runs three to ten times faster than Ncut. Due to equipment and human factors, it is a common problem in real-life the traffic data is incomplete, which is most partition algorithms cannot overcome. However, based on the robustness of the isoperimetric algorithm, the homogeneous parts of the road network can be identified even if the traffic data is incomplete.

In this paper, we proposed a three-step algorithm called Iso-MB (Isoperimetric algorithm - Merging - Boundary adjustment) to carve up a heterogeneous road network into homogeneous subnetworks. First, a weighted graph is established based on the geographic information and link density of the road network, and the isoperimetric algorithm is utilized to partition it. Then, in the premise of without affecting the homogeneity and the spatial compactness of each subregion, considering that a well-defined MFD requires sufficient links in the subnetwork, and implementing the traffic strategy requires the number of links in each subregion be at the same level, a merging algorithm is introduced to implement iteratively until the desired number of new subregions is obtained. Finally, to facilitate the implementation of traffic control strategies and further optimize the partition results after the merge, a boundary adjustment method is presented.

The contributions of the approach include but are not limited to the following points:
- A new three-step algorithm that divides a heterogeneous network into several homogeneous subnetworks.
- Permitting incomplete data as input meets the demand for actual data.
- Low time complexity guarantees real-time road network partition.

The remainder of this paper is organized as follows: Section 2 describes the specific steps of the Iso-MB algorithm. Then, in section 3, the algorithm is applied upon a simulated traffic network to verify the effectiveness of the algorithm, relate experiments on the numerical selection of stop parameter are analyzed, and the sensitivity analysis of the algorithm to different missing rate data are displayed. In section 4 we make a review and summary of this research, the discussion about future work is also presented.

2. Methodology

Our primary objective is to partition an entire heterogeneous transportation network into several homogeneous subnetworks so that the presence of well-defined MFDs with low scatters in each subnetwork is warranted. More concretely, the following goals are required to reach for the algorithm we propose:

- Low variance of link densities in each subnetwork is needed to guarantee the existence of a well-defined MFD.
- Large density variance between the extracted subnetworks ensures that each subnetwork has a different degree of congestion.
- Ensure the number of links in the subnetworks is balanced. Avoid large differences in the number of links in the subinterval or too small numbers of links in the subregion.
- Guarantee the spatial compactness of each subregion to facilitate the implementation of traffic strategies.

Based on the objects above, the following algorithm flow consisting of three steps was designed:

Step1.   Divide the entire road network into several homogeneous subnetworks by applying the isoperimetric algorithm. It can effectively identify local homogeneous components at the link level and guarantee space compactness.

Step2.   Iteratively merge a pair of adjacent similar subregions, until the number of the subregions reaches the desired value.

Step3.   Continuously adjust the boundaries of each subregion, which can reduce the internal link variance of each subregion.

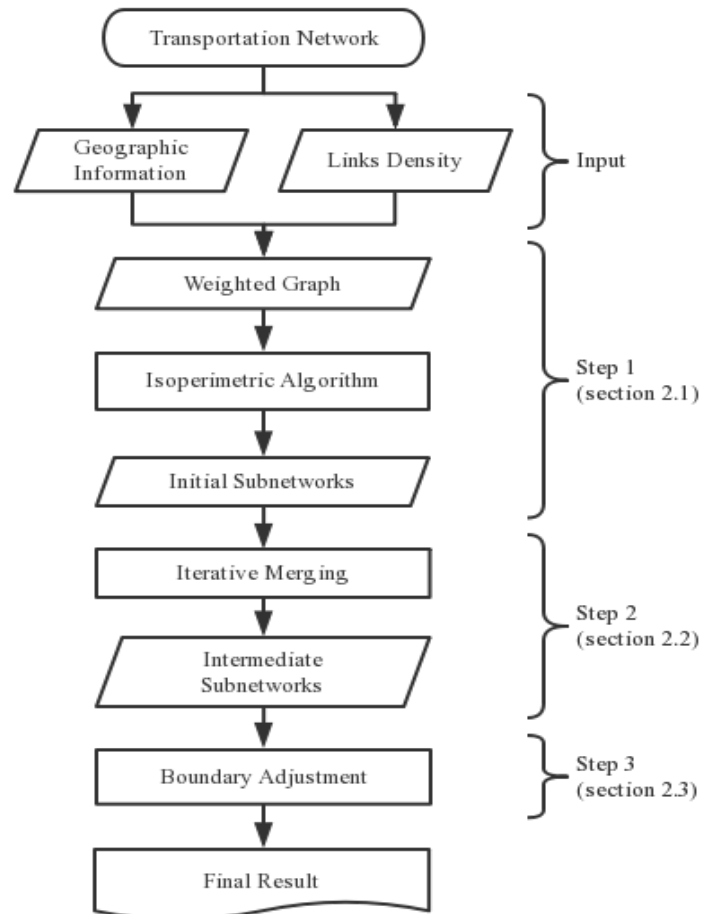The following sections describe the implementation of each step in detai

Fig. 1. Algorithm Flow Chart

2.1 Initial Partitioning

In this step, the isoperimetric algorithm is performed to partition the entire heterogeneous road network into multiple homogeneous subnetworks. The algorithm was originally proposed by Leo Grady et al. [22] and belongs to the spectral clustering algorithm in graph theory clustering algorithms. It utilizes the Laplace matrix for clustering based on the traditional isoperimetric problem. Compared with the Ncut algorithm, the isoperimetric algorithm performs superiority in time complexity and stability [22, 23]. Although the isoperimetric algorithm can effectively extract the homogeneous part, it satisfies first and second goals to a certain extent, but it may divide the road network into many fragmented subregions, which violate the third and fourth goals. The problem would be solved in the next step.

Assume a weighted graph $G = (V, E)$ with nodes $v \in V$ and edges $e \in E \subseteq V \times V$ is established based on the link information of the road network. Each node $v_i \in V$ represents link i in the network and has a density value d_i at an exact time of day. Each edge $e_{ij} \in E$ is regarded as the correlation between the links, there exists e_{ij} if links i and j are adjacent and vice versa. The weighted matrix $\{w(i,j)\}$ measures the value of e_{ij}. We employ the weighting function in [23] define $w(i,j)$ as follows :

$$w(i,j) = \begin{cases} \exp\left(-\beta\left(d_i - d_j\right)^2\right), & \text{if } e_{ij} \in E, \\ 0, & \text{if } e_{ij} \notin E. \end{cases} \tag{1}$$

Where β is an arbitrary constant, the sensitivity of the weight $w(i,j)$ to the link density can be set by this value. The larger β is, the more sensitive $w(i,j)$ is.

The isoperimetric constant h_G [24] of a graph G is defined as

$$h_G = \inf_{S} \frac{|\partial S|}{Vol_S}, \tag{2}$$

Where S is an area in the network, $|\partial S|$ denotes the proportion of the boundary of region S, Vol_S indicates the volume of region S, and h represents the infimum of the ratio over all possible S. Derived from the formula (2) according to [23], h_G can be written as

$$h_G = \min_{x} \frac{x^T L x}{x^T d}, \tag{3}$$

Where x is the indicator vector taking a binary value at each node, $x_i=0$ if the node i is not in region S and $x_i=1$ if the node is in region S. L represents the Laplacian vector of graph G, and d is the degree vector of L.

Solving the Eq. (3) of the isoperimetric problem is an NP-hard problem [24], but through the method corresponding to [22], the problem can be solved by a linear formula instead of the eigenvector. The pseudocode of the isoperimetric algorithm is shown below:

Algorithm 1 Isoperimetric Algorithm

   Initialization:

      $G = \{v \mid$ vertices in Graph$\}$: set of vertices

      $s =$ stop parameter

      $D(G)$: the degree matrix of G

      $maxIndex(D)$: index of the largest element in D

      $L(G)$: the Laplacian matrix of G

      $X(L, D)$: indicate vector x

      $ratioCut(L, x)$: cut graph using ratio cut

Output: partition result R

$\quad D = D(G)$

$\quad L = L(G)$

$\quad i = maxIndex(D)$

$\quad L_0 =$ Remove the i-th row and the i-th column of L

$\quad D_0 =$ Remove the i-th row of D

$\quad x_0 = X(L_0, D_0)$

$\quad$ if $x_0 < s$ then

$\quad R = ratioCut(L_0, x_0)$

$\quad$ end if

return $\boldsymbol{R}$

The detailed steps for initial partitioning by the isoperimetric algorithm are as follows:
- Establish graph G as described above, compute weights of all edges using (1), obtain the Laplace matrix L, and degree vector d.
- Use the isoperimetric algorithm to obtain the corresponding indicator vector x_0, which yields a real-valued solution. Threshold the potentials x at the value that gives partitions corresponding to the lowest isoperimetric ratio if it is less than the stop parameter [23].
- Perform the above steps recursively on each subnetwork until the isoperimetric ratio of each subnetwork is larger than the stop parameter.

Time complexity of the above process is about O($n$log(n)) [23]. As a result of the initial partitioning, the entire road network is divided into many subregions, but some subregions have few links. To obtain a certain number of subregions, and each subregion has enough links to ensure a well-defined MFD, a merging algorithm is required.

2.2 Iterative Merging

We obtain the initial partition results after completing the step above, each subnetwork maintains the homogeneity of the link and the compactness of the space. However, as mentioned in the previous section, the initial partition results have the following disadvantages:
- There exist subregions with too few links.
- The number of subareas obtained is too large to meet the expected number of subareas and the purpose of the traffic network partition.

In the transportation network, subregions with only a few links are not desirable as (1) MFD might reveal high statistical errors and (2) simple perimeter control strategies cannot be easily designed for a network partitioned into a large number of subnetworks as route choice might change [9]. The Iterative merging algorithm is introduced to obtain new subregions with the purpose of acquiring a certain number of subregions we expected, ensuring the number of links in each subregion is sufficient, maintain the homogeneity and space compactness of the subregions after merging, ensure the well-defined MFD existed in each subregion.

Two merge patterns are performed in this step. The first pattern is to merge the smallest subarea with the adjacent subarea closest to its mean of link density, the second pattern is to merge two subregions with the smallest link density difference. Set the lower limit LL as the minimum number of links for a subnetwork and N as the desired number of subnetworks, count the number of links in the smallest subarea $MinS$. If $MinS$ is less than LL, apply the first pattern until $MinS$ is larger than LL, then use the second pattern and repeat merging until the number of subareas reaches N. In this step, we set 90 as the value of LL given the research in [9]. The following is the pseudocode of the iterative merging:

Algorithm 2 Iterative Merging

    Initialization:

      LL= lower limit

      N= desired number of subareas

      N_c= current number of subareas

      R = set of IDs, indicating the subarea each node belongs

      $MinS$ = the number of links in the smallest subarea

      *mergePattern*1(R): merge the smallest subarea with the adjacent subarea closest to its mean of link density

      *mergePattern*2(R): merge two subareas with the smallest link density difference

    Output: R = set of IDs, indicating the subarea each node belongs

      while $N_c > N$ do

        if $MinS < LL$ then

          *mergePattern*1(R)

          update *MinS*

        else

          *mergePattern*2(R)

        end if

      $N_c = N_c - 1$

      end while

      return R

The iterative merging is similar to the Fast Newman algorithm [25] whose complexity is about O($m(m+n)$). However, the merging process is much more sufficient on computation with time complexity of O($k\log(k)$) where k is the number of clusters after the initial partitioning, and usually $k << n$ where n is the number of nodes in the graph.

2.3 Boundary Adjustment

After the steps above, the general partition results have been formed, and the current results have met our predetermined goals. However, the current results can still obtain smaller subnetwork internal link density variances through boundary adjustments to pull in the spatiality of the partitions more compact. The main reasons for implementing boundary adjustments can be attributed to the following points:

- The border between the current two subregions is largely unstable, which means the ownership of a certain link is ambiguous. Even if the link changes the ownership, it will not have a great impact on the overall partition result.
- The isoperimetric algorithm tends to maximize the ratio of the sum weights of the internal nodes of the subregion to the sum weights of the boundary nodes, it does not purely aim at minimizing the variance of link density between subnetworks.
- No consideration for balancing or spatial compactness.

Therefore, by implementing boundary adjustment, the variance of the link density in each subnetwork can be further minimized, optimizes the 1 and 4 objects without violating the others. Y. Ji and N. Geroliminis [9] illustrated how making boundary adjustment correctly, and the corresponding boundary adjustment method of adjusting a group of consecutive spatially links on the boundaries was introduced. It was proved to be effective in decreasing the total variance and improving the spatial compactness of subregions. The time complexity of the boundary adjustment is O(s^2n) where s is the number of subgroups. We took advantage of the

boundary adjustment method in this step to overcome the above questions pointed out in this section. The detail procedures of the method are as follows:

- The spatially adjacent links on the boundaries were built as a sequence in each cluster.
- On each boundary, find a subgroup of spatially adjacent links that can minimize the total variance after moving these links from current subregions to adjacent subregions. The length of the subgroup must be under the constraints of an upper bound and lower bound, if no such subgroup is found, the algorithm stops.
- Select the subgroup that can minimize the total variance among all the boundary subgroups and remove it from the current subregion to adjacent subregions. Update the partitioned result.
- Execute the algorithm in a loop.

3.   Case Study

In this section, the Iso-MB algorithm was performed into a simulated traffic network. The consequences were exhibited and analyzed in detail, the compared outcomes of the NMB algorithm proposed in [9] were implemented to indicate the effectiveness of the algorithm. Furthermore, we showed the MFD curve of each subnetwork to certificate the existence of well-defined MFD. Then, the method for choosing the optimal value of the stop parameters was introduced. Finally, the robustness of the algorithm to missing traffic data was experimentally analyzed.

3.1 Data Source

The simulated transportation network was established on the strength of the available data set which was provided by the TAPASCologne project. The data set generation process in detail was described in [26, 27]. We selected a part of the simulated road network in Cologne, Germany as the experimental site. The selected road network has 707 links. The sumo simulation software was used to simulate the 24-hour time-varying traffic demand, and the spatiotemporal congestion distribution of the road network was also reproduced.

3.2 Partition Results

We verified the effectiveness of the Iso-MB algorithm by applying it to the road network at 8 a.m. **Fig. 2** illustrates the spatial link density of the network at this time. At this time, the whole network presents different levels of congestion, part of the road network is highly congested, which can fully verify the effectiveness of the algorithm. In this network, the algorithm is completed within dozens of seconds, which means the algorithm can be applied to real-time traffic control and management strategies. The network was partitioned into 2-5 subnetworks, and the following indicators were applied to measure the capability of the algorithm:

- Standard Deviation (σ) of Link Density: A metric estimates the homogeneity of each subnetwork, corresponding to the first goal.
- Average Values (μ) of Link Density: A metric measures the congestion level of the subregions, to indicate that each subregion has a different level of congestion.
- Average *NS* [9]: A variance metric that measures the dissimilarity of a subnetwork to its most similar neighbor, corresponding to the second goal. A small value of average *NS* demonstrates a good partition result.
- TV_n [19]: The normalized total variance, measures the ratio between the total variance of the partitioned and unpartitioned network. Smaller TV_n means better partition result.

- Link Count: the number of links in each subnetwork, corresponding to the third goal.

Fig. 2. Grayscale representation of link densities in the road network of cologne at 8 a.m.

Partition results with 2-5 subregions using the Iso-MB algorithm are shown in **Fig. 3**. For each partition outcome in **Fig. 3**, the boundaries between subregions are clear, which guarantees the feasibility of transportation management and control strategies, suggests the last goal is well achieved. Refer to the spatial distribution of link densities showed in **Fig. 2**, it's clear that the degree of congestion between each subregion and its adjacent subregions are different. **Table 1** illustrates the first three indicators of the results with 2-5 subregions. It is observed that the average link density of each subregion is significantly different from that of adjacent subregions, also indicating the road network is divided into subregions with different degrees of congestion and reaches the second goal. **Table 2** shows the link count of each subnetwork, demonstrates the achievement of the third goal, ensures that each subnetwork has sufficient links for well-defined MFD.

NMB method was applied in the cologne network, the partition results with 2-5 subregions are depicted in **Fig. 4**. Four indicators of the results are illustrated in **Table 3** and **Table 4**. By comparing the indicators in the two tables, the excellence of Iso-MB is proved from the following points:

- The partition results of Iso-MB have a lower σ than that of NMB.
- The result with 3 subregions of Iso-MB has the lowest average *NS* value (0.688), which is lower than the lowest average *NS* value (0.737) of NMB partitioned results.
- The TV_n of the partitioned results of Iso-MB are lower than that of NMB. It implies that the algorithm has greater improvement obtained by partitioning compared to NMB.
- The partitioned results of Iso-MB have larger differences in congestion between adjacent subregions.
- It shows better balance over links count of each subnetwork in **Fig. 3**, while some subnetworks in **Fig. 4** have too few links to ensure well-defined MFDs.
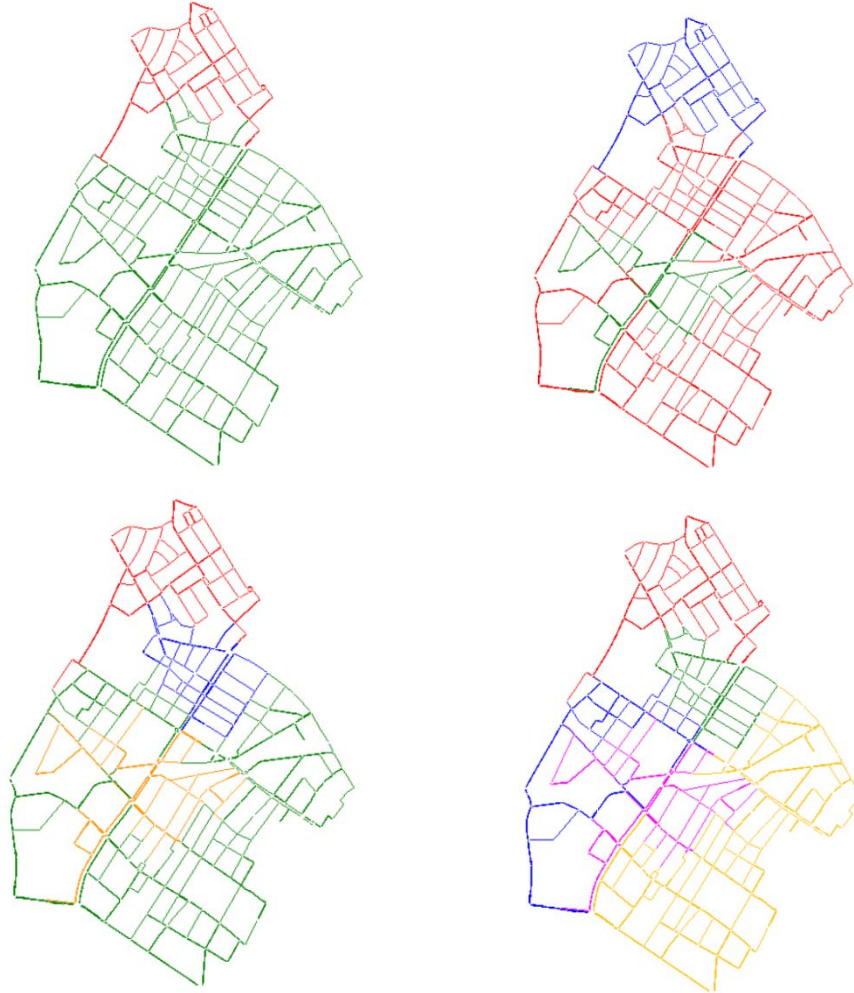
Fig. 3. Partitioned results of the Iso-MB algorithm for cases with 2–5 subregions.

Table 1. Average values (μ) and standard deviation (σ) of link density([veh/lane/km]), average *NS*, TV_n of obtained subnetworks for cases with 2–5 subregions.

(μ/σ)	red	green	blue	yellow	pink	Avg.*NS*	TV_n
2	19.5/41.0	64.46/62.4	-/-	-/-	-/-	0.760	0.923
3	69.4/66.1	33.6/46.5	19.5/41.0	-/-	-/-	0.688	0.915
4	19.3/46.9	73.5/66.5	55.7/62.9	33.6/46.5	-/-	0.807	0.928
5	21.5/43.0	55.0/63.2	81.8/62.1	70.3/67.7	32.6/46.7	0.848	0.911

Table 2. Link count of obtained subnetworks for cases with 2–5 subregions.

Link Count	red	green	blue	yellow	pink
2	124	583	-/-	-/-	-/-
3	456	127	124	-/-	-/-
4	123	353	104	127	-/-
5	131	96	111	243	126

Fig. 4. Partitioned results of the NMB algorithm for cases with 2-5 subregions.

Table 3. Average values (μ) and standard deviation (σ) of link density([*veh/lane/km*]) and average *NS*, TV_n of obtained subnetworks using NMB for cases with 2-5 subregions.

(μ/σ)	red	blue	green	yellow	purple	Avg.*NS*	TV_n
2	62.1/64.0	21.6/45.3	-/-	-/-	-/-	0.737	0.960
3	61.1/63.7	70.4/71.1	16.7/38.7	-/-	-/-	0.845	0.990
4	16.7/38.7	62.5/63.6	63.4/60.9	70.4/71.1	-/-	0.910	0.985
5	16.7/38.7	63.4/65.1	60.8/63.5	57.6/56.5	70.4/71.1	0.913	0.955

Table 4. Link count of obtained subnetworks using NMB for cases with 2–5 subregions.

Link Count	red	green	blue	yellow	pink
2	602	105	-/-	-/-	-/-
3	553	105	49	-/-	-/-
4	108	477	103	49	-/-
5	108	399	109	42	49

The shape of MFDs for the partition result with three regions is investigated, and the outcomes are illustrated in **Fig. 5**. With the network partitioning result, vehicle flow and density data were summated at 10 min intervals during the time from 5:00 to 9:00. The time dimension of this period in the MFD curves is represented by the gradient color from 0 to 24. The congestion distribution of the road network during this period was consistent with 8 a.m. It's clear that all three subnetworks exist MFD with clear shape and low scatter, indicating a satisfactory partitioned result.

The MFD curves of the three subregions are shown in **Fig. 5**. The larger road network (such as red region) has a clearer MFD diagram and fewer scattered points, while the smaller road network (such as green and blue region) has some scattered points, proves that the subnetwork needs a sufficient number of links to ensure a well-defined MFD. In particular, the scatter of MFD for the green region seems to be abnormal. A possible explanation is that the area locates in the suburb which is relatively closed in the simulated road network, its traffic can only pass through the red region. The analysis of the relationship between macroscopic flow and density from the time dimension was also performed. Between 8 and 9 a.m., The scatter of green and red regions have high density and low flow, indicating that the congestion levels of the two regions are high, while the traffic density of the blue region is still not saturated with low density and low flow. In this case, to alleviate traffic congestion and maintain the maximum throughput of the entire road network, it is necessary to introduce the flow of vehicles to the blue region or reduce the inflow of red and green regions.
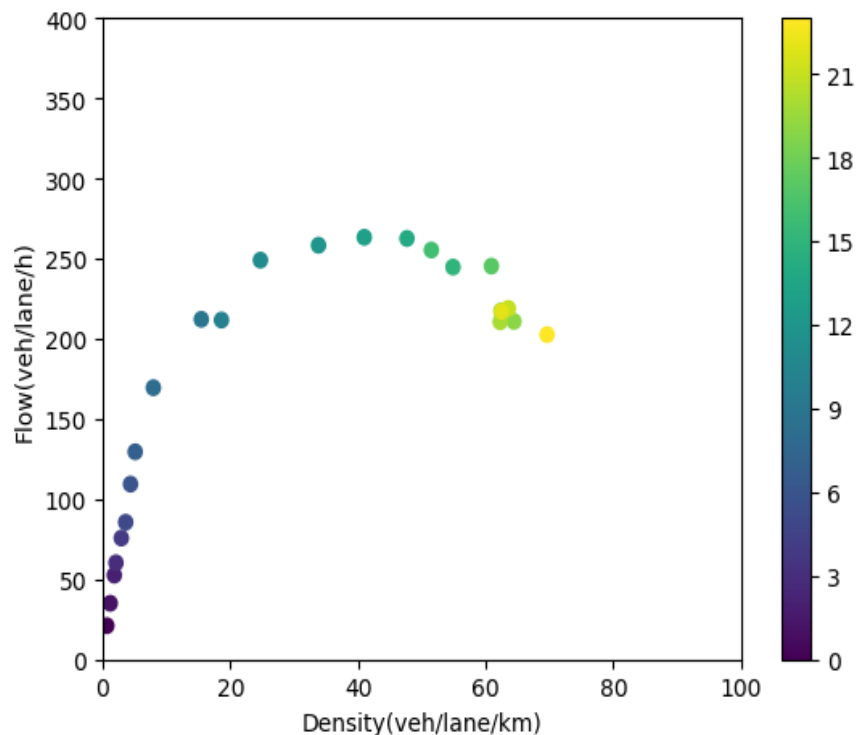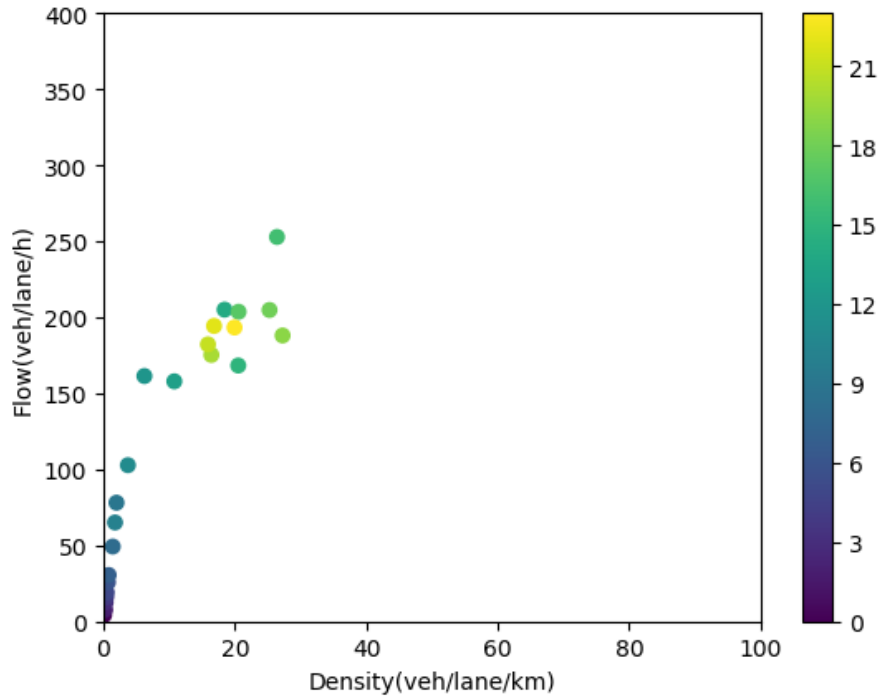

Fig. 5.1. MFD for the red region
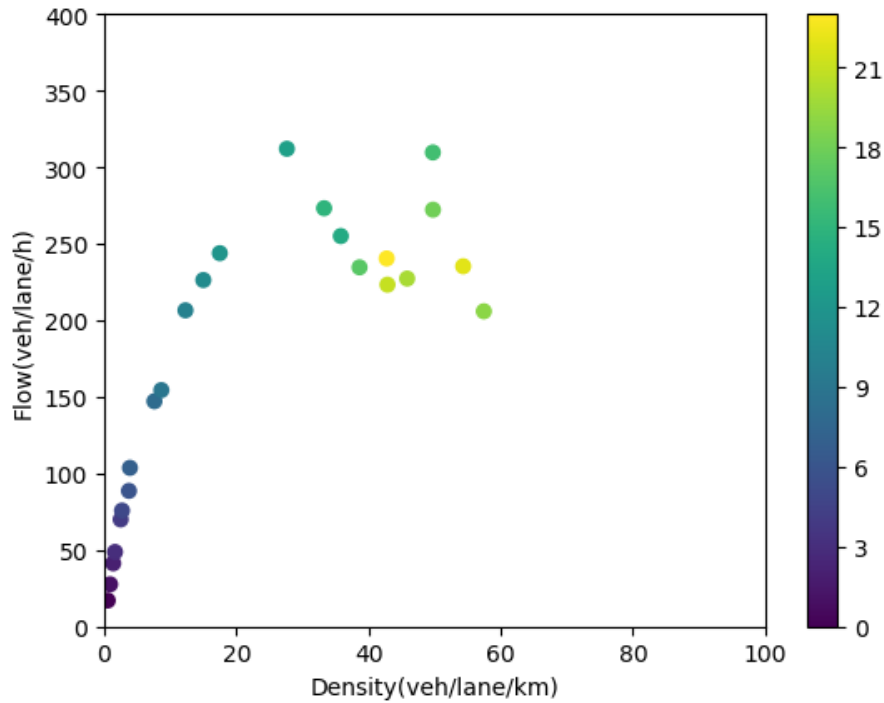
Fig. 5.2. MFD for the blue region

Fig. 5.3. MFD for the green region

Fig. 5. Subnetwork MFDs for Iso-MB partitioned result with 3 subregions

3.3 Selection of Stop Parameters

For the isoperimetric algorithm, the stop parameter is critical to the proposed algorithm, it directly affects the quality of the partitioned results. In this section, to explore the approach obtain the optimal stop parameter, the initial partitioned results employing different stop parameters are analyzed. Considering the evaluation standard in section 2, the description of the stop parameter and the sensitivity analysis for image segmentation in [22, 23], the following criteria are presented to determine the optimal value of stop parameter:

- The Number of Subregions: Fewer subregions are desired. Excessive subregions may obtain subregions with large density variance in the merging process.
- Subnetwork Density Variance: the criterion estimates the difference in average link density between subnetworks. Larger subnetwork density variance is desired, it demonstrates a larger difference in the congestion level between subnetworks.
- Stop Parameter: A lower value of the stop parameter corresponds to more desirable partition results, while a large value permits lower quality partition results. Tending to choose a lower value is required.

Fig. 6 shows the results of sensitivity experiments when the stop parameters are 0.01, 0.0125, 0.015, 0.0175, 0.02, 0.0225, 0.025, 0.03 and **Fig. 7** plots the covariance matrix for the 3 variables in **Fig. 6**. The experimental results are analyzed as follows:

- For the link density variance, With the increase of the stop parameters, the link variance increases rapidly at first, then tend to increase slightly.
- For the number of subregions, the size of the stop parameter is proportional to the number of subregions, and the growth speed is nearly unchanged.

Based on the above analysis, we chose 0.02 as the optimal value of the stop parameter. **Fig. 8** shows the relationship between the average NS of the three partition results and the stopping parameters for the Iso-MB. It's clear that 0.02 is the optimal value of the stop parameter in the current scenario. In practical applications, the optimal value should be selected according to the road and traffic conditions, actual partition requirements of the road networks, etc. The parameter obtained according to the above standards can be applied to different scenarios.
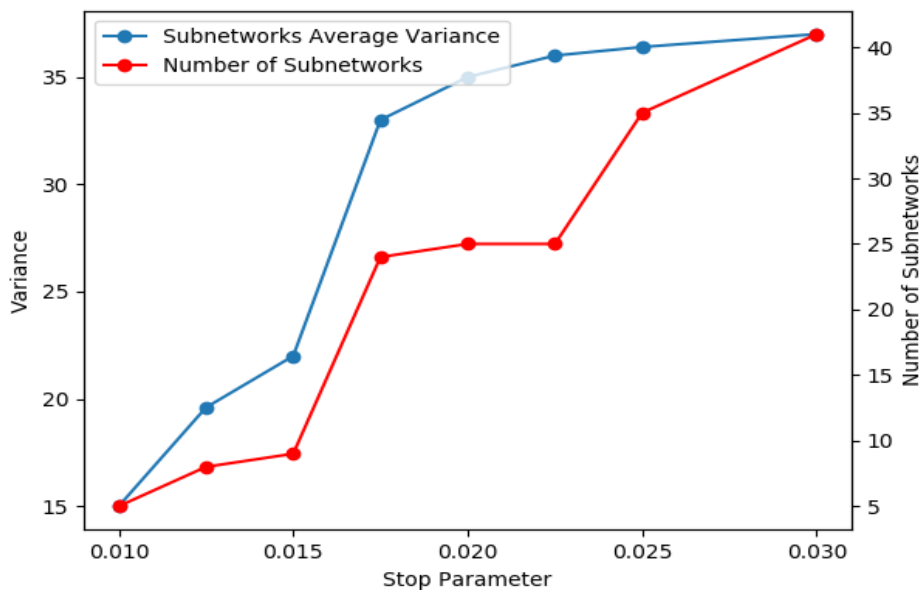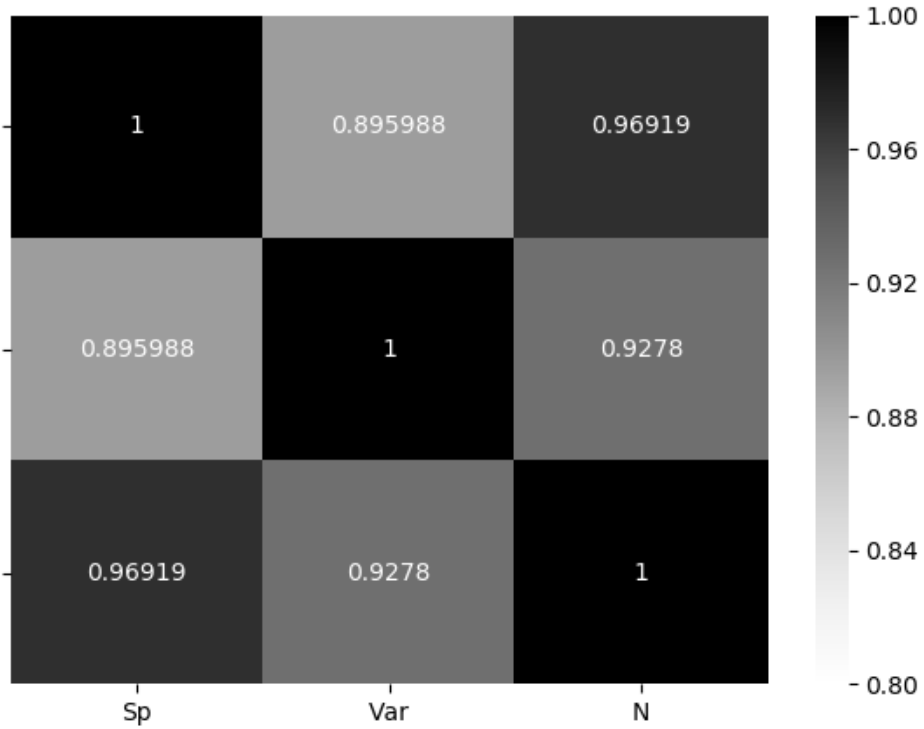

Fig. 6. Sensitive Analysis of the Stop Parameter

Fig. 7. The covariance matrix of the Stop Parameter (Sp), the subnetworks average Variance (Var) and the Number of subnetworks (N)
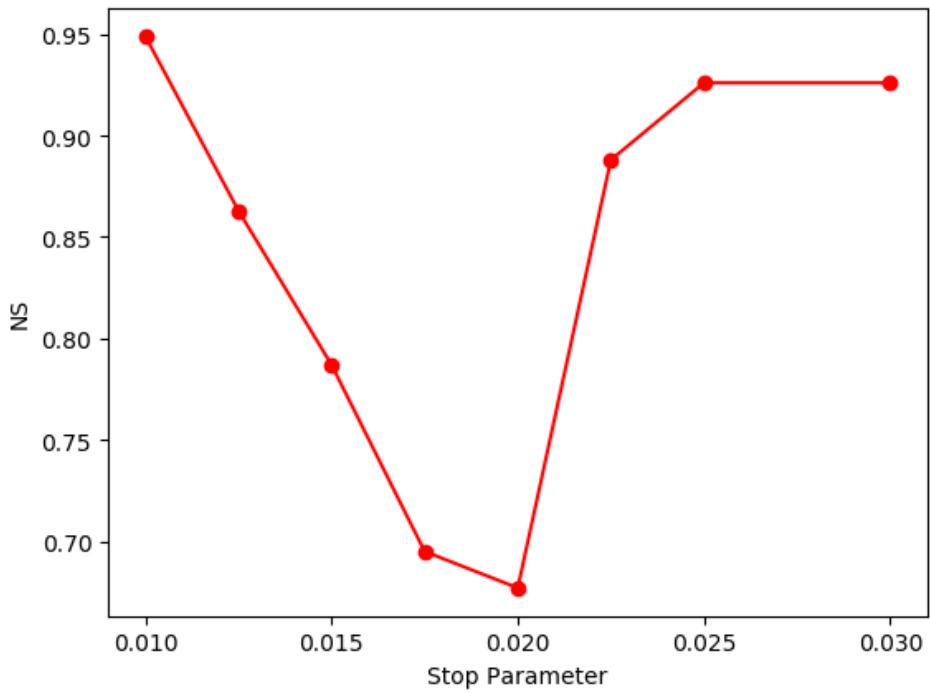

Fig. 8. The average NS of the 3 partition results corresponding to different stop parameters using Iso-MB

3.4 Stability Analysis to Data Missing

As discussed, based on the excellent segmentation stability of the isoperimetric algorithm, the Iso-MB algorithm can eliminate the adverse effects of missing data to some extent. In this section, we measured the similarity of the partitioned results with the missing data and the original intact data based on the criterion for calculating the similarity of the two partitioned results with and without missing data in [21].

To imitate the realistic missing traffic data (refer to [16, 17]), a certain percentage of data was randomly selected. The similarity of the partitioned results using the data with different data missing rate compared to original partitioned results are shown in **Fig. 9**. It illustrates that as the proportion of data loss increases, the similarity becomes smaller, which means that the partitioned results are getting worse. If the minimum acceptable similarity is set to 0.5, at least 40% of the data is required to guarantee satisfactory results. Taking into account the missing proportion of traffic data in reality, if appropriate data imputation technology is used before data input, this requirement can be easily met, which demonstrates the excellence of the proposed algorithm in practicality and effectiveness.
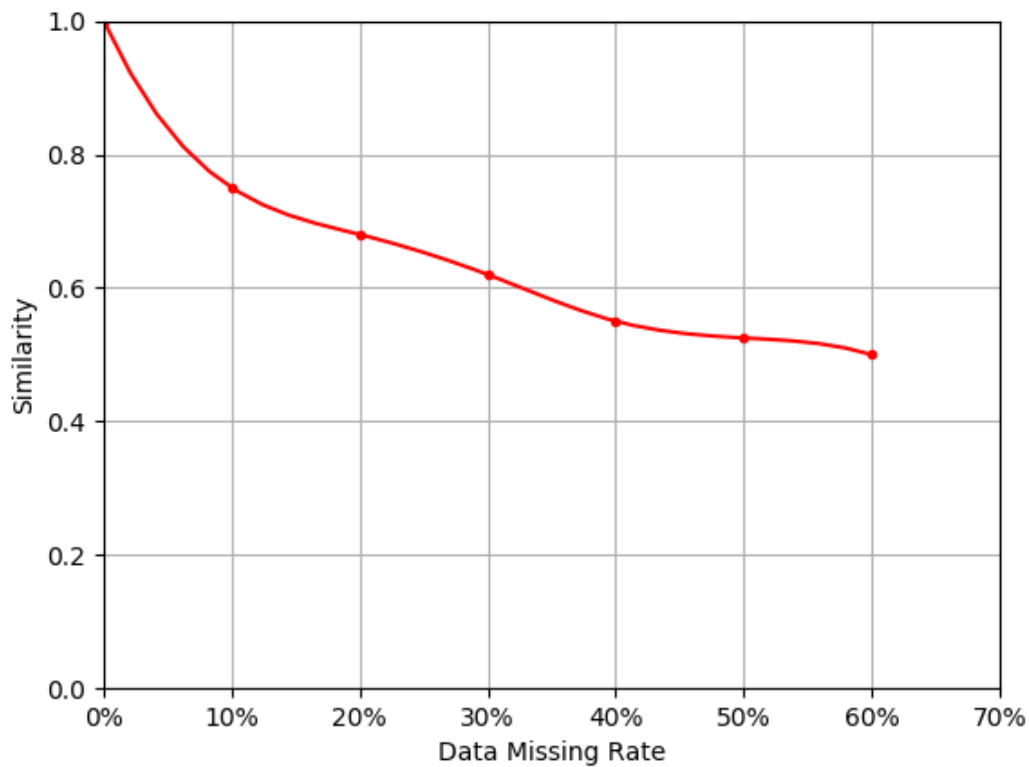

Fig. 9. The Similarity of Different Data Missing Rate

4.　Conclusion

In this paper, to facilitate the implementation of the coordinated control and management strategy of traffic, a three-step methodology called Iso-MB based on the isoperimetric algorithm was proposed to partition a heterogeneous road network into several homogeneous subnetworks. Compared with other partition algorithms, this algorithm mainly solves the

problem that it cannot run in the absence of some data, which is a common problem in reality. The superiority in computational complexity ensures it can satisfy the needs of real-time traffic road network partition. A series of numerical criteria and the MFD of each subnetwork indicated that the partitioned results of the methodology have achieved satisfactory results. Stability analysis experiments performed the effectiveness of the algorithm working with partial data. The algorithm, combines with proper MFD-based traffic coordination strategies, can effectively alleviate traffic congestion in cities.

Despite the above outstanding results, some shortcomings also exist in the algorithm, and there still requires further research and improvement in certain aspects. The calculation of the similarity between adjacent links has a great impact on the quality of the partitioned result, but further discussion about β was not proposed. In the step of Iterative merging, the setting of the lowest limit for the minimum subnetwork size was empirical. In the future research work, the calculation index between nodes will be further explored to improve. The calibration of the lowest limit is also needed in further study by looking at the network homogeneity and realistic condition concerning transportation management. In decree to progress general applicability of the algorithm, the interaction of multiple factors in different scenarios should also be considered. The current algorithm has not yet achieved complete automation of the overall process to obtain the optimal solution, and manual parameter selection and intervention are still called for. Attempting to formulate formulas through certain parameters and relying on convergence optimization methods to obtain the optimal solution is the focus of future work. Furthermore, with the continuous change of road network traffic conditions over time, dynamic adjustment of the stop parameter to ensure the optimal partitioned results is also worth studying in the future.

Acknowledgments

## References

[1] R. J. Walinchus, "REAL-TIME NETWORK DECOMPOSITION AND SUBNETWORK INTERFACING," *Highw. Res. Rec.*, no. 366, pp. 20-28, 1971.

[2] H. N. Yagoda, E. H. Principe, C. E. Vick and B. Leonard, "Subdivision of signal systems into control areas," *Traffic Engineering, Inst Traffic Engr*, vol. 43, no. 12, pp. 42-45, 1973.

[3] E. C.-P. Chang, "How to decide the interconnection of isolated traffic signals," in *Proc. of Presented at Winter Simul Conf.*, pp. 445–453, 1985. Article (CrossRef Link)

[4] R. H. Whitson, B. White, and C. J. Messer, *A study of system versus isolated control as developed on the mockingbird pilot Study*, Dallas, Texas Transportation Institute, 1973.

[5] L. T. Lin and H. J. Huang, "A linear model for determining coordination of two adjacent signalized intersections," *J. Model. Manag.*, vol. 4, no. 2, pp. 162-173, Jul. 2009. Article (CrossRef Link)

[6] Y. Bie, D. Wang, Q. Wei and D. Ma, "Development of Correlation Degree Model between Adjacent Signal Intersections for Subarea Partition," in *Proc. of ICCTP 2011: Towards Sustainable Transportation Systems*, 2011. Article (CrossRef Link)

[7] N. Geroliminis and C. F. Daganzo, "Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings," *Transp. Res. Pt. B-Methodol*, vol. 42, no. 9, pp. 759-770, Nov. 2008. Article (CrossRef Link)

[8]   N. Geroliminis and J. Sun, "Properties of a well-defined Macroscopic Fundamental Diagram for urban traffic," *Transp. Res. Pt. B-Methodol*, vol. 45, no. 3, pp. 605-617, Mar. 2011. Article (CrossRef Link)

[9]   Y. Ji and N. Geroliminis, "On the spatial partitioning of urban transportation networks," *Transp. Res. Pt. B-Methodol*, vol. 46, no. 10, pp. 1639-1656, Oct. 2012. Article (CrossRef Link)

[10] Y. Ma, Y. Chiu and X. Yang, "Urban traffic signal control network automatic partitioning using laplacian eigenvectors," in *Proc. of IEEE Conf Intell Transport Syst Proc ITSC, St. Louis, MO, USA*, pp. 1-5, 2009. Article (CrossRef Link)

[11] A. F. Lentzakis, R. Su and C. Wen, "Time-dependent partitioning of urban traffic network into homogeneous regions," in *Proc. of Int. Conf. Control Autom. Rob. Vis., ICARCV, Singapore, Singapore*, pp. 535-540, 2014. Article (CrossRef Link)

[12] Z. Zhou, S. Lin and Y. Xi, "A dynamic network partition method for heterogenous urban traffic networks," in *Proc. of IEEE Conf Intell Transport Syst Proc ITSC, Anchorage, AK, USA*, pp. 820-825, 2012. Article (CrossRef Link)

[13] B. Wang, L. Tao, C. Gao, D. Xia, Z. Rong, W. Wang and Z. Zhang, "Dividing Traffic Sub-areas Based on a Parallel K-Means Algorithm," in *Proc. of Lect. Notes Comput. Sci.*, pp. 127-137, 2014. Article (CrossRef Link)

[14] T. Anwar, C. Liu, H. L. Vu and C. Leckie, "Partitioning road networks using density peak graphs: Efficiency vs. accuracy," *Inf. Syst.*, vol. 64, pp. 22-40, Mar. 2017. Article (CrossRef Link)

[15] Z. Wang, K. Tan and P. Zhang, "Dynamic Partitioning Methods for Control Sub-area Based on Fuzzy Clustering Analysis," in *Proc. of Chin. Int. Auto. Conf.*, pp. 277-286, 2015.

[16] D. Li, J. Deogun, W. Spaulding and B. Shuart, "Towards missing data imputation: a study of fuzzy k-means clustering method," in *Proc. of 4th RSCTC 2004*, pp. 573-579, 2004. Article (CrossRef Link)

[17] J. Tang, G. Zhang, Y. Wang, H. Wang and F. Liu, "A hybrid approach to integrate fuzzy C-means based imputation method with genetic algorithm for missing traffic volume data estimation," *Transp. Res. Pt. C-Emerg. Technol.*, vol. 51, no. 51, pp. 29-40, Feb. 2015. Article (CrossRef Link)

[18] C. Lopez, P. Krishnakumari, L. Leclercq, N. Chiabaut and H. Van Lint, "Spatiotemporal partitioning of transportation network using travel time data," *Transp. Res. Record*, vol. 2623, no. 1, pp. 98-107, 2017. Article (CrossRef Link)

[19] M. Saeedmanesh and N. Geroliminis, "Optimization-Based Clustering of Traffic Networks Using Distinct Local Components," in *Proc. of Presented at IEEE Conf Intell Transport Syst Proc ITSC*, 2015 Article (CrossRef Link)

[20] M. Saeedmanesh and N. Geroliminis, "Clustering of heterogeneous networks with directional flows based on "Snake" similarities," *Transp. Res. Pt. B-Methodol*, vol. 91, pp. 250-269, 2016. Article (CrossRef Link)

[21] K. An, Y.-C. Chiu, X. Hu and X. Chen, "A network partitioning algorithmic approach for macroscopic fundamental diagram-based hierarchical traffic network management," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 4, pp. 1130-1139, 2017. Article (CrossRef Link)

[22] L. Grady and E. L. Schwartz, "Isoperimetric partitioning: A new algorithm for graph partitioning," *SIAM J. Sci. Comput.*, vol. 27, no. 6, pp. 1844-1866, 2006. Article (CrossRef Link)

[23] L. Grady and E. L. Schwartz, "Isoperimetric graph partitioning for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 3, pp. 469-475, 2006. Article (CrossRef Link)

[24] B. Mohar, "Isoperimetric numbers of graphs," *J. Comb. Theory Ser. B*, vol. 47, no. 3, pp. 274-291, 1989. Article (CrossRef Link)

[25] M. E. Newman, "Fast algorithm for detecting community structure in networks," *Phys. Rev. E*, vol. 69, no. 6, pp. 066133, 2004. Article (CrossRef Link)

[26] S. Uppoor and M. Fiore, "Large-scale urban vehicular mobility for networking research," in *Proc. of IEEE Veh. Netw. Conf., VNC*, pp. 62-69, 2011. Article (CrossRef Link)

[27] S. Uppoor, O. Trullolscruces, M. Fiore and J. M. Barceloordinas, "Generation and Analysis of a Large-Scale Urban Vehicular Mobility Dataset," *IEEE. Trans. Mob. Comput.*, vol. 13, no. 5, pp. 1061-1075, 2014. Article (CrossRef Link)

Haitao Xu is associate professor of Hangzhou Dianzi University. His research interests are in the areas of machine learning and data mining. As a member, he won the second prize of Zhejiang Science and Technology Award and the first prize of Zhejiang Higher scientific research achievement awards.
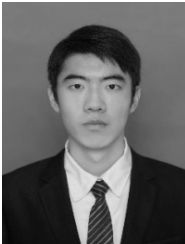
Weiguo Zhang is a postgraduate student in the School of Computer Science, Hangzhou Dianzi University. His research interests are clustering algorithm, graph partitioning and intelligent transportation.

Zuozhang Zhuo is a postgraduate student in the School of Computer Science, Hangzhou Dianzi University. His research interests are machine learning, deep learning and intelligent transportation.