基于 FCM 城市区域道路交通状态时空分层判别方法*

董红召，马　帅，郭明飞
（浙江工业大学 智能交通联合研究所，杭州 310014）

摘　要： 为了解决城市区域路网交通状态的时空分析问题，提出了一种基于模糊 C 均值聚类（FCM）道路交通状态判别模型及分析方法。通过路网的空间单元交通状态的定量分析和对大量的历史数据进行 FCM 分析，挖掘出各空间单元的各类交通状态的聚类中心，并将实时采集的交通数据与聚类中心进行匹配，评判其实时交通状态，最后根据空间单元在路网空间分布，获得各状态下点、线、面的空间分层分析结果。实例结果表明，判别方法能准确地实现区域路网的交通状态时空判别，为交通精细化管理提供辅助决策信息。

关键词： 交通状态；时空分析；模糊 C 均值聚类
中图分类号： U491.3　　**文献标志码：** A　　**文章编号：** 1001-3695(2012)04-1263-04
doi:10.3969/j.issn.1001-3695.2012.04.017

Spatial and temporal model for urban regional traffic state analysis based on fuzzy C-means clustering

DONG Hong-zhao, MA Shuai, GUO Ming-fei
(*ITS Joint Institute*, *Zhejiang University of Technology*, *Hangzhou* 310014, *China*)

Abstract: This paper developed a spatial and temporal hierarchical model based on fuzzy C-means clustering for area traffic state analysis to predict area traffic states. Firstly, it quantitatively analyzed the traffic state of the unit of the road network. Then gained the cluster centers of each traffic state for the unit of the road network based on the method of fuzzy C-means clustering. After this, it recognized and classified the real-time traffic state combining its real-time traffic data. According the units' space distribution in the network, the model presented the different kinds of spatial distribution under different traffic states. The result of example proves that this analytical method obtain the spatial and temporal traffic state accurately. It also supplies the assistant decision-making information for transportation system managers.

Key words: traffic state; temporal and spatial analysis; fuzzy C-means clustering

0　引言

对城市道路交通状态准确的把握，是实行高效的精细化交通管理与控制、缓解车路资源矛盾的重要基础。目前国内外研究大部分集中在对单一路口或路段的交通数据通过模式识别、人工智能及数据挖掘算法进行判别[1~3]，这类方法提高了交通状态判别的精度和效率，但避开了交通路网空间建模的复杂性，不能体现隐含的交通状态的空间关联信息；张和生等人[4~6]提出了一种区域交通状态时空分层模型及其建立方法，结合路段和交叉口的物理邻接关系，得到了交通状态参数的不同预设阈值下交通状态分层模型，但对区域路网采用了统一的判别标准。由于交通状态是交通设施提供的交通供给和交通行为产生交通需求共同作用的外在表现，并且各路段与各交叉口之间的静态特征和动态交通控制参数存在差异，对路网各路段和各交叉口采用统一判别标准可能会存在偏差，因此交通状态的参数选取和判别过程应考虑上述两类因素。

城市道路采集的大量交通流数据，隐含着能真实、客观反映交通运行状态时空变化的信息。交通状态判别实质即将交通流数据分类的过程，但每种交通状态没有绝对的界限。模糊 C 均值聚类（FCM）是一种比较理想的对数据进行软划分的方法，因此采用该方法实现对交通状态的判别。城市路网由若干信号交叉口及相邻上下游交叉口之间的路段组成，将其定义为路网空间单元，并作为区域交通状态判别的最基本物理单元。这两类空间单元静态属性及所分担的交通流的特点不尽相同，因此将历史数据与实时交通流数据相结合作为输入，利用 FCM 判断其实时交通状态。最后根据空间单元之间空间关系，分析各类交通状态子集下的空间分布情况。

1　路网空间单元交通状态参数定量分析

为了更准确地描述城市路网交通运行状态，需对路网空间单元分别选取合适的交通状态参数进行定量分析。车辆在城市道路行驶过程一般可分为两个阶段：路段行驶和通过交叉

收稿日期： 2011-09-06；**修回日期：** 2011-10-28　　**基金项目：** 国家自然科学基金资助项目（61174176）；浙江省科技公益项目（2010C33245）；特种装备制造与先进加工技术教育部/浙江省重点实验室开放基金；杭州市社会发展科研专项资助项目（20110533B02）

作者简介： 董红召（1969-），男，教授，博士，主要研究方向为智能交通机电系统（its@zjut.edu.cn）；马帅（1986-），女，硕士研究生，主要研究方向为交通数据时空分析；郭明飞（1977-），男，讲师，博士研究生，主要研究方向为智能交通系统.

口。交通流在区域路网这两类路网空间单元上分布,可分为连续性交通流和间断性交通流,因此描述这两类交通流的交通特性的参数也不尽相同。因考虑其单位的统一,拟选取平均延误及平均行程时间分别作为评判交叉口及路段的交通运行状态的参数。

1.1 交叉口交通状态定量分析

城市路网大部分均配有信号控制系统,交叉口延误大部分由信号控制引发,包括排队减速延误、排队移动延误、停车延误和启动加速延误。反映交叉路口交通运行状态的参数包括交叉口延误时间、停车时间、停车次数和排队长度等。因交叉口平均延误能客观、真实地反映交通需求、交通供给以及外界干扰(如非机动车流等)的综合影响下的交通运行状态,故选择该参数作为交叉口交通状态定量参数。

对于交叉口平均延误的计算[7],将交叉口的各进口道按服务于一个或几个交通流向的一条或多条车道的原则划分成车道组,车道组作为交叉口延误计算的最小单元;而后根据交叉口的几何条件及信号配时,分别计算其通行能力和饱和流率;最后将交通流检测器采集的交通流数据输入,获取各车道组实时的平均延误。

给定车道组的每辆车的平均延误 d(s/辆)由下式确定:

$$d = d_1 \times PF + d_2 + d_3 \tag{1}$$

$$d_1 = \frac{0.5C(1-\frac{g}{C})}{1-[\min(1,X)\frac{g}{C}]} \tag{2}$$

$$d_2 = 900T[(X-1)+\sqrt{(X-1)^2+\frac{8KIX}{cT}}] \tag{3}$$

其中:d_1 为假设车辆均匀到达下的均匀延误;PF 为均匀延误修正系数,修正信号联动对延误的影响;d_2 为增量延误,即代表车辆随机到达及过饱和排队的效果;d_3 为初始排队延误,即观测期之初的初始排队对观测期的到达车辆产生的影响;C 为信号周期;g 为该车道组的有效绿灯时间;X 为该车道组的 v/c;T 为观测期时长;I 为上游信号交叉口影响修正系数;c 为车道组的通行能力。

1.2 路段交通状态定量分析

在区域交通状态分析中,路段处于连接宏观路网参数和微观路口参数的位置。路段交通参数通常包括流量、占有率、路段行程时间、行驶速度等。在对路段交通状态定量分析时,由于要考虑交叉口的交通状态参数,需采用能够反映这两类空间单元的交通状态的公共变量。行程时间能够反映两类空间单元共同的交通状态,本文拟选取路段行程时间作为路段交通状态参数。目前城市道路的检测器获取的交通流数据为交通流率、车道占有率,因此将利用这两个参数对路段的行驶时间进行估计。BRP 模型[8]是最重要的行程时间估计模型,在该模型中,认为行程时间是流量与通行能力比值的非线性函数,其数学表达式为

$$t_a(t) = t_0[1+\alpha(\frac{q_a(t)}{C_a(t)})^\beta] \tag{4}$$

其中:$t_a(t)$是路段 a 时刻 t 的行程时间;t_0 是交通量为零时路段 a 的行程时间,即等于路段长度与道路的最大限制时速之比;$q_a(t)$是路段 a 时刻 t 的机动车交通流率;$C_a(t)$是路段 a 时刻 t 的实际通行能力;α、β 是参数,建议取 $\alpha=0.15$,$\beta=4$。

2 基于 FCM 路网空间单元交通状态判别

聚类一般是无指导的学习或分割,将数据划分或分割成相交或者不相交的群组过程,根据数据之间的预先指定特征属性的相似性完成聚类任务。FCM 聚类通过引入隶属度来表征每个数据所属各类的不同程度,对数据进行软划分。首先初始聚类中心,反复调整聚类中心,使各类中数据到每个聚类中心的距离之和最小或者满足终止条件。通过 FCM 聚类,获取各类别的聚类中心作为空间单元的各交通状态判断原型,然后输入实时交通数据,计算其与各聚类中心的最小欧几里德距离,判断出最小值,归其所类,实现对路网空间单元的状态判断。

2.1 样本特征属性确定

在时间尺度上对大量的城市交通历史数据分析可知,同一采集地点获取的交通流数据具有显著的周相似特性,如图 1 所示,即不同周同一工作日交通流变化趋势相似,而不同工作日的变化趋势相似性不高。

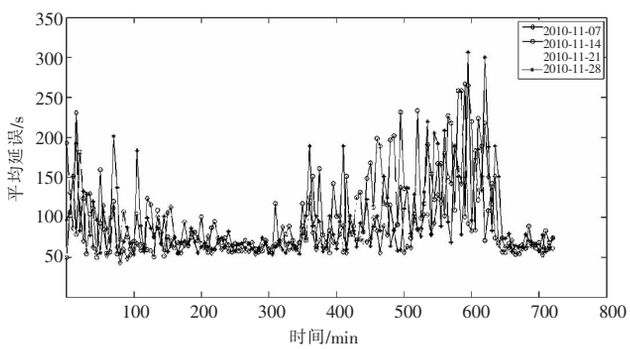

图 1 杭州市某路口周尺度下平均延误变化趋势

FCM 聚类分析样本的特征属性为每周不同工作日,即在时间尺度获得更为精细的判别原型,因此对路网空间单元 k 提取的历史数据样本集合为

$$S = (s_1, s_2, \cdots, s_n) = \begin{pmatrix} s_{11} & s_{12} & \cdots & s_{1n} \\ s_{21} & s_{22} & \cdots & s_{2n} \\ \vdots & \vdots & & \vdots \\ s_{71} & s_{72} & \cdots & s_{7n} \end{pmatrix} \tag{5}$$

其中:$s_i(i=1,2,\cdots,n)$取样时间段内以周为单位的时间序列下的交通状态数据样本;$s_{ki}(k=1,2,\cdots,7)$为采样日期(分别周一到周日)。样本的特征向量 $s_k=(s_{k1},s_{k2},\cdots,s_{k7})^{\mathrm{T}}$。

2.2 基于 FCM 交通状态判别

根据公安部《城市交通管理评价指标体系》将城市道路交通状态划分为通畅、正常、拥挤和堵塞四个等级,FCM 算法中的 c 为数据样本将分成类别的数目,因此 c 取为 4。FCM 具体算法如下:

设目标函数如式(6)所示。

$$J(U,c_1,c_2,\cdots,c_4) = \sum_{i=1}^{4} J_i = \sum_{i=1}^{4}\sum_{j=1}^{n} u_{ij}^m d_{ij}^2 \tag{6}$$

$$u_{ij} = \frac{1}{\sum_{k=1}^{4}(\frac{d_{ij}}{d_{kj}})^{2/(m-1)}} \tag{7}$$

$$d_{ij} = \| c_i - x_j \| \tag{8}$$

$$\sum_{i=1}^{4} u_{ij} = 1 \tag{9}$$

其中:$\forall j=1,2,\cdots,n$,$u_{ij}\in[0,1]$表示第 j 个数据点属于第 i 个

聚类中心的隶属度；c_i 为模糊组 i 的聚类中心，d_{ij} 表示第 j 个聚类中与第 j 数据点之间的欧几里德距离；$m \in [1, \infty)$ 是一个加权 i 指数。构造拉格朗日乘子，建立新的目标函数为

$$J(U, c_1, c_2, \cdots, c_4, \lambda_1, \lambda_2, \cdots, \lambda_n) = J(U, c_1, c_2, \cdots, c_4) + \sum_{k=1}^{n} \lambda_k \left(\sum_{i=1}^{4} u_{ij} - 1 \right) = \sum_{i=1}^{4} \sum_{j=1}^{n} u_{ij}^m d_{ij}^2 + \sum_{k=1}^{n} \lambda_k \left(\sum_{i=1}^{4} u_{ij} - 1 \right) \tag{10}$$

对所有输入参量求导，使原目标函数最小的必要条件为

$$c_i = \frac{\sum_{j=1}^{n} u_{ij}^m x_j}{\sum_{j=1}^{n} u_{ij}^m} \tag{11}$$

FCM 算法的迭代过程如下：

a）随机初始化四个数据聚类中心。

b）用式（7）计算矩阵 U。

c）用式（11）计算四个新的聚类中心 c_i。

d）根据式（10）计算目标函数，若小于某个确定的阈值，或相对上次目标函数改变量小于某个阈值，则算法停止；否则，返回 b）。

通过上述 FCM 聚类分析得到交通状态原型矩阵 P，其元素 p_{ij} 表示各类状态下相对应特征属性的聚类中心。

$$P = (p_1, p_2, p_3, p_4) = \begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ \vdots & \vdots & \vdots & \vdots \\ p_{71} & p_{72} & p_{73} & p_{74} \end{pmatrix} \tag{12}$$

最后将路网空间单元的定量分析后的实时交通状态参数 x_i（$i = 1, 2, \cdots, 7$，为采集的工作日）作为输入，按照最小欧几里德距离原则，根据该空间单元的交通状态原型矩阵，并按式（13）实现交通状态的判别：

$$f_k(x_i) = \min_{1 \leqslant j \leqslant 4} \{ d_{ij} \} = \min_{1 \leqslant j \leqslant 4} \{ \| p_{ij} - x_i \| \} \tag{13}$$

则待判定的交通样本 x_i 所属第 k 类交通状态。

3　区域路网交通状态空间分层分析

3.1　交通状态空间分布特性

区域路网各空间单元在同一时刻也会呈现出不同的交通状态，因此将处于同一交通状态的空间单元归类成集合，观察空间单元在路网上的空间分布特性。分布形式主要有点状、线状、面状三种形式。点状指空间单元是呈现孤立状，与其相邻的空间单元均处在不同状态；线状指空间单元之间相互连通呈现直线状；面状指空间单元相互交叉或重合，呈折线或网状分布。城市交通状态的空间分布由以上三种基本类型组成。在某一时刻某一交通状态下，可能存在其中一种，也可能三种同时存在，反映了交通流在城市路网上的空间分布特性。

3.2　区域路网模型及交通状态分层分析

区域路网交通状态空间分层分析是建立在空间单元交通状态判别的基础之上，从空间单元的空间地理位置及相互之间的物理邻接关系出发，通过立区域路网模型将各个空间单元的空间关联性建立起来。区域路网模型的建立以 m 个路口和路口间的路段组成的交通网络，其中路口集合表示为

$$N = \{ n_1, n_2, \cdots, n_m \} \tag{14}$$

路段集合表示为

$$L = \{ l_{ij} \mid l_{ij} = \langle n_i, n_j \rangle, i, j = 1, 2, \cdots, m \} \tag{15}$$

定义邻接矩阵 $A_C = \{ a_{ij} \}$ 表示路口之间的物理邻接关系：

$$a_{ij} = \begin{cases} 1 & n_i, n_j \in N; \text{and} \langle n_i, n_j \rangle \in L \\ 0 & n_i, n_j \notin N; \text{or} \langle n_i, n_j \rangle \notin L \end{cases} \tag{16}$$

利用交通网络的权矩阵来定义路网单元之间的邻接的地理方位，权矩阵的元素 $D_C = \{ d_{ij} \}$，元素的取值和物理意义

$$d_{ij} = \begin{cases} 0 & a_{ij} = 0 \\ 1 & a_{ij} = 1 \text{ 且 } n_i \text{ 与 } n_j \text{ 南北连接} \\ 2 & a_{ij} = 1 \text{ 且 } n_i \text{ 与 } n_j \text{ 东西连接} \end{cases} \tag{17}$$

区域路网交通状态分层分析过程是各状态的空间单元的子集进行空间关联分析，将处于同一状态的并彼此可达的空间单元根据邻接矩阵和权矩阵的逻辑运算逐一筛选出来，得到区域道路交通状态点、线、面空间分层分布，流程如图 2 所示。

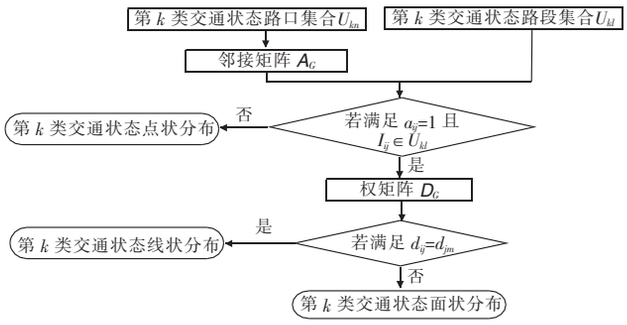

图 2　区域路网交通状态空间分层分析流程

4　实例和结果分析

本文以杭州市区 11 个路口及其之间的路段组成的区域为例，区域路网简化拓扑图如图 3 所示。以 2010 年 7 月的 31 天所采集实际交通流数据作为历史数据样本，以 2010 年 8 月 2 日（周一）8:45 实时采集数据作为输入进行空间单元的交通状态判断的结果如表 1 所示，而后根据其邻接矩阵和权矩阵，该区域交通状态分层结果如表 2 所示。

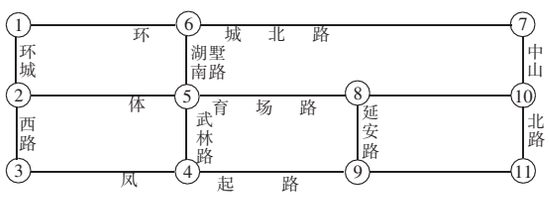

图 3　路网简化拓扑图

表 1　空间单元聚类原型及交通状态判断结果

路网单元	交通状态原型 $P_1(p_{11}, p_{12}, p_{13}, p_{14})$	判断结果	路网单元	交通状态原型 $P_1(p_{11}, p_{12}, p_{13}, p_{14})$	判断结果
n_1	(32,67,121,164)	3	$l_{3,4}$	(24,41,72,102)	2
n_2	(28,58,106,145)	3	$l_{4,5}$	(65,93,134,212)	3
n_3	(37,64,97,152)	2	$l_{5,6}$	(19,32,59,103)	3
n_4	(24,53,87,134)	2	$l_{6,7}$	(64,95,127,193)	2
n_5	(34,72,98,151)	3	$l_{5,8}$	(33,62,89,118)	2
n_6	(29,53,88,143)	3	$l_{4,9}$	(17,36,58,86)	2
n_7	(35,66,91,134)	2	$l_{8,9}$	(58,81,108,154)	2
n_8	(28,70,137,151)	2	$l_{8,10}$	(26,41,74,98)	3
n_9	(32,69,121,178)	3	$l_{7,10}$	(31,56,115,176)	2
n_{10}	(31,67,91,153)	3	$l_{9,11}$	(36,51,97,122)	3
n_{11}	(29,74,103,167)	2	$l_{10,11}$	(49,75,123,160)	3
$l_{1,2}$	(21,37,68,92)	3	$l_{1,6}$	(15,26,49,67)	3
$l_{2,3}$	(64,99,192,243)	3	$l_{2,5}$	(21,34,67,79)	3

注：交通状态类别 1-畅通 2-正常 3-拥堵 4-堵塞

从表 1 可知，通过 FCM 分析得出的各空间单元的聚类结

果中，路段 $l_{1,2}$（环城西路的环北路口—体育场路口路段）的畅通、正常、拥挤，拥堵的交通状态平均行程时间分别为 21 s、37 s、68 s、92 s，而路段 $l_{2,3}$（环城西路的体育场路口—凤起路口路段）的相同交通状态聚类中心分别为 64 s、99 s、192 s、243 s，即使在同一条主干道相邻路段的相同交通状态的聚类中心也存在较大的差异，因此若采用全局区域内同一的交通判断原型，可能会对某些空间单元的交通状态判别准确性有影响。交通管理部门发布的该区域内实际交通状态信息与其他方法进行的比对分析如表 3 所示。

表 2　区域道路交通状态分层结果

交通状态类型	正常	拥挤
点状	11	–
线状	3—4 6—7—10	4—9—11 1—6 2—5
面状	5—8—9	3—4 8—10—11

表 3　路网单元交通状态结果误差分析

项目	统一标准	FCM	FCM
判别参数	定量分析参数	交通流参数	定量分析参数
误判个数	5	4	3
误判率/%	19.23	15.38	11.54

最后由表 2 可知，同一时刻，城市区域内各交叉口及路段也非呈现一致交通状态，各交通状态下的空间分布（包括点状、线状、面状）具有独立分布特性。

5　结束语

本文以平均延误及行程时间为交通状态参数对路网空间单元的交通状态进行定量分析，并利用其历史数据结合周相似性采用 FCM 分析得出其交通状态判断原型，在此基础上实现了城市路网空间单元实时交通状态判别。实例分析表明，该计算方法能在一定程度上降低交通状态误判率。在此基础上，通过建立区域路网模型，从空间尺度上分析判断各种交通状态的实时空间分布结果，为交通管理者和参与者提供了更为直观、细致的交通信息。基于 FCM 道路交通状态时空分层判别方法，实现了对城市道路交通状态从定量到定性的分析，并把历史数据与实时数据相结合，提高了判断的准确率。该模型涉及微观、中观、宏观多层次的路网交通状态空间分布，能直观地反映出城市交通流的时空变化特性，可为城市路网交通控制与交通诱导等实际应用提供理论依据与决策支持。

参考文献：

[1] MONTAZERI-GH M，FOTOUHI A. Traffic condition recognition using the K-means clustering method [J]. Scientia Iranica B，2011，18(4)：930-937.

[2] 杨祖元，徐姣，罗兵，等. 基于 SFLA-FCM 聚类的城市交通状态判别研究[J]. 计算机应用研究，2010，27(5)：1743-1745.

[3] 董红召，吴方国. 基于 FCM 的路段平均行程时间估计[J]. 科技通报，2011，27(3)：426-431.

[4] 张和生，张毅，胡东成，等. 区域交通状态分析的时空分层模型[J]. 清华大学学报：自然科学版，2007，47(1)：157-160.

[5] 陈红，章渺，赵禹乔. 区域高等级公路网交通状态识别与分析模型研究[J]. 公路，2011，2(2)：100-106.

[6] 张和生，张毅，胡东成. 一种区域交通状态定量分析方法[J]. 吉林大学学报：工学版，2009，39(2)：336-342.

[7] 美国交通研究委员会. 道路通行能力手册[M]. 3 版. 任福田，等译. 北京：中国建筑工业出版社，1991.

[8] 姜桂艳，李继伟，张春勤. 城市主干路行程时间估计的 BPR 修正模型[J]. 西南交通大学学报，2010，45(1)：124-129.

（上接第 1233 页）

参考文献：

[1] HU Xiao-hui，EBERHART R. Solving constrained nonlinear optimization problems with particle swarm optimization [C]//Proc of the 6th World Multiconference on Systematics，Cybernetics and Informatics. 2002：203-206.

[2] SEDLACZEK K，EBERHARD P. Constrained particle swarm optimization of mechanical systems[C]//Proc of the 6th World Congresses of Structural and Multidisciplinary Optimization. 2005：1-10.

[3] PARSOPOULOS K E，VRAHATIS M N. Unified particle swarm optimization for solving constrained engineering optimization problems [C]//Lecture Notes in Computer Science，vol 3612. Berlin：Springer，2005：582-591.

[4] HE Qie，WANG Ling. A hybrid particle swarm optimization with a feasibility-based rule for constrained optimization[J]. Applied Mathematics and Computation，2007，186(2)：1407-1422.

[5] 王勇，蔡自兴，周育人，等. 约束优化进化算法研究及其进展[J]. 软件学报，2009，20(1)：2691-2709.

[6] MEZURA-MONTES E，COELLO C A C. A survey of constraint-handing techniques based on evolutionary multi-objective optimization，EVOCINV-04-2006[R]. Mexico：Departamento de Computacion，Evolutionary Computation Group at CINVESTAV，2006.

[7] DEB K. An efficient constraint handing method for genetic algorithms [J]. Methods for Applied Mechanical Engineering，2000，186(2-4)：311-338.

[8] MEZURA-MONTES E，COELLO C A C. A simple multimembered evolution strategy to solve constrained optimization problems [J]. IEEE Trans on Evolutionary Computation，2005，9(1)：1-17.

[9] 林丹，李敏强，寇纪淞. 基于遗传算法求解约束优化问题的一种算法[J]. 软件学报，2001，12(4)：628-632.

[10] KENNEDY J，EBERHERT R. Particle swarm optimization [C]//Proc of the 4th IEEE International Conference on Neural Networks. Piscataway：IEEE Press，1995：1942-1948.

[11] KARABOGA D. An idea based on honey bee swarm for numerical optimization[R]. Kayseri：Erciyes University，2005.

[12] BASTURK B，KARABOGA D. An artificial bee colony (ABC) algorithm for numeric function optimization[C]//Proc of IEEE Swarm Intelligence Symposium. 2006：3-4.

[13] KARABOGA D，BASTURK B，OZTURK C. Artificial bee colony (ABC) optimization algorithm for solving constrained optimization [C]//Advances in Soft Computing：Foundations of Fuzzy Logic and Soft Computing，vol 4529. 2007：789-798.