

## Research Article

# Subdivision of Urban Traffic Area Based on the Combination of Static Zoning and Dynamic Zoning

Jing Luo  and Qingnian Zhang 

*School of Transportation, Wuhan University of Technology, Heping Road No. 1178, Wuchang District, Wuhan 430063, China*

Correspondence should be addressed to Jing Luo; 243252@whut.edu.cn and Qingnian Zhang; zqnwhut@163.com

Received 18 March 2021; Revised 27 June 2021; Accepted 8 July 2021; Published 16 July 2021

Academic Editor: Ya Jia

Copyright © 2021 Jing Luo and Qingnian Zhang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this paper, the traffic area of subzone division in urban road network is studied and a subzone division method based on the combination of static partition and dynamic partition is proposed. The static partition is carried out for the road network when the traffic flow is in a noncongested state, so as to provide the decision-making basis for the traffic green wave signal control strategy. At the same time, aiming at the road network when the traffic flow is congested, the dynamic partition is carried out on the basis of static partition to provide the decision-making basis for the traffic maximum flow signal control strategy. In view of the fact that it is difficult to determine the clustering center point during the initial division, this method proposes to determine the clustering center point according to the value of nodes of betweenness centrality. In order to solve the problem that it is difficult to collect traffic data, a method for estimating traffic flow density is proposed. In order to solve the problem of normalization of different probability distribution among various parameters, Mahalanobis distance is used as the fusion index of subzone division. Model verification shows that the method is feasible and effective.

## 1. Introduction

With the development of cities, the construction of urban road network becomes more and more complicated and finally turns into an extremely huge and complex network system. Because the traffic characteristics of urban road network are stochastic and dynamic [1], the complexity of signal control model calculation increases factorially with each additional intersection in the traffic signal control system. If the traffic guidance and control of the whole road network are unified, first, the signal data calculation system will be too slow or even unable to complete the calculation due to the excessive amount of data. This will reduce the operating efficiency of the whole control system and thus fail to achieve the purpose of traffic guidance and control in the control area according to the real-time traffic flow data [2]. Second, once one part of the control system fails, the whole system will break down. This will shut down the network coordination control at all intersections in the entire urban

road network, greatly reducing the safety and reliability of the traffic signal control system [3]. Third, the urban road network is so huge and complex that different traffic areas may be in different traffic conditions at the same time, which will bring trouble to the optimal timing of traffic signal control at intersections [4]. The best way to tackle the above problems is to divide the traffic area into subareas of the urban road network. In the study of subarea partition, the traditional method is static subarea division based on the historical traffic flow data of the road network. Once the traffic signal control subarea is formed, the number of subareas in the road network and the intersections contained in each subarea will be fixed for a long time. However, with the development of intelligent transportation, the static division method of traffic area subareas is no longer enough to meet the functional requirements of modern intelligent road network. Therefore, it is very necessary research work to find a method that can dynamically partition the urban road network in real time.

## 2. Related Work

*2.1. Existing Work.* With the development of intelligent transportation, there is some updated work on intelligent traffic signal control systems, such as [5–8]. These new research results bring to our control system a better signal timing schedule to effectively reduce traffic delays. With the in-depth research of traffic control system, the research of traffic control subarea division has also attracted the attention of scholars.

The concept of traffic signal control subarea was first proposed by American scholar Walinchus [9] in 1971. Walinchus regarded the significant change of traffic flow characteristics as the standard of subzone division and took the phase difference error, intersection saturation, and significant changes of road physical characteristics as the influencing factors of traffic subzone division. In the subsequent study of subdivision of traffic control, some scholars have specifically studied the various parameters that affect traffic control subpartition. For example, Yagoda et al. [10] analyzed the influence of traffic state, intersection spacing, intersection timing scheme, and other factors on the division of traffic area subareas. Chang [11] and Lin and Tsao [12, 13] studied the control index, threshold, and algorithm of dynamic subarea partitioning. Some scholars have studied the traffic signal control area division method based on a variety of indicators and parameters. For example, Li et al. [14] put forward a dynamic traffic control partition method based on BP (Backpropagation) neural network on the basis of comprehensively considering the three factors of traffic flow, intersection spacing, and cycle. Tian [15] put forward a dynamic subarea division method based on the improved Newman community rapid division by comprehensively considering the distance between adjacent intersections in road network, traffic flow, travel time, discrete characteristics of traffic flow, signal cycle, and traffic flow density of sections. Zhou [16] determined the correlation principle of subsection division from the signal cycle, traffic flow, and queuing ratio of the intersection and established the constraint condition model of dynamic subsection division. Zhao et al. [17] proposed a dynamic subarea division method based on key intersections by taking the period and distance principles as subarea division standard. According to the three principles of distance, flow, and period, Feng et al. [18] proposed a model of secondary molecular region delineation based on green wave coordinated control. Some other scholars have merely used a single index as a reference factor to study the traffic control area division method. Most of these single indexes are supplementary to common metrics such as distance, traffic, and cycles. For example, Xia et al. [19] used the optimal K-means algorithm of parallel clustering and distributed traffic control partitioning method to partition the road network according to the GPS tracks of a large number of taxis. Li et al. [20] proposed the method of spatial statistical analysis to study the automatic division of traffic subareas in urban road network. Ma et al. [2] proposed three automatic partition methods of traffic control network districts: bisection method, division according to mean value method, and division according to distance

method. Lu et al. [4] proposed to use dimensionality reduction processing and genetic algorithm to search for the optimal fast subpartition scheme, aiming at the problem shooting of dimension disaster that may be faced in the process of solving the optimal subpartition scheme. Bie et al. [3], starting from the overall perspective of traffic control system and taking saturation as an indicator, established the target set of subarea division, put forward the idea of establishing the correlation model and the division algorithm, and formed the framework system of subarea dynamic division from three levels: strategic, theoretical, and algorithmic level. Zhu et al. [21] proposed a subarea division method TSAD-HR (Traffic Subarea Division-Hot Region) based on hot spots according to the spatiotemporal trajectory of traffic flow. In these partitioning methods based on a variety of single or compound indicators, the compatibility of traffic control partitioning results to algorithms is rarely considered.

According to the different traffic flow conditions in the road network, the objectives of traffic control can be broadly divided into two kinds. First, when the traffic flow is saturated or supersaturated, the traffic control method takes the maximum total traffic flow of vehicles in the traffic area as the control target. At this time, the density of traffic flow is the main reference index of traffic signal control. Second, in the case of free flow, the traffic control method takes the minimum overall average travel time of vehicles in the traffic area as the control goal. Usually, the traffic signal green wave control method is used to achieve the control purpose when the traffic is in a free flow state. At this time, the maximum green wave bandwidth between any two adjacent intersections in the road network is the main reference index of traffic signal control.

For the two traffic control objectives, Ji and Geroliminis [22] studied road network partition in terms of saturated or supersaturated flow and believed that there was a macroscopic fundamental diagram (MFD) in urban road network, which correlated spatial average network flow and density. As long as traffic flow density in the MFD is at the right position, the average traffic flow in the network can be maximized. Shu et al. [23] designed a demand and supply balance model prediction controller based on the MFD multisubnet model, which optimized the mobility and throughput of road network flow by adjusting the input traffic of the subnet. Zong and Urbanik [24] analyzed the key factors affecting the green wave bandwidth, in which travel time is a key factor affecting the maximum green wave bandwidth. They proposed a traffic control method based on traffic control partition. This method can maximize the green wave bandwidth in the direction of peak traffic flow in each traffic control subarea. To sum up, in MFD, when the road network is in the state of saturated or supersaturated flow, the relationship between density and traffic flow is nonlinear. The nonlinear relationship between flow and density is used to adjust the traffic flow density in the section, so as to reduce the difference of traffic flow density between intersections in the road network and achieve the ideal effect of optimal timing of traffic signal control. When the traffic flow is in the state of free flow, the relationship between flow

and density can be approximately regarded as a linear positive correlation. In the state of free flow, the supply of traffic resources exceeds the demand, and the green wave control is adopted in traffic signal control. At this time, the average travel time of a vehicle is the focus of traffic signal control. In green wave control, green wave bandwidth is an important measure of signal control in road network. The green wave bandwidth is related to the distance between the intersections in the road network.

*2.2. Work in This Paper.* Most of the current traffic partitioning methods are merely aimed at the traffic congestion, how to provide traffic signal control algorithm for partitioning services to alleviate the traffic congestion, and how to provide effective partitioning service for traffic signal green wave control algorithm without considering free flow condition. In addition, in the choice of existing partitioning methods, the main methods include using the multiparameter comprehensive evaluation method to partition the traffic network or using a single index to partition the traffic network or using multi-index integration method to partition the traffic network. The multiparameter comprehensive evaluation method is not conducive to the homogeneity analysis of traffic areas. The zoning method based on a single index cannot fully reflect the differences among different intersections because the zoning index is too simple. Multi-index integration zoning method is more suitable for traffic zoning, but in the current research, most of these comprehensive methods only take into account the dimensional difference between different indicators, and the difference of their probability distribution is insufficient.

In addition, the existing partitioning methods basically divide the sections and intersections with high similarity into the same subarea by using cluster analysis or traversal search method according to the similarity between some traffic characteristics [15]. There are still some defects in these partitioning methods, such as insufficient consideration of traffic network topology in the partitioning process of these partitioning methods. Many parameters of traffic partition are difficult to be collected, such as queue length, average travel time, traffic flow density, etc. It is also difficult to determine many prior parameters required by traffic partitioning, such as the number of subareas, the size of subareas, and so on. To solve these problems, this paper proposes the following solutions:

- (1) This paper takes the static data such as road network topology structure and dynamic data such as traffic flow as easy to collect as indicators; this paper proposes a subarea division method combining static and dynamic data. When the traffic state in the road network is in free flow, the road network is partitioned only according to the static traffic data such as the topology of the road network, so as to improve the operating efficiency of the traffic system in free flow. At the same time, static data is used to solve the problem that some parameters are difficult to determine in the initial partition. When the traffic flow of any section of the road network is large and the

average traffic flow density reaches the threshold, the dynamic division will be carried out on the basis of static division to alleviate traffic congestion.

- (2) In terms of traffic data, this paper proposes an estimation method of traffic flow density in order to deal with the problem of high cost of actual traffic data collection.
- (3) Based on the advantage of Mahalanobis distance considering the correlation of multiple variables in the calculation of multi-index classification system, this paper proposes to use Mahalanobis distance as the fusion index of subarea division to fuse a variety of different partition parameters.
- (4) This paper proposes a method to determine the cluster center point (namely, the “core intersection”) based on road connectivity. This method is simple in calculation and has been verified effective by experiments. At the same time, based on the computing ability of the computing system, the size of the subarea is constrained, so that even when the traffic jam does not occur, a subarea will not be too large.

### 3. Methodology

In this paper, the reference indexes of traffic area partitioning are divided into two categories. The first is static data primarily based on road network structure, and the second is dynamic data related to real-time traffic flow. According to two different types of data, the subarea of traffic area is divided into two stages. When the traffic flow is in free flow state, the road network is statically partitioned according to the road network structure. When the traffic flow of any section of the road network is large and the traffic flow density reaches the threshold, the road network is dynamically partitioned according to the dynamic traffic data on the basis of static partitioning.

In this paper, in the partition method of traffic area, each partitioning is made into two subareas based on road network structure and real-time traffic flow information. If the partitioned area is too large, and the traffic area needs to be divided into three or more traffic areas, repeat the two stages of static partitioning and dynamic partitioning, so as to achieve the best partitioning effect in line with the requirements of the computing system. In addition, this paper takes the intersection as the object of subarea partitioning, so it is necessary to apply the traffic data in the section to the intersection. The method is to calculate the mean value of the traffic data of the adjacent intersections, which can be used to represent the traffic data of the intersection.

*3.1. Static Partitioning Phase.* When the traffic flow in the road network is small, the static partitioning method is used to partition the road network. Because at this time, the intersection traffic lights waiting time is the main factor causing traffic travel delays. Traffic green wave control is an effective traffic control method to solve this problem. The quality of traffic green wave control mainly depends on the average speed of vehicles and the network structure of the

road network. Among them, the average vehicle speed can be replaced by the free flow speed when the traffic flow is small. Free flow speed is closely related to road speed limits. When the traffic flow is small, the main factors affecting the travel time are static data. It is more appropriate to use static partition as traffic area partition method.

*3.1.1. Influencing Factors of Static Partition.* The reference factors of static partition mainly include the shortest path between arbitrary intersections; length of sections between adjacent intersections; and nodes of betweenness centrality (N-BC) at the intersection:

- (1) The shortest path refers to the distance between any adjacent intersections. The shortest path distance between any intersections reflects the position relationship of two intersections in the whole road network, and it is closely related to the average travel time, intersection connectivity, and other factors. There are many mature algorithms for calculating the shortest path distance. In this paper, when the scale of the partitioned road network is small, Dijkstra algorithm is used to realize the calculation of the shortest path between two intersections. When the scale of the partitioned road network is large,  $A^*$  algorithm is used to calculate the shortest path between two intersections.

The basic idea of Dijkstra algorithm comes from dynamic programming. Its advantage is that it can calculate the exact solution of the shortest path, but its disadvantage is that the calculation time is slow and it can only deal with the shortest path search problem with small network scale. It can handle the network size of about 20 intersections below. The essence of  $A^*$  algorithm is A Breadth First Search algorithm, which can quickly deal with the shortest path Search in large network size. The disadvantage is that it may converge to the local optimal solution rather than the global optimal solution. It can handle the network size between 20 and 100 intersections. If the network size is more than 100, Depth First Search algorithm should be considered for shortest path Search. Compared with  $A^*$  algorithm, it is easier to fall into the local optimal solution, but the Search speed is also faster.

- (2) The section length refers to the distance between adjacent intersections. The length of sections between adjacent intersections can be used as one of the criteria to measure the correlation between adjacent intersections. When the distance between two adjacent intersections is too long, the uncertainty of the arrival time will increase when the vehicles arrive at the next intersection, the coordination and control function between the two adjacent intersections will weaken significantly, and the relationship between the two adjacent intersections will show weak correlation. When the distance between two adjacent intersections is short, the vehicles will maintain good

continuity when they arrive at the next intersection, and their coordination and control function will be significantly enhanced. At this time, the relationship between two adjacent intersections is strongly correlated [15].

- (3) The N-BC of intersection points refers to the number of shortest paths through the intersection in the road network. In the entire urban road network, there are numerous key intersections, which are the traffic hubs of the entire urban road network. Its significance is mainly reflected in the connectivity of these intersections in the whole road network. According to the complex network theory and the topology of urban road network, the connectivity of each intersection can be measured by the N-BC at each intersection. In the topology of road network, N-BC can be used to indicate the importance of each intersection in the whole road network. The larger the N-BC at the intersection, the stronger the connectivity of the intersection, the larger the traffic flow through the intersection, and the more prone it is to congestion. Therefore, the core intersections in road network are identified by the numbers of N-BC and can well reflect the actual traffic flow status of the road network.

When calculating the N-BC, it is necessary to calculate the N-BC at each intersection according to the road network data and topology structure and comprehensively consider the number of lanes of each road. The calculation method of the N-BC: the number of lanes on the narrowest section of the shortest path between any two intersections in the road network is taken as the measure value of the N-BC that the path is attached to the intersection. And the intersection cannot be the start and end of the shortest path. The sum of the measure values of the N-BC at each intersection of all the shortest paths in the road network is the total of the N-BC at the intersection. The calculation formula of N-BC is as follows:

$$C_i^b = \frac{1}{(N-1)(N-2)} \sum_{j,k} \frac{n_{jk}(i) \times M_{jk(i)}}{n_{jk}}, \quad (1)$$

where  $n_{jk}$  represents the number of shortest paths between two intersections  $j$  and  $k$  and  $n_{jk}(i)$  represents the number of shortest paths through the intersection  $i$  between two intersections  $j$  and  $k$ . And  $i, j$ , and  $k$  are different intersections, respectively;  $M_{jk(i)}$  stands for the average number of lanes of the narrowest path in the number of shortest paths  $n_{jk}(i)$ ;  $N$  represents the total number of intersections in the network.

*3.1.2. Static Partitioning Methods.* The static division method of traffic area subarea is as follows:

- (1) Find the vertex intersection and make a preliminary subdivision. From the collected static information of the road network, then find the intersection with the maximum and minimum longitude values. Find the intersection of the maximum and minimum

latitudes. When the road network is a straight road, the maximum and minimum longitude values and the maximum and minimum latitude values will be concentrated at the two intersections. When the road network is approximately a triangle, the maximum and minimum longitude values and the maximum and minimum latitude values will be concentrated at the three intersections. In other cases, the maximum and minimum longitude values and the maximum and minimum latitude values are assigned to four different intersections. Then, Dijkstra algorithm, or  $A^*$  algorithm, is used to calculate the shortest path distance between two pairs of each intersection described above. And take the two intersections with the shortest path distance as the vertex intersections in the whole road network. Calculate the shortest path distance from each intersection in the road network to the two vertex intersections. According to the size of the shortest path distance, the intersection is assigned to the area of the nearest vertex intersection. In this way, the whole road network is initially divided into two subareas around the two vertex intersections, known as the “preliminary partition” stage.

- (2) The N-BC of each intersection in each subarea is calculated in each subarea and the core intersection is determined. In each subarea, N-BC at each intersection is calculated through equation (1). The premise is to calculate the number of shortest paths between each intersection  $n_{jk}$ . Dijkstra algorithm or  $A^*$  algorithm is used to achieve this. After finding the two intersections with the most N-BC in each of the two subareas, they can be used as the core intersections of the two subareas to replace the vertex intersections.
- (3) Calculate Mahalanobis distance  $MA_{in}$ , and statically divide subareas. The shortest path distance  $S_{in}$  from each intersection to the two core intersections is calculated. Connection section of each intersection is calculated as the average length  $\bar{L}_n$ . According to the calculation results, two two-dimensional vector matrices were established as  $\vec{A}_{1n} = \begin{bmatrix} \bar{L}_n \\ S_{1n} \end{bmatrix}$ ,  $\vec{A}_{2n} = \begin{bmatrix} \bar{L}_n \\ S_{2n} \end{bmatrix}$ , where  $S_{in}$  is obtained based on the length of each section in the shortest path.  $\bar{L}_n$  stands for the average length of each section connecting intersection  $n$ . Because the shortest path distance from the core intersection to itself is  $S_1 = 0, S_2 = 0$ . The two-dimension vectors of two core intersections, respectively, are  $\vec{A}_1 = \begin{bmatrix} \bar{L}_1 \\ 0 \end{bmatrix}$ ,  $\vec{A}_2 = \begin{bmatrix} \bar{L}_2 \\ 0 \end{bmatrix}$ . Then the Mahalanobis distance from each intersection to the two core intersections is calculated by two-dimensional vector. Mahalanobis distance is a statistical measure method proposed by Indian statistician PC Mahalanobis. Mahalanobis distance

is widely used in the field of multivariate statistics [25]. It represents the distance between a point and a distribution. This is an effective method to calculate the similarity of two unknown sample sets. Its advantage is that when calculating the multi-index classification system, it takes into account the correlation between multiple variables; even if there is a certain correlation between different indicators, it can still get accurate classification metrics. Different traffic parameters often have different degrees of correlation. For example, there is a certain correlation between the distance between the two intersections and the traffic flow between the two intersections. When a traffic accident occurs at an intersection and leads to traffic congestion, the traffic flow of the intersection closer to it will also be affected to a certain extent, and this impact will decrease with the increase of the distance between intersections. Therefore, it is proposed in this paper that it is more appropriate to use Mahalanobis distance as a multi-index classification measure in the traffic zoning. The calculation formula of Mahalanobis distance is

$$MA_{1n} = \sqrt{\left(\vec{A}_{1n} - \vec{A}_1\right)^T \sum_1^{-1} \left(\vec{A}_{1n} - \vec{A}_1\right)}, \tag{2}$$

$$MA_{2n} = \sqrt{\left(\vec{A}_{2n} - \vec{A}_2\right)^T \sum_2^{-1} \left(\vec{A}_{2n} - \vec{A}_2\right)},$$

where  $MA_{1n}$  and  $MA_{2n}$  are Mahalanobis distances from each intersection to the two core intersections in the whole traffic network;  $\vec{A}_{1n}$  and  $\vec{A}_{2n}$ , respectively, stand for two-dimensional vector of intersection  $n$  in traffic network to the two core intersections;  $\vec{A}_1$  and  $\vec{A}_2$  are two-dimensional vectors of two core intersections, respectively;  $\sum_1^{-1}$  and  $\sum_2^{-1}$  are two-dimensional vector matrix of, respectively; and  $\vec{A}_{1n}$  and  $\vec{A}_{2n}$  inverse matrix of covariance matrix.

- (4) In the road network, the  $MA_{1n}$  Mahalanobis distance of each intersection is compared with the  $MA_{2n}$  Mahalanobis distance of each intersection, and then the intersection is classified into the core intersection with a shorter distance. Calculate the N-BC in the new subarea, and determine the core intersection (the intersection with the highest N-BC). Calculate the Mahalanobis distance from each intersection in the road network to the two new core intersections. Then according to the size of Mahalanobis distance, the second static subdivision was carried out. This iterative cycle is carried out until the core intersections are no longer changed and the partition results are constant and stable. At this time, the static partition of the road network will finally be completed.

- (5) When the road network is in free flow state, perform step 6. When the traffic flow of a certain section of the road network is large and the traffic flow density of the section reaches the set threshold value  $P_f$ , the dynamic division stage will be entered.
- (6) When the number of intersections in the divided neutron region exceeds the maximum capacity  $C$  of the calculation system, repeat steps 1–6 for the overlarge subarea. This iteration is carried out until the number of intersections in the subarea meets the calculation requirements of the computing system. At this point, complete the subdivision of the road network.

In traffic partition, regardless of the performance of the traffic signal control algorithm and the computer computing capacity and under the condition of the partition results meeting the partition homogeneity, each subarea should be partitioned as big as possible, with subarea number as less as possible, so as to facilitate the traffic signal control algorithm for global traffic signal overall planning. However, the maximum number of subarea intersections is restricted by the performance of the traffic signal control algorithm and the computing capacity of the computer, so the size of each subarea should not be too large. Therefore, after the comprehensive evaluation based on the performance of the traffic signal control algorithm and the computing capacity of the computer, there must be a maximum of the size of the subarea. Therefore, in the case of traffic partition, different traffic signal control algorithms and computer with different performance differ in the maximum capacity  $C$  of the intersection in the neutron zone, which should be calibrated according to the actual situation. Based on the strong real-time demand of the traffic system, the size of  $C$  should enable the computer to complete the traffic signal timing work in the entire traffic subarea within 0.1 seconds through the traffic signal control algorithm.

**3.2. Dynamic Partitioning Phase.** When the traffic flow in the road network is large, the dynamic partitioning of the road network can be carried out again on the basis of static partitioning, so as to adapt to the optimization control of the overall maximum traffic flow in the traffic signal control when the traffic flow is large.

**3.2.1. Influencing Factors of Dynamic Partition.** The traffic flow density is the key factor affecting the traffic area partition under dynamic partition, and the traffic area partition is still affected by the compact structure of the road network. Therefore, traffic flow density and shortest path distance are the key factors affecting dynamic partitioning.

Traffic flow density refers to the vehicle numbers for every thousand metres on the section. Traffic flow density is the main factor that affects the traffic flow in the section. Traffic flow changes according to density changes in the MFD. As shown in Figure 1, when the traffic flow density  $P$  is close to the saturation flow density  $P_\mu$ , the growth rate of

traffic flow volume  $Q$  will gradually slow down. When traffic flow density  $P$  is in a saturated density  $P_\mu$ , the traffic flow volume comes to the peak value  $\gamma$  in theory. When the traffic flow density  $P$  is in the supersaturated flow density  $P_\nu$ , the growth rate of the traffic flow volume  $Q$  will decrease with the increase of the density, and finally the traffic flow  $Q$  will become zero when it reaches the blocking flow density  $P_{\max}$ . Therefore, as long as the appropriate maximum flow density is found in the MFD and the traffic flow density in the section is maintained at the maximum flow density, the maximum traffic flow can be easily obtained. At this time, for traffic signal control, if the traffic flow density difference of each section of the subarea is small, it will be easier to achieve the effect of traffic flow regulation. At the same time, because the same traffic control strategy is implemented in the same subarea, the dynamic partitioning of the road network also enables the green wave control and the maximum flow control to be carried out in different subareas at the same time.

In actual traffic environment, traffic flow density data is difficult to be obtained directly. In order to solve this problem, this paper adopts a method to estimate the traffic flow density in the section by the data of the blocking flow density, the real-time traffic flow, the green signal ratio, the free flow average speed, and so on. The specific methods are as follows:

$$\lambda_t = \frac{g_t}{c_t}, \quad (3)$$

where  $\lambda_t$  is the green signal ratio at time period  $t$ ;  $g_t$  is the effective green light duration at time period  $t$ ; and  $c_t$  is the traffic signal interval at a certain time period  $t$ .

The relationship between traffic flow and density is usually shown as follows:

$$Q_t = \lambda_t \times \bar{V}_t \times P_t. \quad (4)$$

Change to

$$\bar{V}_t = \frac{Q_t}{\lambda_t \times P_t}, \quad (5)$$

where  $Q_t$  stands for traffic flow volume at  $t$  time period;  $\bar{V}_t$  for the average speed of section vehicles at  $t$  time period; and  $P_t$  for the density for every lane of traffic flow at  $t$  time period ( $P_t$  abbreviated as “traffic flow density on the sections”).

The relationship between traffic flow density and average vehicle speed can be approximated as shown in Figure 2. When the traffic flow density gradually increases, the average vehicle speed will gradually decrease, until the density reaches the blocking flow density, and the average vehicle speed will drop to zero. Therefore, the change relationship between velocity and density can be approximated by the following formula:

$$\frac{\bar{V}_t}{V_{\max}} = 1 - \frac{P_t}{P_{\max}}. \quad (6)$$

It can be changed into

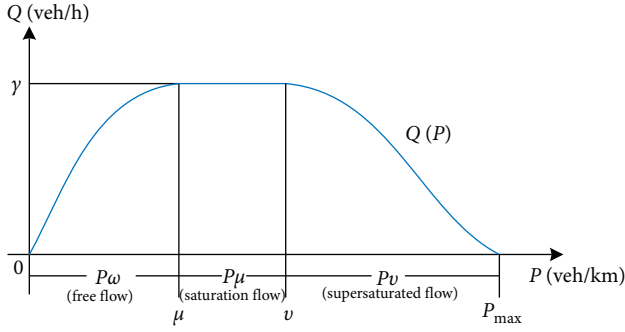


FIGURE 1: MFD similar to a trapezoid.

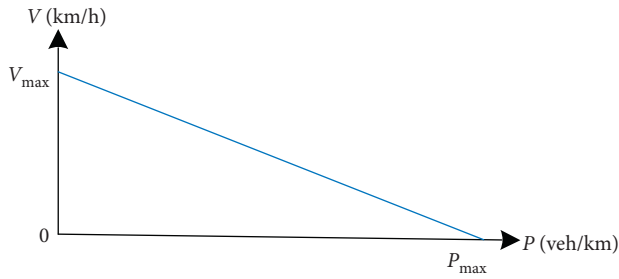


FIGURE 2: Relationship between vehicle average speed and traffic flow density.

$$\bar{V}_t = V_{\max} \left( 1 - \frac{P_t}{P_{\max}} \right), \quad (7)$$

where  $V_{\max}$  is the free flow speed and  $P_{\max}$  is the blocked flow density.

From equations (5) and (7), we obtain

$$\frac{Q_t}{\lambda_t \times \bar{P}_t} = V_{\max} \left( 1 - \frac{P_t}{P_{\max}} \right). \quad (8)$$

The quadratic function can be obtained after the variation

$$P_t^2 - P_{\max} \times P_t + \frac{P_{\max} \times Q_t}{\lambda_t \times V_{\max}} = 0. \quad (9)$$

According to the quadratic equation of one variable to find the root formula,

$$P_t = \frac{P_{\max} \pm \sqrt{P_{\max}^2 - ((4 \times P_{\max} \times Q_t) / (\lambda_t \times V_{\max}))}}{2}. \quad (10)$$

Usually, the blocking flow density is  $P_{\max} = 124$  veh/km. According to the speed limit standard of urban road, the average speed of traffic free flow can be set as  $V_{\max} = 60$  km/h. Real-time traffic flow  $Q_t$  and real-time green ratio  $\lambda_t$  are obtained in real time through intelligent transportation systems such as V2X. In this way, the traffic flow density within the section in time period  $t$  can be estimated in real time.

### 3.2.2. Dynamic Partitioning Methods.

- (1) Calculate Mahalanobis distance  $MD_{in}$ . According to the results of static partition, the shortest path distance  $S_{in}$  between all intersections in the road network and the two core intersections is calculated. According to the collected dynamic traffic data, the average  $P_n$  of the traffic flow density is calculated of all its connection section of the intersection ( $P_n$ , abbreviated as "traffic flow density at intersections"). According to the shortest path distance from the intersections to the two core intersections and the average length of all its connection section of the intersection two two-dimensional vector matrices are set up, marked as  $\vec{D}_{1n} = \begin{bmatrix} P_n \\ S_{2n} \end{bmatrix}$ ,  $\vec{D}_{2n} = \begin{bmatrix} P_n \\ S_{2n} \end{bmatrix}$ . The two-dimensional vectors of core intersection in the two subareas are  $\vec{D}_1 = \begin{bmatrix} P_n \\ 0 \end{bmatrix}$ ,  $\vec{D}_2 = \begin{bmatrix} P_n \\ 0 \end{bmatrix}$ , respectively. The Mahalanobis distance  $MD_{in}$  from each intersection to two core intersections in the road network is calculated by two-dimensional vector. The calculation formula of Mahalanobis distance is

$$MD_{1n} = \sqrt{\left( \vec{D}_{1n} - \vec{D}_1 \right)^T \sum_1^{-1} \left( \vec{D}_{1n} - \vec{D}_1 \right)}, \quad (11)$$

$$MD_{2n} = \sqrt{\left( \vec{D}_{2n} - \vec{D}_2 \right)^T \sum_2^{-1} \left( \vec{D}_{2n} - \vec{D}_2 \right)}.$$

- (2) The Mahalanobis distances  $MD_{1n}$  and  $MD_{2n}$  of each intersection in the road network to two core intersections were compared, and then the intersection was classified into the closer core intersections. This method divides the whole road network into two subareas again.
- (3) Calculate N-BC of each intersection in the new subarea, take the intersection with the most N-BC as the core intersection of the subarea, and calculate the Mahalanobis distance from each intersection to the new core intersection. According to the magnitude of Mahalanobis distance, the road network is subdivided again. This iterative cycle is carried out until the partition results are stable and the core intersections are no longer changed. At this time, the dynamic partition of the road network will be completed.
- (4) For the subareas with too many intersections in the partition results, all the steps of the static and dynamic division stages are repeated. This iteration is carried out until the number of intersections in the subarea meets the calculation requirements of the computing system. At this point, the subdivision of the road network is completed.

Figure 3 shows the traffic area subarea partition flowchart.

## 4. Model Validation

4.1. *Validation of the Data.* The data are validated as follows:

- (1) Road network: take Heping Avenue and Youyi Avenue in Qingshan District of Wuhan City as the main axes and 40 intersections in the traffic area of Garden Road-Beiyangqiao West Road-Renhe Road-Jianshe Third Road-Fushun Street as the edge of the model to validate the traffic network. The intersection distribution is shown in Figure 4. For static data such as section distance, see Appendix 1 in Supplementary Material.
- (2) Traffic flow data: collect traffic volume every 10 minutes from 7:20 to 9:20:00. Its unit is vehicle per hour (veh/h). See Appendix 2 in Supplementary Material for the traffic flow data at each intersection.
- (3) Subdivision frequency: set to partition every 10 minutes.
- (4) The maximum number of intersections in the area are set to  $C=20$ .
- (5) The threshold for triggering dynamic partition is set as  $P_f \geq 40$  veh/km.

In practical application, the division frequency of the subarea and the number of intersections that can be accommodated in the subarea should be determined according to the traffic condition of the area and the capacity of the calculation system.

4.2. *Verification Process.* This includes the following:

- (1) According to the road network structure shown in Figure 4, find the intersection with the maximum and minimum longitude values. Find the intersection of the maximum and minimum latitudes. The intersections with maximum and minimum longitude and latitude are  $a_1$ ,  $a_5$ ,  $a_{35}$ , and  $a_{40}$ . After calculation, the distance from  $a_5$  to  $a_{35}$  is 3968 metres, and its shortest path distance is the largest. Therefore,  $a_5$  and  $a_{35}$  intersections are the top intersections of the road network. According to the distance from each intersection to two vertices  $a_5$  and  $a_{35}$ , preliminary partitioning is conducted. The results are shown in Table 1.
- (2) The first subarea partition process is divided into static and dynamic stages. According to the preliminary partition results, the first static partition is conducted, and after iterations, the static partition result is stable. The results are shown in Table 2. According to the collected traffic flow data at each time period,  $P_t$  at 7:30, 7:40, 7:50, and 8:00 is greater than the set threshold  $P_f \geq 40$  veh/km. Therefore, the first dynamic partition is needed on the basis of the first static partition, and the dynamic partition result is stable after iterations. The results

are shown in Table 3. In other periods, due to the traffic flow density, there is no threshold  $P_f$  triggered, so dynamic partitioning is not required.

- (3) As can be seen from Tables 2 and 3, in the first partition results, subareas with intersections greater than 20 are stored in the partition results of each period. Therefore, based on the results of the first partition, subareas with intersections greater than 20 should be partitioned for the second time. In the second partition, dynamic partitioning is still needed for the four periods of 7:30, 7:40, 7:50, and 8:00, while static partitioning is only needed for the other periods. The zoning process is the same as that of the first partition. The results of the second partition are shown in Table 4.
- (4) As can be seen from Table 4, the road network is divided into three traffic subareas in each period of time, and there are no more than 20 intersections in each traffic subarea. The partition results are stable, and the division of traffic subareas is completed. The partition results are shown in Figure 5.

4.3. *Performance Optimization Analysis.* In this part, the partitioning method proposed in this paper is analyzed from four aspects: standard deviation, optimization of calculation time, partitioning method, and ablation experiment.

4.3.1. *Standard Deviation.* Variance is the most commonly used index to measure the difference degree of statistical data which has been used by many experts [22, 26–28]. Standard deviation is the arithmetical square root of the variance. Standard deviation reflects the degree of dispersion of a data set. Two sets of data with the same mean do not necessarily have the same standard deviation. A large standard deviation represents a large difference between most values and their mean; a small standard deviation means that these values are closer to the mean. When traffic area is partitioned, we need to minimize the difference of section length in the subarea after static partition, so as to adapt to green wave control. The difference of traffic flow density in the section of the subareas after dynamic partition is minimized as far as possible to adapt to the maximum traffic flow control. Here we use standard deviation to analyze the above partition results. The standard deviation of section length in this paper represents the compactness of the road network. The standard deviation of traffic flow density represents the congestion difference. Calculating formula of the section length standard deviation and section traffic flow density standard deviation before subdivision:

$$\mu = \sqrt{\frac{\sum_{n=1}^m (L_n - \bar{L})^2}{m}}, \quad (12)$$

$$\rho = \sqrt{\frac{\sum_{n=1}^m (P_t - \bar{P})^2}{m}}. \quad (13)$$

In equations (12) and (13),  $\mu$  is the standard deviation of the length of all sections in the road network before



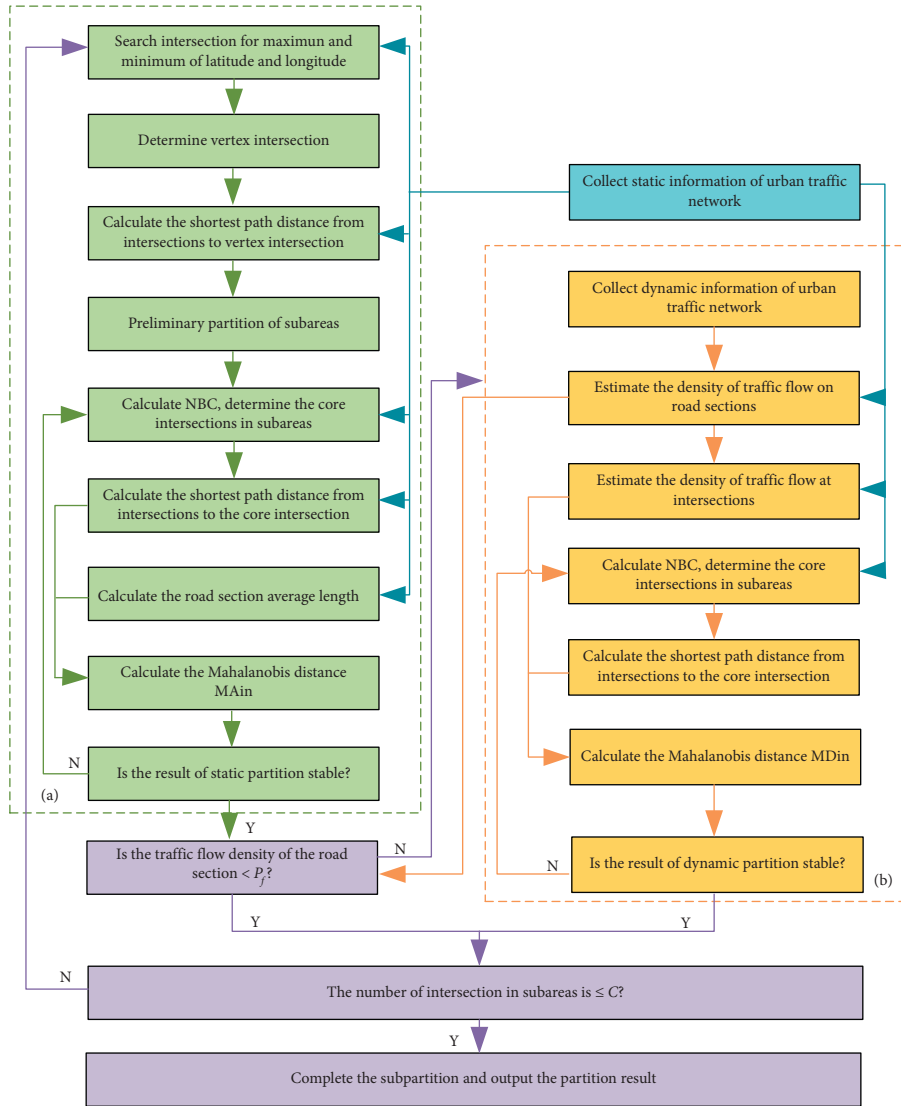


FIGURE 3: Traffic area subarea partition flowchart: (a) static stage; (b) dynamic stage.

partition.  $\rho$  is the standard deviation of the traffic flow density of all sections in the road network before subdivision.  $L_n$  and  $P_i$  are, respectively, the length of the section connected by the  $N$ th intersection and the traffic flow density of the section.  $\bar{L}$  and  $\bar{P}$  stand for the average length of all sections and mean value of traffic flow density in the network, respectively;  $m$  is the total number of intersections in the road network. The standard deviation of section length and standard deviation of traffic flow density in each subarea after partitioning can also be figured out by equations (12) and (13), respectively.

After partition, the standard deviation of section length and standard deviation of section traffic flow density are calculated as follows:

$$\bar{\mu} = \frac{\sum_{i=1}^n m_i \times \mu_i}{m}, \quad (14)$$

$$\bar{\rho} = \frac{\sum_{i=1}^n m_i \times \rho_i}{m}. \quad (15)$$

In equations (14) and (15),  $\bar{\mu}$  refers to, after partition, the length standard deviation of all sections in the road network;  $\bar{\rho}$  refers to, after partition, traffic flow density standard deviation of all sections in the road network;  $m_i$  is the number of intersections in the  $I$  subarea after partition;  $\mu_i$  is the standard deviation of section length in the  $I$  subarea after partition; and  $\rho_i$  is the standard deviation of section traffic flow density in the  $I$  subarea after partition. According to the static partition result, the standard deviation of section length before and after partition is calculated, as shown in Table 5.

According to the results of dynamic partition, the standard deviations of traffic flow density before and after partition were calculated, as shown in Table 6.

It can be seen from Tables 5 and 6 that the standard deviation of the section length or the section traffic flow density in the subarea at the back of the partition is increased compared with the previous partition. This increase is due to the fact that there are always some singular intersections in the road network, and the number of intersections sharing

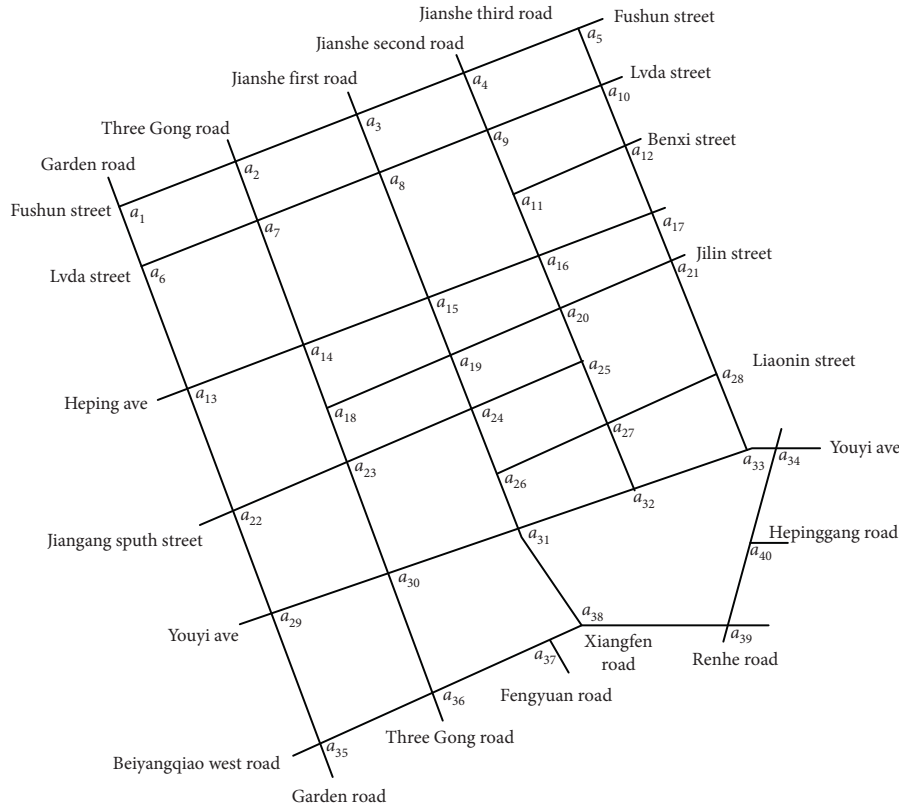


FIGURE 4: Model validation of road network. Note: a stands for intersection.

TABLE 1: Preliminary partition results.

Vertex intersection	Number of intersections	Intersection number
5	26	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 17, 19, 20, 21, 24, 25, 27, 28, 32, 33, 34
35	14	13, 18, 22, 23, 26, 29, 30, 31, 35, 36, 37, 38, 39, 40

TABLE 2: Static partition results of the first subzone.

Core intersection	Number of intersections	Intersection number
16	24	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 25, 27, 28
31	16	22, 23, 24, 26, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40

TABLE 3: Dynamic partition results of the first subzone.

Time period	Core intersection	Number of intersections	Intersection number
7:30	16	21	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 20, 21, 25
	31	19	19, 22, 23, 24, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40
7:40	16	22	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 25
	31	18	22, 23, 24, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40
7:50	16	22	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 25
	31	18	22, 23, 24, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40
8:00	16	22	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 25
	31	18	22, 23, 24, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40

these singular values in the subarea after partition decreases. The rise is inevitable because the structure of the road network is immutable. However, the standard deviation of

the whole road network section length and the section traffic flow density decreases after the partition compared with the previous partition. This indicates that the whole road

TABLE 4: Partition results of the second subzone.

Time period	Core intersection	Number of intersections	Intersection number
7:30	9	17	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 16, 17, 20, 21, 25
	14	4	13, 14, 15, 18,
	31	19	19, 22, 23, 24, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40
7:40	9	17	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 16, 17, 20, 21, 25
	14	5	13, 14, 15, 18, 19,
	31	18	22, 23, 24, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40
7:50	9	17	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 16, 17, 20, 21, 25
	14	5	13, 14, 15, 18, 19
	31	18	22, 23, 24, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40
8:00	9	18	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 16, 17, 19, 20, 21, 25
	14	4	13, 14, 15, 18
	31	18	22, 23, 24, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40
8:10-9:20	14	7	1, 2, 6, 7, 13, 14, 18
	16	17	3, 4, 5, 8, 9, 10, 11, 12, 15, 16, 17, 19, 20, 21, 25, 27, 28
	31	16	22, 23, 24, 26, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40

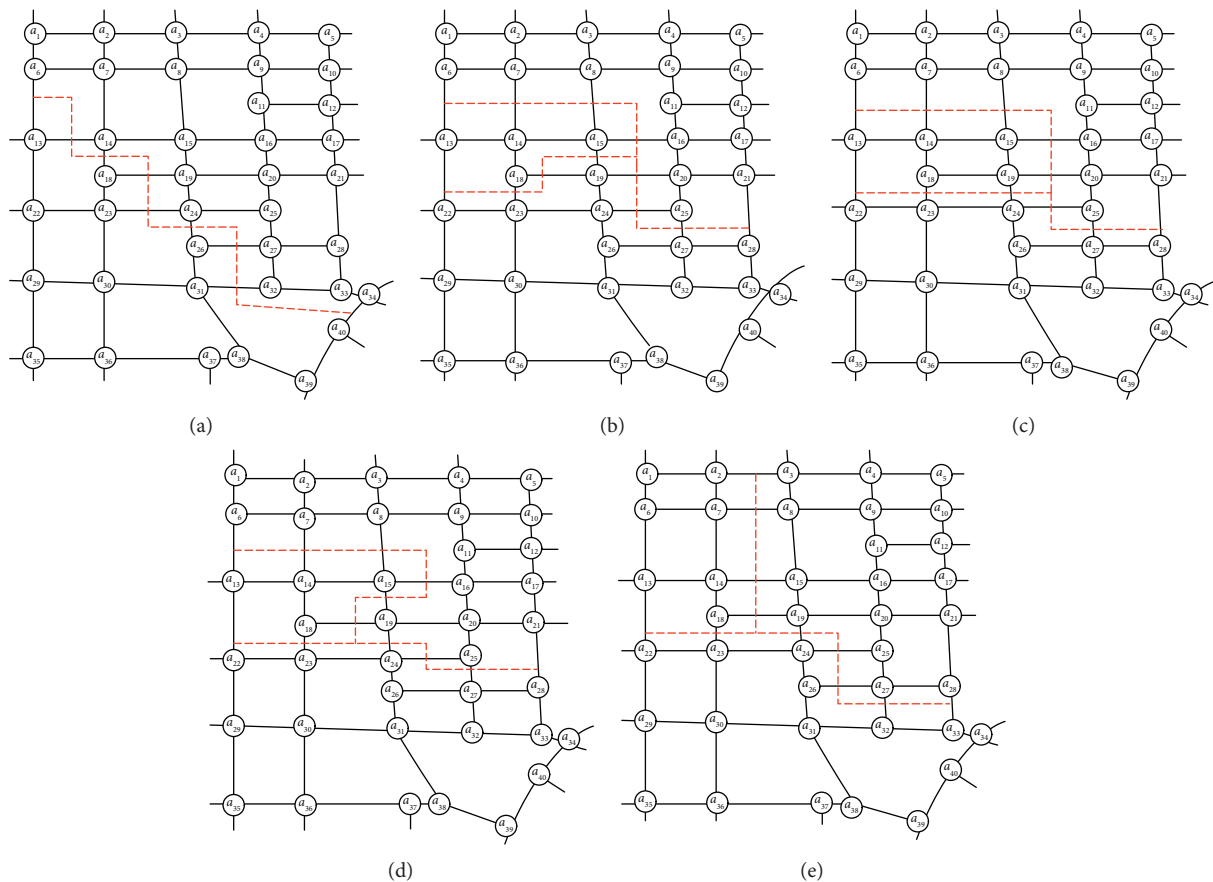


FIGURE 5: Traffic area subarea partition result. (a) Results of preliminary partition of subareas. (b) 7:30 time period. (c) 7:40 and 7:50 time periods. (d) 8:00 time period. (e) 8:10-9:20 time period.

network after partition can better adapt to different traffic control systems compared with previous partition .

4.3.2. Optimization of Calculation Time. The problem of integrated traffic signal control in a large-scale road network is NP-hard. In most cases, heuristic search algorithm will be

used to achieve the optimal traffic signal control timing results. Therefore, the size of the solution vector space will directly determine the computational efficiency of the signal control system. In this experiment, the size of the solution vector space before partition is  $(T \times \lambda)N$ , where  $T$  represents the size of the solution space of the signal period length of the intersection,  $\lambda$  represents the size of the solution space of

TABLE 5: Traffic section length standard deviation before and after static partition.

Core intersection	State	Standard deviation of section length	Number of intersections
—	Road network before partition	84.3	40
14	Road network after partition: Subarea 1	60.35	7
16	Road network after partition: Subarea 2	41.55	17
31	Road network after partition: Subarea 3	102.35	16
—	Average standard deviation of section length after partition in road network	74.57	40
—	Optimization rate (percentage of decrease in standard deviation)	11.54%	40

TABLE 6: Traffic flow density standard deviation before and after dynamic partition.

Time period	Core intersection	State	Standard deviation of section traffic flow density	Number of intersections
7:30	—	Road network before partition	5.48	40
	9	Road network after partition: Subarea 1	3.53	17
	14	Road network after partition: Subarea 2	7.31	4
	31	Road network after partition: Subarea 3	2.1	19
	—	Average standard deviation of traffic flow density after partition in road network	3.57	40
	—	Optimization rate (percentage of decrease in standard deviation)	34.85%	40
7:40	—	Road network before partition	5.56	40
	9	Road network after partition: Subarea 1	3.59	17
	14	Road network after partition: Subarea 2	6.07	5
	31	Road network after partition: Subarea 3	2.42	18
	—	Average standard deviation of traffic flow density after partition in road network	4.55	40
	—	Optimization rate (percentage of decrease in standard deviation)	18.17%	40
7:50	—	Road network before partition	4.78	40
	9	Road network after partition: Subarea 1	3.05	17
	14	Road network after partition: Subarea 2	5.42	5
	31	Road network after partition: Subarea 3	2.44	18
	—	Average standard deviation of traffic flow density after partition in road network	3.05	40
	—	Optimization rate (percentage of decrease in standard deviation)	36.19%	40
8:00	—	Road network before partition	5.05	40
	16	Road network after partition: Subarea 1	2.79	18
	14	Road network after partition: Subarea 2	6.03	4
	31	Road network after partition: Subarea 3	2.58	18
	—	Average standard deviation of traffic flow density after partition in road network	3.18	40
	—	Optimization rate (percentage of decrease in standard deviation)	37.03%	40

the green signal ratio of the signal period of the intersection, and  $N$  represents the number of intersections. In this experiment,  $N = 40$ . The size of the solution vector space after partition is  $\sum_0^i (T \times \lambda)^{N_i}$ , where  $i$  is the number of partitions and  $N_i$  is the number of intersections in the subregion  $i$ . In this experiment  $\sum_0^i N_i = 40$ , it can be concluded that  $\sum_0^i (T \times \lambda)^{N_i} \leq (T \times \lambda)^{40}$ , if and only if the number of partitions  $i = 1$ ,  $\sum_0^i (T \times \lambda)^{N_i} = (T \times \lambda)^{40}$ . Taking this experiment as an example,  $T = 5$  and  $\lambda = 10$  were set to calculate the size of solution space before and after partitioning. The results are shown in Table 7.

As shown in Table 7, the size of the solution vector space after partitioning is significantly smaller than that before partitioning. Therefore, the performance of the traffic signal control system is improved after the partition.

The experimental results show that the subdivision method can not only meet the demand of traffic signal control, but also better adapt to various traffic control strategies to some extent.

**4.3.3. Method Comparison.** This paper, together with literature [27–30], believes that there is a macroscopic basic diagram (MFD) in road network, which plays an important role in dynamic traffic management, traffic signal control, and urban traffic congestion mitigation. However, the large-scale road network is usually a heterogeneous network, and the heterogeneity has a strong influence on the shape of MFD and the mode of the system in MFD. Therefore, the urban road network must be divided into multiple

homogeneous regions to achieve the purpose of carrying out different traffic signal control strategies in multiple subareas. However, it is difficult to collect such data as travel time, queue length, and vehicle density in practical application, and the data that can be effectively collected must be used for partitioning.

Snake algorithm is used in [27] for preliminary clustering of road network, and mixed integer linear programming model is used for partitioning. They divide the data into different sets with high similarity within the sets and low similarity between the sets. Clusters with similar service characteristics are extracted by using this partitioning idea. When clustering traffic data, the spatial adjacency constraints are also considered, and the traffic data are clustered into spatially connected homogeneous partitions. After clustering, a certain number of traffic partitions are obtained. The traffic conditions in these partitions are similar, and the internal roads are also connected. In [28], a fast network division method based on modular optimization was adopted as the standard to classify the advantages and disadvantages of the results, and the isomorphic subnetworks in large-scale urban road networks were identified. They proposed a method to calculate the correlation degree of traffic conditions at adjacent intersections quantitatively. Then, the modularized and optimized fast network partition method is used as the criterion to identify the isomorphic subnetworks in the large-scale urban road network. Literature [29] proposed an evolutionary spectral clustering algorithm to divide time-varying heterogeneous networks into connected homogeneous regions. The sparse matrix is obtained by calculating the similarity, which simplifies the complexity of the algorithm. They also applied Evolutionary Spectrum Clustering algorithms to road speeds to obtain clustering results that fit current traffic conditions without deviating from past history. Based on the self-set zoning index in [30], weighted K-means clustering, K-harmonic means clustering, and spectral clustering are used, respectively. Three clustering algorithms are used to perform traffic partitioning and are compared. From the experimental results, that K-means clustering will lead to unstable clustering results due to different selection of clustering center points. The partitioning method proposed in the above literature is compared with the partitioning method proposed in this paper in various aspects (see Table 8).

The following can be seen from Table 8:

- (1) From the traffic parameters and zoning indexes used in the zoning process, Pearson Product-Moment Correlation Coefficient (PCC) was calculated using the average driving speed and average density of vehicles on the road in [27]. It is used as the homogeneity index of partition. PCC is usually used to calculate the similarity of two indicators, while homogeneity partition pays more attention to the deviation degree of a sample in the whole sample, which is not well described by PCC. At the same time, this method lacks relevant traffic parameters to describe the homogeneity of the road structure, such as the length of adjacent roads. Literature [28] used the traffic flow between intersections and the length of sections to calculate the similarity between sections and believed that such similarity could also describe the degree of road congestion. From the calculation formula of similarity index, this index can describe the homogeneity of traffic flow density more, but the ability to describe the homogeneity of road structure is a little insufficient. Literature [29] only uses the average speed of a section as a clustering index, and the average speed of a section can only describe the traffic flow density of a section indirectly, which is not intuitive. At the same time, there is also a lack of relevant traffic parameters that can reflect the description of the homogeneity on the road structure. Literature [30] calculated a custom zoning index CV based on the variance of traffic flow density in each subarea and the variance of traffic flow density in the whole road network. The purpose is to minimize the density difference of traffic flow within all subregions and maximize the difference of traffic flow within each subregion. This method can describe the homogeneity of traffic flow density well. However, there is also a lack of relevant traffic parameters to describe the homogeneity of road structure.
- (2) The zoning indexes adopted in this paper are “shortest path distance,” “average distance of road sections,” and “average density of traffic flow at intersections.” Among them, the “shortest path distance” ensures the continuity of the road network structure of each subarea. At the same time, “shortest path distance” and “average section distance” jointly ensure the homogeneity of each subarea in the road structure (adjacent section length). The “average density of traffic flow at intersection” ensures the homogeneity of traffic flow density in each subarea.
- (3) From the perspective of partitioning objects, literature [27] takes sections as partitioning objects. And the traffic signal control is indeed located in the intersection of the traffic signal lights as the control body. If the section is taken as the partition object, there will inevitably be boundary intersections in the two traffic subsections. At this time, it is impossible to determine which traffic subareas these boundary intersections should be classified into, so it is easy to cause difficulties in decision-making of traffic signal control strategies at boundary intersections. In this paper, the intersection is considered as the partition object, to solve the boundary intersection traffic signal control difficulties.
- (4) In terms of the basic data of the initial partition, in the methods such as in [27], basic data such as the size and number of subareas and clustering standards are artificially determined by historical experience information. However, the partitioning method itself does not calculate and control these aspects. This makes these partitioning methods difficult to implement in practical applications.

TABLE 7: The size of the solution vector space before and after partition.

Time period	The size of the solution vector space
Before partition	$9.095 \times 10^{67}$
7:30	$1.908 \times 10^{32}$
7:40	$1.945 \times 10^{32}$
7:50	$1.945 \times 10^{32}$
After partition	$7.629 \times 103^0$
8:00	$7.629 \times 103^0$
8:10-9:20	$7.782 \times 10^{28}$

This paper proposes a method to determine the cluster center point (the core intersection) by the road connectivity degree (N-BC) and takes the cluster center point as the clustering standard. The method is simple to calculate, and the experiment demonstrates that the method is effective. At the same time, based on the performance of the computer, the size of the partition is constrained, and the number of partitions is controlled by the size of the partition, so that even when the traffic jam does not occur, a partition will not be too large, and the number of subareas in the whole traffic area can be controlled within a reasonable range.

- (4) Zoning methods such as in [27] can only play an auxiliary role in traffic signal control when traffic congestion occurs. However, in the non-traffic congestion state, these zoning methods can play a very limited role in assisting the traffic signal control system and may even fail to achieve the correct zoning. This is because the main purpose of these methods of zoning is to distinguish between traffic congestion areas and non-traffic congestion areas. When the entire traffic area is free of congestion, there is no definite zoning standard.

Based on this, this paper adopts the method of combining dynamic and static partitions to partition traffic. When the traffic flow in the road network is small, static partitioning is adopted for traffic partitioning, while when the traffic flow in the road network is large, dynamic partitioning will be carried out on the basis of static partitioning. The “homogeneity of sections structure” and “homogeneity of traffic flow density” were used as the evaluation criteria of zoning performance. When the traffic flow is small, the main influence on the green wave control of traffic signals is also the average travel time of each section, which is closely related to the length of the sections. Therefore, “homogeneity of sections structure” is taken as one of the evaluation criteria for static zoning. When traffic congestion occurs, traffic flow density will be the most direct parameter to be used in the traffic signal control strategy. Therefore, “homogeneity of traffic flow density” is taken as one of the evaluation criteria of dynamic partition.

- (5) From the perspective of computational time complexity, the computational time complexity needed

to carry out an iteration in this paper is similar to that in [29, 30], but better than that in [27, 28]. However, the number of iterations converging to the optimal solution in this paper will be better than those in [29, 30]. This is because this paper takes the initial partition result as the initial solution of the static partition and carries out iterative search on this basis. In dynamic partitioning, static partitioning results are used as the initial solution of dynamic partitioning, and iterative search is carried out on this basis. The correct choice of the initial solution will greatly accelerate the convergence rate of the calculation. It is verified by experiments that the number of iterations of this method can converge to the optimal solution between 1 and 2 times.

4.3.4. *Ablation Experiment.* To analyze the performance of every step in partitioning algorithm, three types of ablation experiments were designed in this paper, and the data of 7:30, 7:40, 7:50, and 8:00 were used for verification:

- (i) Ablation experiment 1: in the zoning algorithm, the step of finding the core intersection and iteratively searching the best zoning result based on the core intersection is omitted. Based on the vertex intersection, the road network is partitioned by using Mahalanobis distance.
- (ii) Ablation experiment 2: eliminate the step of static partition in the partitioning algorithm. Directly use the dynamic partitioning method to partition the road network.
- (iii) Ablation experiment 3: eliminate the calculation of Mahalanobis distance in the zoning algorithm. Directly use the sum of the shortest path distance and the difference of traffic flow density to replace the Mahalanobis distance for partition.

The ablation experimental results are shown in Table 9.

It can be seen from Table 9 that although the standard deviation of the average traffic flow density of ablation experiment 1 and ablation experiment 3 was reduced compared with that before the partition, the optimization rate was very small. This means that the iterative calculation of the core intersection and the calculation of Mahalanobis distance are necessary steps in the division. In ablation experiment 2, although the average traffic flow density was significantly reduced, the number of iterations calculated to the approximate optimal solution was significantly increased and the convergence rate was significantly slower compared with that without ablation experiment. This will reduce the overall operation efficiency of the traffic system, and the traffic system is a high real-time system, and the slow convergence speed of the partition is intolerable for the whole intelligent transportation system. By ablation experiments, all stages and steps used in the zoning methods are verified effective and efficacious.

TABLE 8: The partition method in this paper is compared with other partition methods.

Name	Literature [27]	Literature [28]	Literature [29]	Literature [30]	This paper
Zoning parameter	(1) Average speed (2) Average density	(1) Traffic flow (2) Section length	Average speed	(1) Average distance (2) Variance of density	(1) N-BC (2) Section average length (3) Shortest path length (4) Traffic flow density
Zoning index	PCC	Section attribute similarity	Speed similarity	CV	Mahalanobis distance
Basic data of the initial partition	Empirical data	Empirical data	Empirical data	Empirical data	Static data of road network
Partition object	Section	Section	Section	Section	Intersection
Homogeneity of network structure	Not supported	Supported	Not supported	Supported	Supported
Homogeneity of dynamic traffic data	Supported	Supported	Supported	Supported	Supported
Partition without congestion	Not supported	Not supported	Not supported	Not supported	Supported
Standardization of probability density space in multi-index evaluation	Not supported	Not supported	Not supported	Not supported	Supported
Computational time complexity of partition method	$O(2^n)$	$O(n^2)$	$O(t*k*n)$	$O(t*k*n)$	$O(2*k*n)$

$n$  for the scale of the road network;  $t$  is the number of partitions;  $k$  is the number of iterations.

TABLE 9: Comparison of ablation experiment results.

Time	Ablation algorithm	Cumulative iteration	Intersection	Density standard deviation before partition	Density standard deviation after partition	Optimization change (%)
7:30	No ablation	3	40	5.48	3.57	34.85
	Ablation experiment 1	2	40	5.48	5.41	1.28
	Ablation experiment 2	10	40	5.48	3.60	34.31
	Ablation experiment 3	3	40	5.48	5.48	0
7:40	No ablation	2	40	5.56	4.55	18.17
	Ablation experiment 1	2	40	5.56	5.53	0.54
	Ablation experiment 2	8	40	5.56	4.55	18.17
	Ablation experiment 3	3	40	5.56	5.47	1.62
7:50	No ablation	2	40	4.78	3.05	36.19
	Ablation experiment 1	2	40	4.78	4.65	2.72
	Ablation experiment 2	7	40	4.78	3.31	30.75
	Ablation experiment 3	3	40	4.78	4.51	5.65
8:00	No ablation	3	40	5.05	3.18	37.03
	Ablation experiment 1	2	40	5.05	4.86	3.76
	Ablation experiment 2	7	40	5.05	3.23	36.04
	Ablation experiment 3	3	40	5.05	4.98	1.39

The ‘‘cumulative number of iterations’’ is the sum of the numbers of iterations for each partition.

### 5. Conclusion

This paper discusses the importance of traffic area partition and expounds the influencing factors of traffic area subdivision. In view of the two main traffic control strategies, a method combining static partition and dynamic partition of road network is proposed. On the premise of balancing the computation scale of the traffic control system, the results of the partition are more consistent with the two main traffic control strategies. At the same time, because the traffic flow density data is difficult to collect, a traffic

flow density estimation method is proposed. The experiment verifies that this method can accurately divide the subarea of any shape traffic area. Although traffic partitioning does not directly and effectively alleviate traffic congestion, it can provide a reasonable traffic control area for traffic signal control and play a certain role in the optimization of traffic control effect. It is hoped to provide a new way of thinking for the subdivision technology of traffic signal control.

In future, we will continue to analyze the state-of-the-art research results related to the existing traffic signal control

and combine them with a variety of traffic partitioning strategies and further analyze and make comparisons under different traffic signal control methods to adopt different means of traffic subarea so as to improve the performance of the transportation system.

## Data Availability

The data used to support the findings of this study are included within the supplementary information file(s).

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Supplementary Materials

1. In Appendix 1, “Road section”: for example, “ $a_1$ - $a_2$ ” refers to the section between  $a_1$  and  $a_2$  intersections; “number of two-way lanes” refers to the number of two-way lanes in the section between these two intersections; and “section length” refers to the distance between these two intersections. 2. The data in Appendix 2 show the traffic flow of 40 intersections in Figure 4 from 7:20 AM to 9:20 AM, for a total of 12 time periods, each of which is 10 minutes. (*Supplementary Materials*)

## References

- [1] Y. Zhang, X. X. Jing, and J. M. Hu, “A preliminary study on the complexity analysis and application of urban traffic system,” in *Proceedings of the 2005 China Automation Industry Summit: China Enterprise Automation and Informatization Construction Forum-- Report of the Conference Sub-Forum IV: Intelligent Transportation*, pp. 103–110, Beijing, China, July 2005.
- [2] Y. Y. Ma, X. G. Yang, and Y. Zen, “Urban traffic signal network partitioning using spectral method,” *Systems Engineering—Theory & Practice*, vol. 30, no. 12, pp. 2290–2296, 2010.
- [3] Y. M. Bie, L. H. Wang, D. H. Wang et al., “Strategy of dynamic traffic control subarea partition in urban road network,” *China Journal of Highway and Transport*, vol. 26, no. 6, pp. 157–168, 2013.
- [4] K. Lu, J.-M. Xu, S.-J. Zheng, and S.-M. Wang, “Research on fast dynamic division method of coordinated control sub-area,” *Acta Automatica Sinica*, vol. 38, no. 2, pp. 279–287, 2012.
- [5] Z. Qiao, L. Ke, G. Zhang et al., “Adaptive collaborative optimization of road network signal timing based on immune-fireworks algorithm and hierarchical strategy,” *Applied Intelligence*, pp. 1–17, 2021.
- [6] X. Wang, L. Ke, Z. Qiao, and X. Chai, “Large-scale traffic signal control using a novel multiagent reinforcement learning,” *IEEE Transactions on Cybernetics*, vol. 51, no. 1, pp. 174–187, 2021.
- [7] S. Yang and B. Yang, “A semi-decentralized feudal multi-agent learned-goal algorithm for multi-intersection traffic signal control,” *Knowledge-Based Systems*, vol. 213, Article ID 06708, 2021.
- [8] Z. Zeng, “Graph light: graph-based reinforcement learning for traffic signal control,” in *Proceedings of the International Conference on Computer and Communication Systems*, pp. 645–650, Chengdu, China, April 2021.
- [9] R. J. Walinchus, *Real-time Network Decomposition and Subnetwork Interfacing*, pp. 20–28, Transportation Research Board, Washington, DC, USA, 1971.
- [10] H. N. Yagoda, E. H. Principe, C. E. Vick et al., “Subdivision of signal systems into control areas,” *Traffic Engineering, Institute of Traffic Engineers*, vol. 43, no. 12, pp. 42–47, 1973.
- [11] E. C. P. Chang, “How to decide the interconnection of isolated traffic signals,” in *Proceedings of the 17th Conference on Winter Simulation*, pp. 445–453, Association for Computing Machinery, New York, NY, USA, December 1985.
- [12] L. T. Lin and S. M. Tsao, “A System Approach on Signal grouping for areawide control of computerized Traffic System,” in *Proceedings of the 79th Transportation Research Board Annual Meeting*, pp. 1–21, TRB, Washington, DC, USA, January 2000.
- [13] L. T. Lin and H. J. Huang, “A linear model for determining coordination of two adjacent signalized intersections,” *Journal of Modelling in Management*, vol. 4, no. 2, pp. 162–173, 2009.
- [14] C. G. Li, Y. Xie, H. L. Zhang et al., “Dynamic division about traffic control sub-area based on back propagation neural network,” in *Proceedings of the 2nd International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC'10)*, pp. 22–25, Nanjing, China, August 2010.
- [15] X. J. Tian, *Research on Traffic Signal Control Optimization Methods under Unsaturated State*, Jilin University, Changchun, China, 2018.
- [16] Y. L. Zhou, *Research on Coordinated Control Method of Traffic Signal Timing in Urban Area*, Tianjin Industry University, Tianjin, China, 2016.
- [17] W. T. Zhao, X. J. Qian, Y. Zhu et al., “Research on the dynamic traffic zoning based on the key junctions,” *Computer and Applied Chemistry*, vol. 29, no. 10, pp. 1241–1244, 2012.
- [18] Y. J. Feng, M. Shan, H. C. Yue et al., “Subarea dynamic division algorithm based on green wave coordinated control,” *Control Theory and Application*, vol. 3, no. 8, pp. 1034–1046, 2014.
- [19] D. W. Xia, B. F. Wang, Y. T. Li et al., “An efficient Map-Reduce-based parallel clustering algorithm for distributed traffic subarea division,” *Discrete Dynamics in Nature and Society*, vol. 1, no. 1, pp. 1–18, 2015.
- [20] X. D. Li, X. G. Yang, and H. J. Chen, “Division method of traffic area in urban road network,” *Computer Engineering and Application*, vol. 45, no. 5, pp. 19–22, 2009.
- [21] J. H. Zhu, X. Z. Niu, and C. S. Q. Wu, “On a clustering-based approach for traffic sub-area division,” in *Proceedings of the International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems*, Graz, Austria, June 2019.
- [22] Y. Ji and N. Geroliminis, “On the spatial partitioning of urban transportation networks,” *Transportation Research Part B: Methodological*, vol. 46, pp. 1639–1656, 2012.
- [23] L. Shu, J. K. Qing, and Q. M. Huang, “A model-based demand-balancing control for dynamically divided multiple urban subnetworks,” *Journal of Advanced Transportation*, vol. 50, pp. 1046–1060, 2016.
- [24] T. Zong and T. Urbanik, “System partition technique to improve signal coordination and traffic progression,” *Journal of Transportation Engineering*, vol. 133, no. 2, pp. 119–128, 2007.
- [25] H. Ghorbani, “Mahalanobis distance and its application for detecting multivariate outliers,” *Facta Universitatis Series*



- Mathematics and Informatics*, vol. 34, no. 3, pp. 583–595, 2019.
- [26] A. N. Kang, Y. C. Chiu, X. B. Hu, and X. H. Chen, “A network partitioning algorithmic approach for macroscopic fundamental diagram-based hierarchical road network management,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 4, pp. 1130–1139, 2018.
  - [27] D. Liu, M. Wang, and G. Shen, “A new combinatorial characteristic parameter for Clustering-Based road network partitioning,” *IEEE Access*, vol. 7, pp. 40175–40182, 2019.
  - [28] Z. Zhou, S. Lin, and Y. Xi, “A dynamic network partition method for heterogenous urban road networks,” in *Proceedings of the 2012 15th International IEEE Conference on Intelligent Transportation Systems*, pp. 820–825, Anchorage, AK, USA, September 2012.
  - [29] P. Al Alam, D. Hamad, J. Constantin et al., “Dynamic partitioning of transportation network using evolutionary spectral clustering,” in *Proceedings of the Smart Applications and Data Analysis (SADASC 2020)*, pp. 178–186, Marrakesh, Morocco, June 2020.
  - [30] A. F. Lentzakis, R. Su, and C. Wen, “Time-dependent partitioning of urban road network into Homogeneous Regions,” in *Proceedings of the 2014 13th International Conference on Control, Automation, Robotics & Vision (ICARCV)*, pp. 535–540, Singapore, December 2014.