The Optimization Method of Dynamic Traffic State Identification on Urban Arterial Road

Shuo Tong
School of Traffic and Transportation
Beijing Jiaotong University
Beijing, China
E-mail: tongd126yx@126.com

Abstract—Dynamic traffic state identification is important to traffic control and guidance on urban arterial road, according to the fuzzy characteristics of the traffic flow, this paper firstly establishes a dynamic traffic state identification model based on improved Fuzzy C-Means (FCM) clustering to divide traffic states into four phases. Secondly, because the FCM algorithm is difficult to distinguish the states good or bad in the same phase, this paper establishes a dynamic traffic state identification model based on Data Envelopment Analysis (DEA), which uses the comprehensive index to evaluate each state sample, and uses similarity between the results of improved FCM and DEA to evaluate this model. Thirdly, this paper analyzes the statistical features of traffic state changes from relative smooth to congestion (congestion to severe congestion) combining with time sequence, and proposes the pre-identification number of long-time congestion. Finally, the actual data of a urban arterial road in H city as an example are verified the effects of the method, the similarity of two models is 86.5%, which illustrates the method of this paper is feasible, and the pre-identification number is three.

Keywords—dynamic traffic state identification; improved FCM algorithm; Data Envelopment Analysis (DEA); long-time congestion; congestion pre-identification

I. INTRODUCTION

Dynamic traffic state identification is the basis of city road traffic control and management measures, timely and reasonable traffic state identification can provide reference information to traffic management and road users, which is propitious to optimize traffic condition in real time and improve the efficiency of road. Dynamic traffic state classification is based on traffic flow theory, combined with the mathematics, system engineering and so on to the variation characteristics and laws of dynamic traffic parameters, so as to get the threshold of traffic state change to distinguish.

Over the last few decades, domestic and foreign scholars have done a lot of research on this problem, some scholars have designed the California algorithm [1], the traffic state identification based on the basic graph theory [2] and Boris S. Kerner proposed three phases traffic flow theory [3]; some scholars think traffic is chaotic, so the Fuzzy C-Means(FCM) clustering can be used based on analysis on the similarity and dissimilarity between different traffic state parameters in the samples. Reference [4] used FCM to divide traffic states to 6 classes based on the flow and speed data detected; Reference [5] used genetic algorithm to improve FCM, and analyzed the scope values of the traffic state based on simulation traffic data. This paper proposes an improved FCM algorithm to identify traffic state phases.

Clustering method not only can describe the chaos of the traffic flow, but also can achieve the recognition of all sample points, but it is difficult to distinguish between the different traffic states in the same phase. Reference [6] used fuzzy comprehensive evaluation method to determine the degree of congestion, which can specifically grade each state, but the weight of this method is subjective, the more objective evaluation methods are needed, so a new traffic state identification method based on Data Envelopment Analysis(DEA) is presented.

The most important significance in identification of traffic state is pre-identification for traffic congestion, previous studies obtain that congestion is not only the change of traffic state, but also should last for a period of time, which kind of congestion is needed to control. Congestion is given the definition that should continue for 15 minutes or more in Japan[7], and it should be collected three consecutive periods of data to determine whether the congestion or not in McMaster algorithm [1]. So the congestion categories of domestic arterial road are analyzed in this paper.

II. DYNAMIC TRAFFIC STATE IDENTIFICATION MODEL

A. Dynamic Traffic State Identification Model Based on Improved FCM

Fuzzy C-Means (FCM) clustering is a typical fuzzy clustering method, which was proposed by Bezdek (1984)[8], and it is the process of dividing data elements into classes or clusters so that items in the same class are as similar as possible, and items in different classes are as dissimilar as possible. In view of the fuzziness and uncertainty of the traffic state, according to the characteristics of the parameters of different traffic states, FCM can use membership to describe the relation degree that the data point belongs to which phase.

This method needs appropriate classification parameters: the initial value of cluster centers, the fuzzy exponent (m) and the number of clusters (c)[8]. To optimize the selection of appropriate parameters, in this study, we first use subtractive clustering to optimize c and the initial value of cluster centers, and use partition entropy index to optimize m, the partition entropy is used as the
evaluation index of judging the validity of clustering, the specific steps of the algorithm are as follows:

**Step1** Set \( \varepsilon_1 \in [0.1, 1], m \in [1.1, 2.5] \) [9]

**Step2** The initial clustering center selection -- each sample point as a potential cluster center

**Step3** Calculate the measured value \( P_i \) of each sample point \( x_i \) : 
\[
P_i = \frac{4}{\sum_{j=1}^{n} e^{-d_{ij}}}; \quad \text{where, } \alpha = \frac{4}{r_{a}^{2(m-1)}} ,
\]
where \( r_a \) is neighbor radius, let \( r_a \equiv r_m \).

\[
r_m = \frac{1}{2} \sqrt{\frac{1}{n(n-1)} \sum_{k=1}^{n} \sum_{i=1}^{n} \|x_j - x_k\|^2}
\]
(1)

calculate the value of \( P_i \), select the highest value \( P_i \) of the corresponding point \( x_i \) as the initial clustering center point \( x_i^* \).

**Step4** Determine the next cluster center:
\[
P_i = P_i - P_i^* e^{-\beta |x_i^* - x_i|}
\]
(2)
where \( \beta = \frac{4}{r_\beta}, r_\beta = 1.5 r_a \), the farther from the sample point \( x_i^* \) is, the greater \( P_i \) is, which becomes a cluster center more possibly. So select the sample point of the highest value \( P_i \) as the cluster center, and remove the cluster center from the candidate cluster center sample points next time.

**Step5** Calculate the other clustering centers in accordance with the method of Step 3 with \( k \) calculations,
\[
P_i = P_i - P_k^* e^{-\beta |x_i^* - x_k|}
\]
(3)
repeat Step 5 until the objective function for the optimal solution meets the following conditions, that is \( P_k^* < \varepsilon_1 P_i \), then find the initial clustering centers and the number of clusters (c)

**Step6** Set the number of iterations and \( m \)

**Step7** Calculate the membership degree matrix \( U^{(r)} \):
\[
U^{(r)} = \begin{cases} 
\frac{1}{\sum_{k=1}^{c} \left( \frac{d_{ij}^{(r)}}{d_{ij}^{(r)}} \right)^{2(m-1)}} & d_{ij}^{(r)} > 0 \\
\frac{\sum_{k=1}^{c} \left( \frac{d_{ij}^{(r)}}{d_{ij}^{(r)}} \right)^{2(m-1)}}{d_{ij}^{(r)}} & d_{ij}^{(r)} = 0 and \ i = j \\
0 & d_{ij}^{(r)} = 0 and \ i = j
\end{cases}
\]
(4)

**Step8** Use the following (5) update the clustering center:
\[
y^{(r+1)}_j = \frac{\sum_{j=1}^{n} (u^{(r)}_{ij})^m x_j}{\sum_{j=1}^{n} (u^{(r)}_{ij})^m}
\]
(5)

**Step9** If \( \|y^{(r)} - y^{(r+1)}\| \leq \varepsilon_2 \), stop the iteration, if not satisfied, return Step 6

**Step10** Calculate partition entropy \( H \):
\[
H = -\frac{1}{n} \sum_{j=1}^{n} u_{ij} \log u_{ij}(u_{ij})
\]
(6)
calculate the corresponding partition entropy and return Step 1, then compare the partition entropy to find the minimum value and stop the algorithm.

### B. Dynamic Traffic State Identification Model Based on Data Envelopment Analysis(DEA)

The improved FCM uses membership to describe the relation degree that one data point belongs to one phase, which can solve the problem that some methods are unable to distinguish the all points to phases, however, it is difficult to know how far the sample is from the cluster center and to evaluate states in the same phase. Maybe it can be seen by means of two-dimensional map, but it is difficult to identify for three-dimensional or more than three, so it is necessary to further quantify.

Data Envelopment Analysis (DEA) is a comprehensive evaluation method of multi parameters, without the need for dimensionless method or subjective weight for parameters [10], which is better than the other method of comprehensive evaluation.

The DEA model is originally derived from the evaluation of engineering efficiency, and later extend to become an evaluation system which can evaluate multiple input and multiple output parameters, this method has been used in the traffic field, the most is the static evaluation of the overall road, this paper introduces the method for dynamic traffic state evaluation.

1) DEA principle

The DEA model is defined for each sample as a decision making unit (DMU), there are a set of input parameters and output parameters for each decision making unit, when shadowed into multidimensional space; n decision making units are composed of a polyhedral cone, the plane for multi cone of the most optimal decision making unit is the relative surface of DEA (i.e., efficient production frontier)[10], which comprehensive index \( \theta \) is 1, other decision making unit (DMU) can be compared with the efficient production frontier by comprehensive index \( \theta \), the farther the deviation degree is, the smaller value of \( \theta \) is \( \theta \in (0, 1] \), DEA is a kind of relative comparison among samples, where the weight of each parameter is according to the optimal weight distribution of each decision making unit.
2) Dynamic traffic state identification model thinking based on DEA

According to the principle of DEA, each sample can be seen as a DMU, different traffic states will get different comprehensive index $\theta$ when identified. The $\theta$ of the smoothest state sample is 1, which is the point on efficient production frontier in DEA, samples of the other states are deviated from the surface successively from the states of smooth to serious congestion, and their comprehensive indexes decrease successively, so they can effectively describe the different traffic states.

For which phase the state concretely belongs to, the point of the most extreme smooth and each clustering center of traffic state phases can be used as reference values to construct a framework, so we can calculate the deviation between the comprehensive index of DMU and above reference values, and judge which phase the DMU belongs to by the minimum distance from which reference value (i.e., principle of minimum deviation). Based on the assumption that the states are divided into 4 phases, then we select a smooth extremum and clustering centers of smooth, relative smooth, congestion and serious congestion to calculate their comprehensive indexes, their corresponding index are $\theta_1, \theta_2, \theta_3, \theta_4$, and DMU $i$ obtains a comprehensive index $\theta_i$, if $\|\theta_i - \theta_1\| < \|\theta_i - \theta_2\|$, DMU $i$ is identified smooth, if $\theta_i - \theta_1 < 0$, it shows that the state of DMU $i$ is between the smooth clustering center and relative smooth clustering center, if $\theta_i - \theta_1 > 0$, the state of DMU $i$ is between the extreme smooth value and smooth clustering center.

3) Parameters and variables

The selected parameters include input and output parameters, the input parameters are positive parameters, expected the smaller the better; the output parameter is negative parameters, expected the bigger the better; this paper selects flow $q$, and travel-time delay $d$, as input parameters, and selects speed $v$, as output parameter.

Suppose $n$ samples need to identify, so the definition of variables are as follows:

- $x_{ij}$: input parameter $i$ in sample $j$
- $y_{ij}$: output parameter $k$ in sample $j$
- $u_i$: the weight coefficient of the output parameter $i$
- $v_i$: the weight coefficient of the input parameter $i$

4) DEA model —— $C^2R$ model

$$
\begin{align*}
\max h_{j0} &= \frac{\mu y_{j0}}{\sum_{i=1}^{n} v_i x_{ij}} \\
C^2R &\text{ s.t. } \frac{\mu y_{j0}}{\sum_{i=1}^{n} v_i x_{ij}} \leq 1, j = 1, 2, \ldots, n \\
v_i &\geq 0
\end{align*}
$$

(7)

Set $t = \frac{1}{\sqrt{\mathbf{x}_0^T \mathbf{x}_0}}, \mathbf{w} = tv, \mathbf{\mu} = tu$ this fractional program can be equivalently stated as the following linear programming problem:

$$
\begin{align*}
\max h_{j0} &= \max \mu^T y_{j0} = V_p \\
P &= \text{ s.t. } \mathbf{w}^T \mathbf{x}_j - \mu^T \mathbf{y}_j \geq 0, j = 1, 2, \ldots, n \\
\mathbf{w}^T \mathbf{x}_{j0} &= 1 \\
\mathbf{w} &\geq 0, \mu \geq 0
\end{align*}
$$

(8)

where, $\mathbf{x}_j = (x_{ij}, y_{kj})^T, j = 1, 2, \ldots, n, \mathbf{w} = (w_1, w_2)^T$.

Let the above model into its dual problem:

$$
\begin{align*}
\min \theta &= V_p \\
P &= \text{ s.t. } \sum_{j=1}^{n} \lambda_j + S^- = \theta x_{j0} \\
\lambda_j &\geq 0, j = 1, 2, \ldots, n \\
S^- &= (s_1^-, s_2^-) \geq 0
\end{align*}
$$

(9)

Find out $\theta$, when $\theta_0 = 1, \quad s^0 = 0, s^{\omega} = 0$, the decision unit $j_0$ is effective, if $\theta_0 < 1$, the DMU is invalid.

C. Analysis of Congestion Classification and Congestion Pre-identification

According to the above introduction of some researches on congestion identification, we can see that the changes of traffic state phase do not necessarily continue, some may be restored within a short period of time, this kind of change may be due to a large inflow into this road section, but the next road is a less traffic, so the congestion is caused short time, this kind of congestion on the road can be adjusted automatically without the need of traffic control and guidance. Conversely, if the changes to more severe phases continue long, it is known for long time congestion, road itself is difficult to adjust, so the case should be promptly identified of the traffic state and pre-management.

So it is necessary to analyze the characteristics of long time congestion, from the definitions of congestion by Japan and the McMaster algorithm, we can know that identification number or threshold can be used to describe the category of congestion, this paper uses the method of...
data statistics to determine the required number of discriminant analysis and threshold congestion.

III. Case Analysis

A. Data Description

This analysis is carried out on an arterial road located in city H, China, which is 4 lanes per direction. And the data of flow, speed and travel time obtained directly by the devices are complete. The collecting time continues for about a week, from 0 a.m. to 12 p.m., each day in the research[2], and the time slice is 5 minutes. The number of samples is 2016 in total, which is more than the minimum required (60)[1], sufficient traffic data are necessary to ensure all states can be demonstrated.

B. Data Pre-processing

The text assumes these samples can represent the whole states of the traffic on this road. Firstly, we identify and process the error data of flow, speed and travel-time, and turn the travel time into time delay. Secondly, we use the “3σ” rule to process the error data, finally, it is totally removed 13 error data, remained 1992 samples. Thirdly, we use the means of dimensionless processing to the data[11], these means can retain the variable information and enhance the comparability among the variables, which is more suitable for clustering method.

C. Solutions and Analysis Based on the Improved FCM Model

According to the improved FCM algorithm, we use MATLAB 2014 (a) to implement the program, and the partitioning results are as follows:

When \( m = 1.2, \epsilon_1 = 0.2, \epsilon = 4 \) and the partition entropy is minimum, the data are divided into four clusters, i.e., clusters 1, 2, 3 and 4, the clustering center matrix \( V \) is obtained:

\[
V = (q, v, d) = \begin{pmatrix}
112 & 31.20 & 20.48 \\
282 & 19.36 & 61.41 \\
305 & 10.09 & 157.25 \\
312 & 6.08 & 292.38
\end{pmatrix}
\]

The first column of the matrix \( V \) represents flow, the second column represents the speed, and the third column represents the travel-time delay; each row represents a different phase, we name them smooth(phase1), relative smooth (phase2), congestion (phase3) and serious congestion (phase4) from the top to down. Fig.1 is a three-dimensional graph of the clustering and classification results, state points in blue area are identified as smooth, state points in green area are identified as relative smooth, state points in yellow area are identified as congestion, state points in red area are identified as serious congestion.

D. Solutions and Analysis Based on DEA Model

We select each clustering center in matrix \( V \) and a extremal point in the most smooth value, and solve the traffic state identification model based on DEA by using DEAP software, obtaining the comprehensive index of the cluster centers under different conditions and the extreme points of in the most smooth value, as shown in TABLE II.

![Figure 1. A three-dimensional graph of the clustering and classification results](image)

So when identifying the traffic state dynamically, we just get the data of flow, speed and travel-time delay, and combine with the cluster center to calculate the membership degree of each phase, then use the principle of maximum membership degree state to identify. For instance, we choose some samples to identify the traffic state in accordance to the above method, the results are shown in TABLE I:

<table>
<thead>
<tr>
<th>Sample</th>
<th>Flow</th>
<th>Speed</th>
<th>Travel-time delay</th>
<th>( u_1 )</th>
<th>( u_2 )</th>
<th>( u_3 )</th>
<th>( u_4 )</th>
<th>Phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>162</td>
<td>24.30</td>
<td>37.80</td>
<td>0.9783</td>
<td>0.0217</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>98</td>
<td>264</td>
<td>21.57</td>
<td>46.73</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>138</td>
<td>300</td>
<td>11.87</td>
<td>119.73</td>
<td>0</td>
<td>0.0048</td>
<td>0.9952</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>185</td>
<td>360</td>
<td>6.87</td>
<td>238.73</td>
<td>0</td>
<td>0</td>
<td>0.0225</td>
<td>0.9778</td>
<td>4</td>
</tr>
</tbody>
</table>

\( u_1, u_2, u_3, u_4 \) in TABLE I. represent separately the membership degree of four state values: smooth, relative smooth, congestion and serious congestion. From the table, we can see the membership degree of smooth phase is maximum in sample 13, so the state of the sample 13 can be judged as smooth.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Extreme smooth phase point</th>
<th>Phase1 center point</th>
<th>Phase2 center point</th>
<th>Phase3 center point</th>
<th>Phase4 center point</th>
</tr>
</thead>
<tbody>
<tr>
<td>speed</td>
<td>46</td>
<td>31.20</td>
<td>19.36</td>
<td>10.09</td>
<td>6.08</td>
</tr>
<tr>
<td>flow</td>
<td>69</td>
<td>112</td>
<td>282</td>
<td>305</td>
<td>312</td>
</tr>
<tr>
<td>travel-time delay</td>
<td>0.3</td>
<td>20.48</td>
<td>61.41</td>
<td>157.25</td>
<td>292.38</td>
</tr>
<tr>
<td>( \theta )</td>
<td>1</td>
<td>0.247</td>
<td>0.056</td>
<td>0.017</td>
<td>0.01</td>
</tr>
<tr>
<td>rank</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>
TABLE III. TRAFFIC RESULTS BASED ON DEA MODEL

<table>
<thead>
<tr>
<th>Sample</th>
<th>Flow</th>
<th>Speed</th>
<th>Delay</th>
<th>$\theta_i$</th>
<th>$|\theta_i - \theta_1|$</th>
<th>$|\theta_i - \theta_2|$</th>
<th>$|\theta_i - \theta_3|$</th>
<th>$|\theta_i - \theta_4|$</th>
<th>Traffic phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>165</td>
<td>29.50</td>
<td>23.00</td>
<td>0.176</td>
<td>0.07</td>
<td>0.10</td>
<td>0.16</td>
<td>0.166</td>
<td>1</td>
</tr>
<tr>
<td>98</td>
<td>264</td>
<td>21.57</td>
<td>46.73</td>
<td>0.073</td>
<td>0.17</td>
<td>0.02</td>
<td>0.05</td>
<td>0.063</td>
<td>2</td>
</tr>
<tr>
<td>138</td>
<td>300</td>
<td>11.87</td>
<td>119.73</td>
<td>0.023</td>
<td>0.22</td>
<td>0.03</td>
<td>0.006</td>
<td>0.013</td>
<td>3</td>
</tr>
<tr>
<td>185</td>
<td>360</td>
<td>6.87</td>
<td>238.73</td>
<td>0.009</td>
<td>0.23</td>
<td>0.04</td>
<td>0.008</td>
<td>0.001</td>
<td>4</td>
</tr>
</tbody>
</table>

TABLE II. shows the comprehensive index of the smooth extreme point is 1, which is the highest, the comprehensive indexes of worse traffic condition decrease successively, so they form a framework to the state, when recognizing, we only need calculate the comprehensive index of sample, and compare with the comprehensive indexes of these five points. The recognition results of the samples in TABLE II. are presented in TABLE III.

In TABLE III., $\|\theta_i - \theta_1\|$, $\|\theta_i - \theta_2\|$, $\|\theta_i - \theta_3\|$, $\|\theta_i - \theta_4\|$ represent the departure distance between comprehensive index of sample $i$ and each cluster center, and sample $i$ belongs the phase which is the degree of deviation is minimum. Such as the sample13, $\|\theta_{13} - \theta_1\|$ is the minimum in $\|\theta_{13} - \theta_1\|$, $\|\theta_{13} - \theta_2\|$, $\|\theta_{13} - \theta_3\|$, $\|\theta_{13} - \theta_4\|$, so the state is distinguished for the smooth state. The samples can also compare the state is good or bad with other samples directly by the comprehensive index, where the comprehensive index value is larger, where the state is more smooth.

E. The Evaluation of Identification Model Based on DEA

We use the improved FCM model to evaluate the results of identification model based on DEA by using the above results, the evaluation standard is the similarity of two models.

We select 288 samples on Monday, then use the two models to identify for the state, Fig. 2 shows the state estimation results of the two model, which phase 1, 2, 3, 4 respectively represent the smooth, relative smooth, congestion and serious congestion.

Figure 2. Similarity of identification results of two models

Fig. 2 shows that a total of 288 samples are tested, including the inconsistent classifications are 39 samples, the similarity is 86.5%, the different identification of samples are divided into two adjacent phases, so it can illustrate the correctness of the DEA method.

F. Analysis of Congestion Classification

We use the method of improved FCM to identify the traffic states of 2017 samples a week, corresponding to the original time series, we use the method of statistics to discuss congestion identification number and threshold criterion of congestion.

From the results of the improved FCM algorithm, we can see the general regulation of state changes: phase1 → phase2 → phase3 → phase4, but in the phase2 → phase3 and phase3 → phase4, there are some kinds of different situations, in addition, there are still a few changes from phase2 to phase4. Focus on the samples of state changes from phase2 to phase3 and phase3 to phase4, we found a total of 49 sample points of phase2 to phase3, in a total of 25 sample points of phase3 to phase4, of which the changes from phase2 to phase3 can be summarized as under several conditions:

1) condition1: a sample point of phase2 changes to phase3, then changes to phase2 again and continue: phase2 → phase3 → phase2…
2) condition2: a sample point of phase2 changes to phase3, then continue once in phase3 and continues in phase2 after changing back to phase2: phase2 → phase3 → phase2 → phase2…
3) condition3: a sample point of phase2 changes to phase3, then continues twice and continues in phase2 after changing back to phase2: phase2 → phase3 → phase3 → phase2 → phase2…
4) condition4: a sample point of phase2 changes to phase3, then continues three times and continues in phase3 or changes back to phase2: phase2 → phase3 → phase3 → phase3 → phase3 → phase2 → phase2…

The conditions changing from phase3 to phase4 are the same as the conditions changing from phase2 to phase3, We put these conditions summarized in TABLE IV. and TABLE V.

TABLE IV. SUMMARY OF STATE CHANGES FROM PHASE 2 TO PHASE 3

<table>
<thead>
<tr>
<th>Changes</th>
<th>Phase 2 to phase 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>conditions</td>
<td>1 2 3 4</td>
</tr>
<tr>
<td>sample number</td>
<td>11 11 6 21</td>
</tr>
</tbody>
</table>
### Table V. Summary of State Changes from Phase 3 to Phase 4

<table>
<thead>
<tr>
<th>Changes Phase 3 to phase 4</th>
<th>conditions 1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>sample number</td>
<td>10</td>
<td>3</td>
<td>4</td>
<td>8</td>
</tr>
</tbody>
</table>

G. Determine the Congestion Pre-identification Number

We can see from the TABLE IV. and TABLE V. that the phases of traffic state after changing do not always continue, some may be restored within a short period of time, which are proposed before. However, the changes that become more severe phase and continue 3 times or more may will continue for a long time, such as condition 4 in the changes of phase 2 to phase 3 and phase 3 to phase 4, the road itself is difficult to adjust, so the case should be promptly to the traffic state of prejudgment and management. To sum up, this paper thinks that traffic managements can forecast three state samples in the next time and identify the phases they belong to, if the three state samples are identified as congestion or severe congestion phases, which represent the jam will last for a long time and should be taken corresponding control measures.

H. Determine Congestion Identification Threshold

In this paper, the samples of state changes from phase 2(3) to phase 3(4) which continue for 3 times or more are counted, and Fig. 3 (Fig. 4) shows the scatter diagram of the samples belongs to phase 2(3) and samples will change from phase 2(3) to phase 3(4).

**Figure 3. Phase changes in phase 2 sample points distribution**

**Figure 4. Phase changes in phase 3 sample points distribution**

From Fig. 3 and Fig. 4, we can see that samples will change from phase 2(3) to phase 3(4) are highly dispersed, so it is difficult to determine the specific threshold. In this paper, we argue that the dynamic analysis can be carried out with emphasis on the status of the phase 2 sample points, combining with forecast of the downstream traffic flow three times, we can identify the states of the 3 times to know whether congestion will occurs or congestion will be serious.

IV. Conclusion

This paper firstly analyzes the dynamic traffic state identification methods based on the existing FCM algorithm, and proposes an improved FCM algorithm to optimize identification model, using partition entropy to evaluate the algorithm; secondly, this paper establishes the traffic state identification model based on DEA to identify samples in the same state phase, which is difficult for FCM algorithm; thirdly, the congestion formation processes are analyzed by timing sequence, and we find the number of pre-identification for congestion is better than congestion identification threshold, due to the chaos characteristic of traffic states.

### References


