PAPER • **OPEN ACCESS**

Genetic and Simulated Annealing Algorithms-based Traffic State Identification

To cite this article: Yuan Yuan *et al* 2018 *IOP Conf. Ser.: Earth Environ. Sci.* **170** 032117

View the article online for updates and enhancements.

IOP Conf. Series: Earth and Environmental Science **170** (2018) 032117 doi:10.1088/1755-1315/170/3/032117

Genetic and Simulated Annealing Algorithms-based Traffic State Identification

Yuan Yuan[1, 2, a], Chunfu Shao[1, *], Shiping Zhang[2, b], Lamtam Mei[3, c]

[1]MOE Key Laboratory for Urban Transportation Complex Systems Theory and Technology, Beijing Jiaotong University Beijing 100044, China
[2]School of Automotive and Transportation Engineering, Shenzhen Polytechnic, Shenzhen 518055, China
[3]Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hong Kong, China

*Corresponding author e-mail: cfshao@bjtu.edu.cn, [a]peggyyuan@szpt.edu.cn, [b]754642483@qq.com, [c]karen.tam.mei.lam@polyu.edu.hk

Abstract. Accurate and scientific traffic state identification is the basis of traffic navigation system, traffic control system and traffic organization optimization. In this paper, the dynamic traffic data collected by geomagnetic detector are firstly used to identify traffic state. We proposed SAGA-FCM clustering algorithm which is combined simulated annealing algorithm (SA) with genetic algorithm (GA) for urban expressway traffic state identification. This method can overcome the problems those the dynamic data of other detectors are not accurate and the time interval is not uniform. Meanwhile, it can overcome the instability of FCM algorithm clustering center and it is easy to fall into the local extreme value and "premature" problem of genetic algorithm. Research results show that, Compared with FCM algorithm and GA-FCM algorithm, SAGA-FCM clustering algorithm can be more effective and fast convergence, so as to improve the accuracy of urban traffic state identification.

1. Introduction

With the improvement in urbanization level, the explosive increase in motor vehicle population and road traffic flow turns traffic congestion into a common problem in big cities [1]. Accordingly, accurate and rational traffic state identification should be the foundation for traffic navigation system, traffic control system and traffic organization optimization.

Automatic traffic state identification, also known as automatic traffic congestion identification, involves direct method and indirect method. Direct method is a video image recognition-based state identification method, while indirect method is intended to indirectly identify traffic state based on the traffic flow parameter data obtained with detector[1-4].

Ruspini took the lead in describing and systematically studying fuzzy clustering method[5]. A lot of researchers make much account of efficient clustering algorithm during data mining[6]. Clustering analysis is a method, where a few groups are so automatically established based on the metrics of correlation between samples that the samples within the same group are similar to each other while the ones from different groups are different from each other. This method has been extensively used for data mining, image processing and pattern recognition etc. Jiang Guiyan[3] reported traffic state identification performed through fuzzy C-means clustering (FCM) based on the flow, speed and occupancy data acquired with loop-coil detector. Since FCM algorithm is substantially designed to seek optimal solution through gradient descent, local convergence is inevitable[6]. To make up for the deficiency of FCM algorithm in traffic state identification, some scholars introduced some evolutionary algorithms to fuzzy clustering and developed new clustering methods for traffic state identification; for instance, Yang Zuyuan[4] fused shuffled frog leaping algorithm (SFLA) with fuzzy C-mean algorithm (FCM) based on data acquired with microwave detector to have established the SFLA -FCM clustering algorithm, which effectively prevents FCM being sensitive to initial value and prone to local minimum. Yue Li[7] performed traffic state identification based on speed, flow and occupancy data as acquired with microwave detector, having prevented FCM algorithm from being excessively sensitive to initial clustering center, and identified the optimal number of clusters and the m value of weight coefficient. The integration of multiclass support vector machines with fuzzy C-means clustering (FCM) proposed by Ma Faping [2] enables intelligent analysis and discrimination of traffic state with an accuracy rate as high as 95%.

As a new-style traffic flow detection device, geomagnetic detector is easy to install and weather-proof, brings about slight damage to ground, and acquires traffic flow parameters in a complete, accurate and continuous manner [8]. However, only a few studies on traffic state identification address the data of geomagnetic detector.

First, this paper performs comparative analysis of the features of different traffic detectors within the same road segment, and gain traffic flow parameter data through detectors that could accurately and completely acquire data at uniform time interval. Then, the SAGA-FCM clustering algorithm that fuses simulated annealing algorithm (SA) with genetic algorithm (GA) is used to identify traffic state of urban expressway. Genetic coding method and fitness function are designed based on physical circumstances of clustering problem to overcome the algorithm instability and frequent orientation towards local optimum resulting from the randomness of initial clustering center selection in original fuzzy C-means clustering; simulated annealing algorithm is integrated with genetic algorithm to prevent premature convergence in traditional genetic algorithm. Finally, geomagnetic data is used to verify the effectiveness of the method stated in this paper, and compare that method with FCM algorithm and GA-FCM algorithm.

2. Comparison between traffic detectors and selection

The accurate identification of traffic state is dependent on dynamic basic traffic information, which reflects time-varying traffic condition. Dynamic basic traffic information must be acquired in a timely, accurate and complete manner. From this point of view, it's important to choose appropriate detector for acquisition of dynamic traffic information.

Along with the instant development of detector technology, more and more vehicle detectors have been used to collect dynamic basic traffic information[1][3-4][7][9], including loop-coil detector, GPS equipped floating vehicle, microwave detector, and video detector. All these detectors have both advantages and disadvantages. This paper compares the traffic parameter data acquired with fixed traffic detection device arranged at a cross section of an urban expressway with the traffic parameter data acquired with GPS equipped floating vehicle arranged in that segment of road. Fixed traffic detection devices include loop coil, geomagnetic detector and license plate recognition detector; GPS equipped floating vehicles include bus and taxi.

The comparison between traffic parameter data acquired with principal detectors is shown in Table 1:

IOP Conf. Series: Earth and Environmental Science **170** (2018) 032117 doi:10.1088/1755-1315/170/3/032117

Table 1. Comparison between traffic parameter data acquired with principal detectors

Detector technology	Traffic volume	Occup -ancy	Vehicle speed	Accu-racy	Data integrity	Acquisition interval	Other parameters
Video image (license plate recognition)	√	√	√	*	NO	1 minute	License plate number, detector ID, time, time headway (time consumed, traffic volume, and license plate number)
Loop-coil detector	√	√	√	√	NO	5 minutes	Detector ID, time, and time headway
Geomagnetic detector	√	√	√	√	√	1 minute	Detector ID, lane number, time, time headway and vehicle type etc.
GPS equipped floating vehicle	*	*	√	*	*	Uneven	Acquisition time, i.e. the travel time consumed to indirectly attain the data of a certain car

Note: √ - Direct detection with extremely complete and accurate data; * - Indirect detection with relatively complete and accuracy data; NO - Detection failure, serious data missing and extremely inaccurate data

To sum up, geomagnetic detector is superior to other detectors in terms of accurate and complete acquisition of basic parameters data (flow, speed and density) of traffic flow at uniform interval. As a new traffic information detection technology, it features outstanding interference immunity, stable and reliable operation, and suitability for a variety of complex weathers. Therefore, this paper studies section traffic state identification with the dynamic traffic information acquired with geomagnetic detector as object of study.

3. Saga-fcm algorithm

Fuzzy clustering is one of the important research branches in a many fields like knowledge discovery and pattern recognition etc. In company with the expansion of research scope, both scientific research and practical application raises higher requirement for clustering result in many aspects. As a currently popular type of clustering, fuzzy C-means clustering (FCM) employs the concept of determination of "geometric closeness" at data point in Euclidean Space, allocates the data to different clusters, and determines the distance between clusters. Fuzzy C-mean clustering algorithm lays a foundation for other fuzzy clustering analysis techniques theoretically and practically, and has the most extensive application. FCM algorithm, however, is substantially a local search optimization algorithm, which may converge to local minimum point in case of the improper selection of initial value. Accordingly, this shortcoming of FCM algorithm restricts its application [9].

3.1. Simulated Annealing Algorithm

This paper combines simulated annealing algorithm and genetic algorithm (SAGA) for clustering analysis; simulated annealing algorithm and genetic algorithm draw on each other's merits, thereby effectively overcoming the premature convergence of traditional genetic algorithm; additionally, genetic coding method and fitness function are designed based on physical circumstances of clustering problem so that this algorithm could more effectively and rapidly converge to global optimal solution.

Simulated annealing algorithm was successfully applied to combinatorial optimization in 1983 by finding the global optimal solution or approximate global optimal solution to optimization problem through the simulation of annealing process of high-temperature object. First, it generates an initial solution as current solution, and then selects a non-local optimal solution in neighboring region of current solution using probability P(T), and repeats such solution, thereby avoiding the evolution towards local optimum. In the beginning, objective function is allowed to evolve towards the direction of increase (corresponding to the increase in energy) along with parameter adjustment so as to facilitate the departure from local minimum area. Along with the reduction in fictive temperature (corresponding to annealing of object), the system activity decreases; finally, probability 1 is stably realized in the global minimum area.

Simulated annealing algorithm is described below [10]:

(1) Take S_0 as initial state; let $S(0) = S_0$, and assume initial temperature is T; let i=0;

(2) Let $T = T_i$, use T and S_i to call Metropolis sampling algorithm, and return state S as its current solution, $S_i = S(0)$;

(3) Reduce the temperature in a certain way, i.e. $T = T_i + 1$, where $T < T_{i+1}$, $i = i+1$;

(4) Check termination condition; if such condition is satisfied, go to step (5); otherwise, go to step (2);

(5) The current solution S_i is optimal solution; output the result and stop.

Metropolis sampling algorithm is described below:

(1) Let current solution $S(0) = S_0$ when k=0 , and perform the following operations at temperature T;

(2) Generate a neighbor subset $N(S(\mathrm{k})) \in S$ based on the state S of current solution $S(\mathrm{k})$ in a certain specified way, get a new state S' randomly from $N(S(\mathrm{k}))$, take it as next candidate solution, and calculate the energy $\Delta C' = C(\mathrm{S'}) - \mathrm{C(S(k))}$;

(3) If $\Delta C' < 0$, accept S' as next current solution; otherwise, accept S' as next current solution with probability $\exp(-\Delta C'/T)$. If S' is accepted, let $\mathrm{S(k+1)} = S'$; otherwise, $\mathrm{S(k+1)} = S(\mathrm{k})$;

(4) If $\mathrm{k} = \mathrm{k}+1$, check if the algorithm meets termination condition; if such condition is satisfied, go to step (5); otherwise, go to step (2);

(5) Return $S(\mathrm{k})$, and stop.

3.2. Genetic Algorithm

Genetic algorithm is principally composed of the following five steps[10]:

(1) Coding method: In genetic clustering algorithm, the parameter to be optimized is c initial clustering centers; binary coding is used here, where each chromatosome is composed of c clustering centers; for m-dimension sample vector, the number of variables to be optimized is $c \times m$. Assuming each variable employs k-bit binary encoding, the chromatosome should be a binary code string of $c \times m \times k$ in length;

(2) Fitness function: The measure for individual fitness is fitness function, of which the role is similar to the measure for the ability of creatures to adapt to the environment in the world of nature. Artificial objective function is obtained based on Equation (20-1) for each individual; its value is inversely proportional to individual fitness value; hence, fitness function is sorted fitness assignment function: $FintV = ranking\left(J_b\right)$;

(3) Selection operator: Selection operator is subjected to stochastic universal sampling (sus);

(4) Crossover operator: The most straightforward single-point crossover operator is employed;

(5) Mutation operator: Generate the number of mutant genes at a certain probability, and select mutated genes by stochastic method. If the code of selected gene is 1, the number becomes 0; otherwise, it becomes 1.

3.3. Simulated Annealing Genetic Algorithm-based Fuzzy C-means Clustering

Modification of the clustering centers and degrees of membership is the core of FCM. When the algorithm is convergent, clustering centers a_i can be found, and degrees of membership μ_{ij} can be calculated. Therefore, fuzzy clustering result can be achieved [1].

The procedure of simulated annealing genetic algorithm-based fuzzy C-means clustering is as follow:

(1) Initialization control parameter: Population individual size (sizepop); maximum number of generations (MAXGEN); crossover probability; mutation probability (P_m); initial annealing temperature; cooling factor (T_0); ending temperature (T_{end});

(2) Perform random initialization of c clustering centers, and generate initial population Chrom; calculate the membership degree of each sample and the individual fitness value f_i for each clustering center, where, $i = 1, 2, \ldots,\ sizepop$;

(3) Let cycle count variable gen=0;

(4) Subject population Chrom to genetic operations like selection, crossover and mutation etc., and work out the c clustering centers, the membership degree of each sample, and the individual fitness value f_i' for new individuals. If $f_i' > f_i$, replace old individuals with new ones; otherwise, accept new individuals with probability $P = \exp((f_i' > f_i)\mathrm{T})$ and abnegate old individuals;

(5) If gen<MAXGEN, then gen = gen+1, when it's necessary to go to step (4); otherwise, go to step (6).

(6) If $T_0 < T_{\text{end}}$, the algorithm comes to a successful end, when it's possible to return the global optimal solution; otherwise, perform cooldown operation and go to step (3).

4. Calculation example analysis

4.1. Data Selection

The geomagnetic data of occupancy (O), flow (Q) and speed (V) are highly sensitive to congestion time and degree, thus satisfactorily reflect the variation of road traffic state during occurrence, persistence and dissipation of congestion. Therefore, the above-noted 3 input variables are used for traffic state identification.

This paper takes an interchange front section of a certain expressway in a certain city as the object of research. The data was acquired at an interval of 5 minutes. The research span was 0:00-24:00 every day during period from January 5 to 11, 2015, where (288×7=)2016 sets of data were collected, and completely covered the four traffic flow states, i.e. "clear", "slow", "congested" and "seriously congested", on working days and rest days.

Geomagnetic detection device generates data for each vehicle, including mean lane speed, lane traffic within period, and lane occupancy within period. Since the study on individual lanes is not representative in road traffic state identification, this paper slightly improved the original traffic parameters when selecting parameters. Take integrated account of single-lane section mean speed v_s , single-lane section flow q_s and single-lane section mean occupancy o_s .

Their computing formulas are as follows:

Single-lane section mean speed:

$$v_s = \frac{\sum_{i=1}^{n} v_i}{\sum_{i=1}^{n} q_i} \tag{1}$$

Individual lane flow:

$$q_s = \sum_{i=1}^{n} q_i \tag{2}$$

Mean individual lane occupancy:

$$o_s = \frac{\sum_{i=1}^{n} o_i}{n} \tag{3}$$

Where, v_i, q_i and o_i represent the speed, flow and occupancy of individual lanes within detection period, respectively.

4.2. Computation Result and Analysis

Clustering is performed with SAGA-FCM algorithm respectively for working days (January 5-9) and rest days (January 10-11); the clustering result for January 5 and 10 is shown in Fig.1, and the clustering center is shown in Tables 2 and 3.

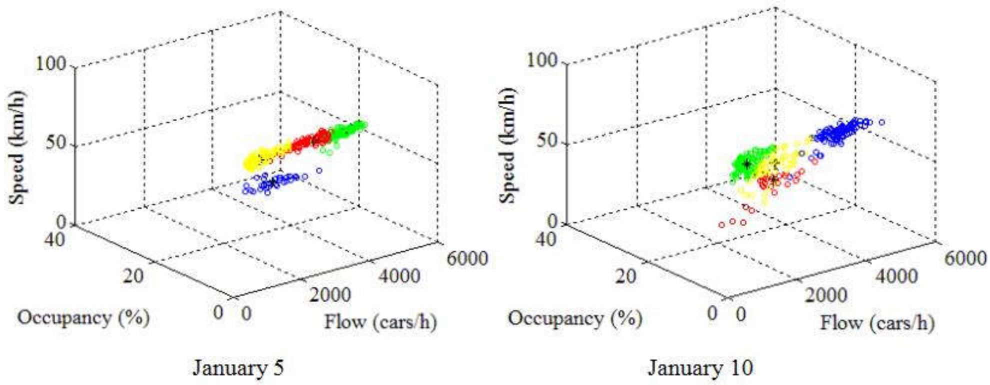

Fig. 1 SAGA-FCM clustering result on working days and rest days

Table 2. SAGA-FCM clustering center for January 5

Service level	Flow (cars/h)	Speed (km/h)	Occupancy (%)
Very clear	699	78.53	1.29
Clear	1791	69.72	3.77
Slightly congested	4222	69.07	8.64
Congested	3656	29.11	20.29

Table 3. SAGA-FCM clustering center for January 10

Service level	Flow (cars/h)	Speed (km/h)	Occupancy (%)
Very clear	961	80.1	1.79
Clear	3002	75.52	5.78
Slightly congested	4162	71.15	8.2
Congested	3642	26.9	21.67

As shown in Fig. 1, the traffic condition is more complicated on working days than on rest days, and the clustering center on rest days is more obvious. According to Tables 2 and 3, the clustering center for working day is slightly different from that for rest day.

4.3. Result Comparison between Algorithms

4.3.1. Precision Analysis. Take Equation (5) as estimated value of error rate.

$$a^* = \frac{n_1^* + n_2^* + n_3^* + n_4^*}{n_1 + n_2 + n_3 + n_4} \times 100\% \tag{4}$$

This paper divides geomagnetic detector data set X into 4 categories, i.e. c_1, c_2, c_3 and c_4; assuming c_1 contains n_1 samples; c_2 contains n_2 samples; c_3 contains n_3 samples, and c_4 contains n_4 samples; the number of errors for each category is denoted as n_1^*, n_2^*, n_3^* and n_4^*.

FCM algorithm is slightly different from GA-FCM and SAGA-FCM in respect of traffic state identification result; the traffic state identification result of working day is significantly different from that of rest day: The working days with clear traffic state are less than rest days, while the working days with congested traffic state are more than rest days.

For comparative analysis of the effect of three traffic state identification algorithms, i.e. FCM algorithm, GA-FCM algorithm and SAGA-FCM algorithm, this paper comprises cross assessment of error rate for Algorithms 1, 2 and 3 at a time interval of 5 minutes. According to Table4, the traffic state identification outcomes of FCM, GA-FCM and SAGA-FCM algorithms are partially different from each other; as shown in Equation 5 and Table 4, the error rates of FCM, GA-FCM and SAGA-FCM for working day are 5.14%, 0.35% and 0.21% respectively, so the traffic state identification accuracy of SAGA-FCM algorithm for working day is higher than that of FCM algorithm and GA-FCM algorithm; the error rates of FCM, GA-FCM and SAGA-FCM for rest day are 0.17%, 0% and 0.34% respectively, so the traffic state identification accuracy of SAGA-FCM algorithm for rest day is slightly inferior to that of FCM and GA-FCM algorithms, which means FCM and GA-FCM are more suitable for traffic state identification in case of uncomplicated traffic flow change, while SAGA-FCM is suitable for traffic state identification in case of complicated traffic flow change; as a whole, the error rates of SAGA-FCM algorithm and GA-FCM algorithm are 0.25%, while the error rate of FCM algorithm is 3.72%; with respect to overall accuracy, SAGA-FCM algorithm is remarkably superior to FCM algorithm and substantially equal to GA-FCM algorithm; furthermore, SAGA-FCM algorithm is superior to GA-FCM algorithm in case of complicated traffic flow change.

Table 4. Comparison in terms of traffic state identification result

Algorithm	Number of classification samples	Traffic state			
		Very clear	Clear	Slightly congested	Congested
FCM	2016	683	396	598	339
SAGA-FCM	2016	659	453	645	259
GA-FCM	2016	657	455	645	259

Table 5. Comparison between SAGA-FCM, GA-FCM and FCM in terms of error rate

<table>
<tr><th rowspan="2">Date</th><th rowspan="2">Category</th><th colspan="3">Number of errors</th><th colspan="6">Error rate</th></tr>
<tr><th>SAGA-FCM</th><th>GA -FCM</th><th>FC-M</th><th>SAGA-FCM</th><th>Tota-l</th><th>GA-FCM</th><th>Tota-l</th><th>FCM</th><th>Tota-l</th></tr>
<tr><td>January 5</td><td>Working day</td><td>0</td><td>1</td><td>1</td><td rowspan="5">0.21%</td><td rowspan="7">0.25%</td><td rowspan="5">0.35%</td><td rowspan="7">0.25%</td><td rowspan="5">5.14%</td><td rowspan="7">3.72%</td></tr>
<tr><td>January 6</td><td>Working day</td><td>2</td><td>1</td><td>19</td></tr>
<tr><td>January 7</td><td>Working day</td><td>1</td><td>2</td><td>0</td></tr>
<tr><td>January 8</td><td>Working day</td><td>0</td><td>1</td><td>40</td></tr>
<tr><td>January 9</td><td>Working day</td><td>0</td><td>0</td><td>14</td></tr>
<tr><td>January 10</td><td>Rest day</td><td>1</td><td>0</td><td>1</td><td rowspan="2">0.34%</td><td rowspan="2">0.00%</td><td rowspan="2">0.17%</td></tr>
<tr><td>January 11</td><td>Rest day</td><td>1</td><td>0</td><td>0</td></tr>
</table>

4.3.2. Search Performance. Fig. 2, 3 and 4 show the objective function value convergence curves of SAGA-FCM, GA-FCM and FCM. Three experiments were carried out for each algorithm .The comparison between Fig. 2, 3 and 4 shows that the convergence of FCM is slower than that of SAGA-FCM; after three tests on FCM, complete convergence is realized after 19 iterations, GA-FCM after 16 iterations, and SAGA-FCM after 12 iterations; objective function is used to calculate individual cluster values, which are inversely proportional to individual fitness. The comparison between Fig. 2 and 3 shows that the mean objective function value of FCM is significantly greater than mean objective function value of SAGA-FCM, which means the global search performance of SAGA-FCM and SAGA is improved as compared with FCM, and avoids local convergence to a certain extent; it is observed from Fig. 3 and 4 that the objective function value of SAGA-FCM is close to that of GA-FCM, but SAGA-FCM is more stable and effectively overcomes the premature convergence of traditional genetic algorithm; due to the fact that the initial clustering center of FCM algorithm is randomly allocated, the objective convergence value is unstable; when it comes to SAGA-FCM, due to the fact that the optimal clustering center obtained through global search of genetic algorithm is taken as initial value of initial clustering center, the result of SAGA-FCM algorithm is near optimal objective value after each execution; as a result, the algorithm stability has been remarkably improved, for example: SAGA-FCM does not change any more substantially after the 13th iteration, while FCM does not change any more after approx. 17th iterations. For FCM, the optimal value starting point varies with search, but when it comes to SAGA-FCM, the optimal value starting point substantially remains unchanged for each search.

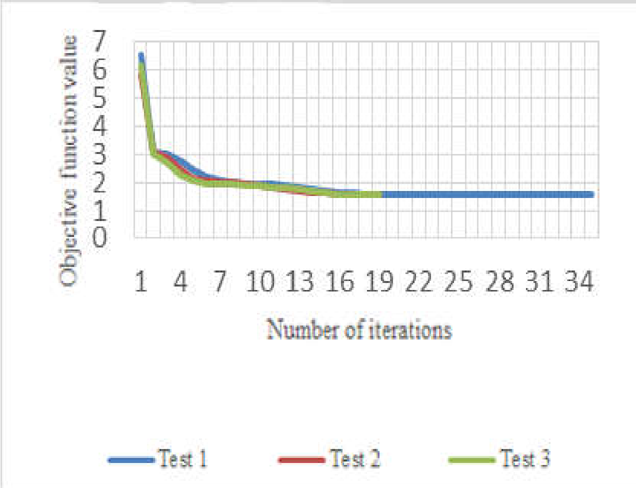

Fig. 2 Objective function value convergence curve of FCM

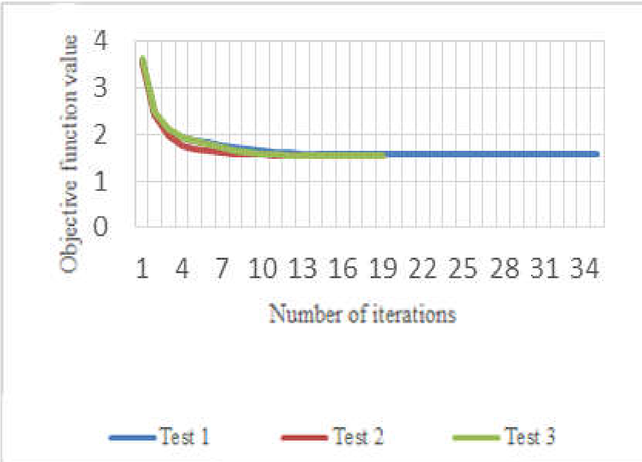

Fig. 3 Objective function value convergence curve of SAGA-FCM

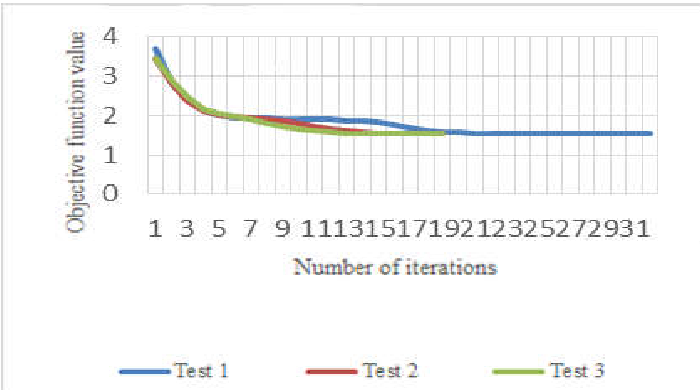

Fig. 4 Objective function value convergence curve of GA-FCM

5. Conclusion

This paper uses the combination of simulated annealing algorithm and genetic algorithm for purpose of fuzzy C-means clustering, and effectively and efficiently identifies traffic states by taking advantage of the strong local search capability of simulated annealing algorithm and the strong global search capability of genetic algorithm. The concrete conclusion reached is as follows:

1. Upon the traffic state's being divided into clear, slow, congested and seriously congested states, cross-over identification is used to determine the error rates of SAGA-FCM algorithm and GA-FCM algorithm (0.25%) and FCM algorithm (3.72%), according to which SAGA-FCM algorithm is remarkably superior to FCM algorithm and substantially equal to GA-FCM in terms of overall accuracy; additionally, SAGA-FCM algorithm is superior to GA-FCM algorithm in case of complicated traffic flow change.
2. In average, the convergence iteration speed of SAGA-FCM is 4 times faster as compared with FCM and GA-FCM.
3. SAGA-FCM is more stable than FCM and GA-FCM, because its results stably gather around optimal value, i.e. 1.57.
4. The global search performance of SAGA-FCM is significantly improved when compared with FCM and GA-FCM; the optimal value starts from around 3.7, which helps to avoid local convergence and the premature convergence of GA-FCM to a certain extent.
5. Comparative analysis is performed over different traffic detectors, and the traffic data acquired with geomagnetic detector is used for traffic state identification to attain satisfactory result.

Despite the improvement of traffic state identification accuracy by SAGA-FCM algorithm, FCM algorithm may bring about such problems as local optimization and evenness; simulated annealing method could avoid the problems of FCM algorithm, but its search result exhibits randomness, which leads to a high error rate up to 12.5% when traffic state is divided into five levels. Genetic and simulated annealing algorithms will be integrated with particle swarm optimization algorithm for FCM clustering in future, and the particle swarm optimization algorithm will be embedded based on genetic simulated annealing algorithm to establish genetic and simulated annealing-particle swarm hybrid optimization algorithm. Hybrid optimization algorithm is used to perform particle swarm optimization algorithm search within temperature range of simulated annealing so as to avoid probabilistic kicking of simulated annealing and prevent particle swarm optimization algorithm from getting trapped in local extreme point. This algorithm is compared with current algorithms and expected to be more satisfactory. In addition, in view of the fact that the traffic state identification is currently performed based on cross-sectional geomagnetic data, road-segment traffic state identification is expected to be performed with multi-source detectors along the segment of road so as to provide travelers with more accurate and effective information.

Acknowledgements
This paper was supported by the Hebei Natural Science Foundation of China (No. E2016513016).

References
[1] Hua-pu Lu, Zhi-yuan Sun, and Wen-cong Qu (2015). Big Data-Driven Based Real-Time Traffic Flow State Identification and Prediction [J]. Discrete Dynamics in Nature and Society, Volume 2015, Article ID 284906, 11 pages.
[2] Ma Faping (2014). Research on urban road traffic state identification method based on support vector machine [D]. Qingdao: Qingdao University of Science & Technology, 2014.
[3] Jiang Guiyan, Guo Haifeng, Wu Chaoteng et al (2008). Urban road traffic state identification method based on induction coil data [J]. Journal of Jilin University (Engineering Science), 2008 (2):37-42.
[4] Yang Zuyuan, Xu Jiao, Luo Bing, Du Changhai (2010). Study on urban traffic state identification based on F L A-F CM [J]. S computer application research, 2010 (27), 1743-1745.
[5] EH Ruspini (1970). Numerical methods for fuzzy clustering [J]. Information Sciences, 1970, 2 (3):319-350.
[6] Baoju Zhang, Shan Qin, Wei Wang, Dan Wang, Lei Xue (2016). Data stream clustering based

on Fuzzy C-Mean algorithm and entropy theory [J]. Signal Processing, (2016), 111–116.
[7] Yue Li (2013). Study on dynamic identification of urban expressway traffic status based on FCM [D].Chengdu: Southwest Jiao Tong University, 2013.
[8] Cai Xiaoqing (2012). Research on Key Technologies of traffic system identification based on magnetic induction [D]. Guangzhou: South China University of Technology, 2012.
[9] Toru Seo, Takahiko Kusakabe,Yasuo Asakura (2015). Estimation of flow and density using probe vehicles with spacing measurement equipment [J].Transportation Research Part C53, (2015): 134–150.
[10] Shi Feng, Wang Hui, Hu Fei, Yu Lei. MATLAB intelligent algorithm 30 case analysis [M]. Beijing: Beihang University press, 2011.