

文章编号: 1001-0920(2008)08-0915-04

## 融合粗糙集和 DS 方法的空中目标类型识别算法

丛 蓉<sup>1,2</sup>, 王秀坤<sup>1</sup>, 杨南海<sup>1</sup>, 孟祥宇<sup>2</sup>

(1. 大连理工大学 电子与信息工程学院, 辽宁 大连 116024; 2. 海军大连舰艇学院 教育技术中心, 辽宁 大连 116018)

**摘 要:** 通过分析粗糙集和 Dempster-shafe (DS) 方法在描述不确定信息上的相似性, 将这两种理论融合建立统一框架, 提出了融合粗糙集和 DS 方法的空中目标类型识别算法. 该算法通过知识库推导得出基本概率分配函数, 采用基于粗糙集的信任函数作为可信度量, 使得改进后的算法更有理论深度. 实验结果表明, 算法具有良好的运行效果.

**关键词:** 粗糙集; DS 证据理论; 目标识别

**中图分类号:** TP18      **文献标识码:** A

### Recognition algorithm of aerial targets type integrating rough set and DS theory

CONG Rong<sup>1,2</sup>, WANG Xiurkun<sup>1</sup>, YANG Nanhai<sup>1</sup>, MENG Xiang-yu<sup>2</sup>

(1. School of Electronic and Information Engineering, Dalian University of Technology, Dalian 116024, China; 2. Education Technology Center, Dalian Naval Academy, Dalian 116018, China. Correspondent: CONG Rong, E-mail: congrong1002@sina.com)

**Abstract:** Based on analyzing the similarity of rough set and Dempster-shafe (DS) evidence theory, a recognition algorithm of aerial targets type integrating the two theories is proposed. In the algorithm, basic probability assignment function can be deduced from platform database, and the rough belief function is taken as the reliability measure, which is more theoretical than ever. Experiments show that the algorithm has satisfying recognition results.

**Key words:** Rough set; DS evidence theory; Target recognition

### 1 引 言

在实际应用中,所有融合方法都面临处理各种不确定信息的问题,而 DS 方法为不确定信息的表达和合成提供了自然而强有力的方法,使得它在数据融合领域获得了广泛的重视.但由于 DS 方法存在以下 3 个方面的原因,使得它在实际应用中并没有取得预想的效果<sup>[1]</sup>:1) 要求证据必须是独立的,而这一条件有时不易满足;2) 基本概率分配赋值通过经验得到,信任函数在此基础上计算出来,缺少非常坚实的理论支持;3) 计算上存有潜在的组合爆炸问题.妨碍 DS 方法在数据融合领域更加广泛应用的最主要原因是该方法可能导致巨大的计算量问题.

通过研究发现,在实际的融合系统中,采用任意一种单一的数据融合方法均不能很好地解决存在干扰、噪声等不确定因素条件下的目标识别问题,很难得到理想的识别效果<sup>[2]</sup>.因此,将多种融合方法分层使用,功能互补是很好的选择.从当前 DS 方法的发展趋势来看,将相关理论(如模糊集理论、随机集理

论、粗糙集理论等)与之结合是一条重要途径.文献[2]表明,将 DS 方法与模糊集理论和专家系统方法相结合,能提高融合系统中的目标识别的精确性和可靠性.粗糙集理论为 DS 方法的拓展提供了新的手段,两者都是用于处理不确定问题的理论,但它们描述不确定性的角度不同,即 DS 方法使用 mass 函数,而粗糙集理论采用集合划分的概念.在文献[3]中提出了结合粗糙集理论和 DS 方法的数据融合模型,将粗糙集放在融合层次的像素级和特征级,DS 方法放在决策级.

本文的目的是在算法上融合粗糙集理论和 DS 方法,充分利用粗糙集具有强大的理论支持、不需要先验知识、具有强大的数据分析能力等特点,弥补 DS 方法的不足,使得改进后的可信度描述更具理论化,算法具有更高的计算效率和更高的识别准确性.

### 2 基于 DS 方法的空中目标识别模型

基于 DS 方法对多传感器身份数据进行融合的原理框图如图 1 所示.每个传感器把观测数据从观

收稿日期: 2007-06-12; 修回日期: 2007-09-24.

作者简介: 丛蓉(1974—),女,辽宁大连人,副教授,博士生,从事数据挖掘算法的研究;王秀坤(1946—),女,辽宁辽阳人,教授,博士生导师,从事数据挖掘、数据库应用等研究.

测空间变化到证据空间,对每一个命题或传感器所给出的“粗糙”的身份报告分配一个证据,即对每一个命题分配一个基本概率分配赋值,计算相应的信任度和似然度.身份融合系统首先根据 Dempster 的组合规则计算各个命题组合后的概率赋值和相应的信任度区间,然后计算综合概率赋值和信任度区间,最后按照一定的判决规则选择信任度最大的假设作为多传感器信息融合后的身份类型<sup>[4]</sup>.

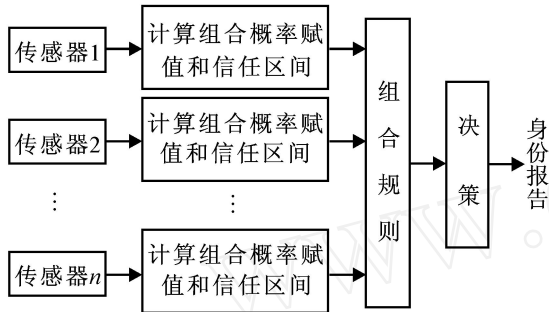


图1 基于DS方法的融合模型

这里需要说明的是,身份融合要求有一个能够提供可能目标先验知识的平台数据库配合系统进行工作.

### 3 相关基本概念

Pawlak 提出的粗糙集理论为证据理论的发展提供了新的机制,它使得无限框架上的证据处理向有限框架上的近似转换成为可能.

#### 3.1 DS方法的不确定性描述

DS方法<sup>[4]</sup>利用信任函数(Bel)和似然函数(Pl)来描述命题的不确定性.根据基本概率分配函数(BPA),信任函数和似然函数分别定义如下:

$$\text{Bel}(A) = \sum_{B \subseteq A} m(B), \quad (1)$$

$$\text{Pl}(A) = 1 - \text{Bel}(\bar{A}) = \sum_{B \subseteq A} m(B). \quad (2)$$

区间(Bel(A), Pl(A))可以描述事件A的不确定性,(1,1)表示A为真,(0,0)表示A为假,(0,1)表示对A一无所知.

#### 3.2 粗糙集理论的不确定性描述<sup>[5,6]</sup>

在粗糙集理论中,知识被认为是一种将现实或抽象的对象进行分类的能力.对决策系统DS=(U, C, D, {V<sub>a</sub>}, f<sub>a}), B ⊆ C, D是属性集合的一个子集,对于每个子集X ⊆ U,它的B下近似集和B上近似集分别定义如下:</sub>

$$B_-(X) = \{x \in U \mid [x]_B \subseteq X\} = \{Y \in U/B \mid Y \subseteq X\}, \quad (3)$$

$$B_+(X) = \{x \in U \mid [x]_B \cap X \neq \emptyset\} = \{Y \in U/B \mid Y \cap X \neq \emptyset\}. \quad (4)$$

集合区间(B<sub>-</sub>(X), B<sub>+</sub>(X))被称为X的B边

界线集,是由那些根据知识B既不能判断肯定属于X又不能判断肯定属于~X=U-X的U中元素组成的集合.X为B粗糙集当且仅当B<sub>-</sub>(X) ⊆ B<sub>+</sub>(X).

如果B ⊆ C是条件属性集合的一个子集,决策属性D的B正域定义为

$$\text{POS}_B(D) = \sum_{x \in U/D} B_-(x). \quad (5)$$

决策属性D的B正域是U中所有根据条件属性B的等价类U/B的信息,可以准确地划分到决策属性D的等价类U/D中的对象集合.

### 4 融合粗糙集和DS方法的空中目标类型识别算法

从定义可知,DS方法的信任函数和似然函数与粗糙集理论的上下近似集描述的都是不确定性,因此本文融合这两种理论,建立统一的框架.

#### 4.1 粗糙集DS方法的相关定义

为描述方便,先给出下面两个表达式:

$$\text{POS}_B(D = j) = \{x \in U \mid x \in \text{POS}_B(D) \wedge D(x) = j\}, \quad (6)$$

$$\text{POS}_{B=i}(D = j) = \{x \in U \mid x \in \text{POS}_B(D = j) \wedge B(x) = i\}. \quad (7)$$

式(6)为U中所有根据条件属性B的等价类U/B的信息,可以准确地划分到决策属性D的等价类{x:D(x)=j}<sub>D</sub>中的对象集合;式(7)为U中所有根据条件属性B的等价类{x:B(x)=i}<sub>B</sub>的信息,可以准确地划分到决策属性D的等价类{x:D(x)=j}<sub>D</sub>中的对象集合.

定义1 粗糙基本概率分配函数(RBPA).对于决策系统DS=(U, C, D, V, f),条件属性C称为证据,决策属性D称为命题,假设B ⊆ C,决策属性集D只包含一个类型属性d, d(x) ∈ {1, ..., N}是决策属性d关于U上的个体x的属性值.在这种意义下,粗糙基本概率分配函数可表示为

$$\begin{cases} m(\emptyset) = 0, \\ m(i) = |\{x \in U \mid d[x]_B = i\}| / |U|; \\ \emptyset \subseteq N = \{i \mid d(x) = i, x \in U\}. \end{cases} \quad (8)$$

其中:|·|表示集合中元素的个数, d[x]<sub>B</sub>表示条件属性集B上的x等价类在决策属性d上的取值集合.

证明

$$\begin{aligned} m(i) &= \sum_{x \in N} |\{x \in U \mid d[x]_B = i\}| / |U| = \end{aligned}$$

$$\frac{|\text{POS}_B(d) \cap U|}{|U|} = 1. \quad (9)$$

定义 2 任意  $B \subseteq N$  的粗糙信任函数 (RBel) 和粗糙似然函数 (RPl) 用粗糙集可以表示为

$$\text{RBel}(B) = \frac{|B \cap \bigcup_{i \in N} X_i|}{|U|}, \quad (10)$$

$$\text{RPl}(B) = \frac{|B^c \cap \bigcup_{i \in N} X_i|}{|U|}. \quad (11)$$

显然,在粗糙集意义下,信任函数  $\text{Bel}(A)$  和似然函数  $\text{Pl}(A)$  的计算较之用 DS 方法计算更加理论化<sup>[7]</sup>.

#### 4.2 空中目标类型识别算法

空中目标类型识别算法是通过多传感器传回的目标特征信息,识别出该组目标的类型.假设输入的目标特征信息已进行了关联处理<sup>[4]</sup>.

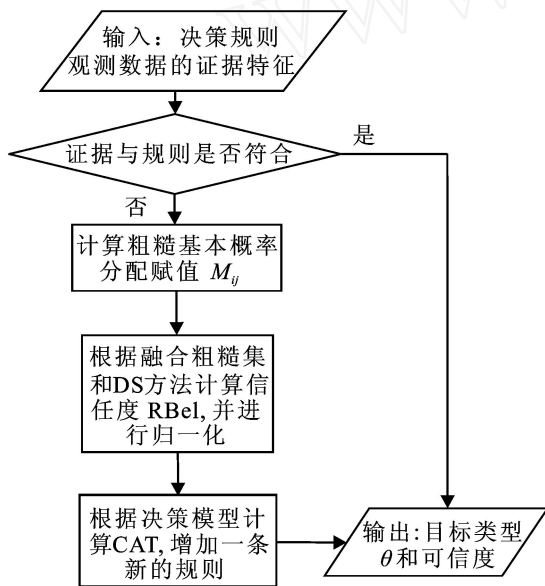


图 2 空中目标类型识别算法流程图

目标识别算法流程图如图 2 所示,详细描述如下:

输入: 1) 目标先验知识的平台数据库 PDB, 相当于一个决策系统  $DS = (U, \{X_1, X_2, \dots, X_M, d\}, V, f)$ , 该决策系统有  $M$  个条件属性表示目标的特征, 1 个决策属性表示目标的类别;

2) 决策规则  $(X_i = a_{x_i}) \Rightarrow (d = ) \{1 \leq i \leq M, 1 \leq \leq N\}$ , 可信度为  $\mu(a, )$ ;

3) 传感器观测数据的证据特征  $(b_1, b_2, \dots, b_M)$ ,  $b_i = \text{null}$  表示某个特征为空值.

输出: 可能的目标类型 CAT 及其可信度  $\mu$ .

Step1: 如果证据特征  $(b_1, b_2, \dots, b_M)$  与某条决策规则符合, 则输出目标类型  $\theta$  和可信度, 算法结束; 否则, 转 Step2.

Step2: 计算根据第  $i$  个证据特征  $(X_i = b_i)$  判断该目标类型  $(d = )$  的粗糙基本概率分配赋值

$$M_{ij} = \frac{|\text{POS}_{X_i=b_i}(d = )|}{|[x: d(x) = ]_d|}. \quad (12)$$

Step3. 1: 按照基本 DS 方法, 根据 Dempster 的组合规则计算各个命题组合后的概率分配函数和相应的信任度区间, 然后计算综合概率分配函数和信任度区间, 计算公式参见文献[4];

Step3. 2: 按照融合粗糙集和 DS 方法, 计算证据组合后的信任度

$$\text{RBel}(B) = \frac{|\text{POS}_{X=(b_1, b_2, \dots, b_M)}(d = )|}{|[x: d(x) = ]_d|}. \quad (13)$$

关于信任函数值和似然函数值的计算, 可采取两种方法: 一种是根据 DS 的组合规则计算, 对应 Step3. 1; 另一种是按照融合粗糙集和 DS 方法计算, 对应 Step3. 2.

Step4: 对粗糙信任函数结果进行归一化, 得到目标属于各类型的概率分布

$$\text{RBel}(i) = \frac{\text{RBel}(i)}{\sum_{i=1}^N \text{RBel}(i)}. \quad (14)$$

Step5: 建立空中目标类型识别的决策模型

$$\text{CAT} = \begin{cases} \text{RBel}(i) = \max_{i=1}^N \text{RBel}(i), & N; \\ \text{RBel}(i) > \theta_1; \\ \text{RBel}(i) - \text{RBel}(i) > \theta_2; \end{cases} \quad N.$$

其中  $\theta$  为目标识别的分辨阈值. 空中目标类型识别的决策模型的判定算法应保证: 判定的目标类型应具有最大的信任函数值, 并且该信任函数值应大于  $\theta_1$ ; 判定的目标类型与其他已知类型的信任函数值的差必须大于  $\theta_2$ .

Step6: 如果  $\text{CAT} \neq N$ , 在 PDB 中增加一条信任度为  $\text{RBel}(i)$  的决策规则:

$$(X_i = b_i) \Rightarrow (d = \text{CAT}), \quad 1 \leq i \leq M, b_i \neq \text{null}.$$

融合后的算法具有以下特点:

1) 融合算法中, 基本概率分配赋值是通过先验知识平台数据库 PDB 推导得出; 在 DS 方法中, 基本概率分配赋值是根据经验得到.

2) 在粗糙集理论意义下, 信任函数和似然函数的计算较之用 DS 方法计算更加理论化<sup>[7]</sup>.

#### 4.3 算法时间复杂度分析

融合算法的时间复杂度是  $O(|U| \cdot \log_2(|U|))$ . 假设数据库 PDB 根据类型和特征建立索引, 特征个数为  $M$ , 类型个数为  $N$ , 数据库的记录数为  $|U|$ , 一般情况下  $M \cdot N \ll |U|$ . 经计算,  $|[x:$

$d(x) = |J_d|$  的时间复杂度是  $O(\log_2 |U|)$ ,  
 $|POS_{X_j=b_j}(Y)|$  的时间复杂度是  $O(|U|)$ ,  
 $|POS_{X=(b_1, b_2, \dots, b_M)}(Y)|$  的时间复杂度是  $O(|U| \cdot \log_2 |U|)^{[6]}$ .

### 5 应用实例与比较

#### 5.1 算法应用实例

建立目标类型识别框架  $\Omega = \{1, 2, \dots, 7\}$ , 其中: 1 目标为  $A_1$  类目标, 2 目标为  $A_2$  类目标, 3 目标为  $B_1$  类目标, 4 目标为  $B_2$  类目标, 5 目标为  $C$  类目标, 6 目标为  $D$  类目标, 7 目标为  $E$  类目标<sup>[8]</sup>.

根据 DS 数据融合的全局方法, 首先需要提取相应的目标特征参数作为空中目标类型的判别证据. 不同类型的空中目标所反应出的战术特征和物理特征众多, 根据对空中目标的分类和编队所能掌握的目标信息, 选取目标的距离特征、速度与高度特征、加速度特征和回波特征作为空中目标类型区分的证据. 为考察本文算法的有效性, 并与 DS 方法进行比较, 采用文献[9]的实验数据, 采用3个批次, 每批次20个目标, 各算法识别的准确率和总体识别时间如表1所示.

表1 空中目标类型识别仿真结果

目标批次	DS方法		算法1		算法2	
	准确率 %	识别时间 / s	准确率 %	识别时间 / s	准确率 %	识别时间 / s
1	100	0.41	100	0.32	100	0.25
2	95	0.72	95	0.6	90	0.49
3	95	0.78	95	0.64	95	0.52

表1中: 算法1对应 Step3.1, 算法2对应 Step3.2. 从实验结果可以看出, 本文所建立的目标类型识别模型具有与 DS 方法同样的识别准确度, 以及更快的识别时间. 在战场环境中, 准确性和实时性是很重要的.

#### 5.2 算法性能比较

实验数据采用6个UCI<sup>[10]</sup>分类问题数据集: diabetes, heartdisease, iris, monk1, soybean 和 satellite 对算法的有效性进行测试, 6个集合的属性如表2所示. 其中 heartdisease 和 soybean 集合因含

表2 数据集属性

数据集	记录数	条件属性个数	决策属性个数	类别数目	属性中是否有空值
diabetes	768	8	1	2	无
heartdisease	270	13	1	2	有
iris	150	4	1	3	无
monk1	432	6	1	2	无
soybean	307	35	1	19	有
satellite	4 435	36	1	7	无

有噪声, 造成数据缺失、信息不完整, 部分属性值为空.

实验软件采用华沙大学开发的粗糙集系统 RSES(Rough Set Exploration System) 和 Delphi 平台开发的融合算法. 实验过程将每个数据集分成训练集(80%)和测试集(20%), 先通过对训练集训练得到决策规则, 再采用本文的两种算法、KNN 和 LTF\_C 神经网络对测试集进行分类. 采用各分类算法的分类精确度结果如图3所示, 其中: 算法1对应 Step3.1, 算法2对应 Step3.2. 结果表明, 算法对于含有噪声类型的数据也具有较好的精确度.

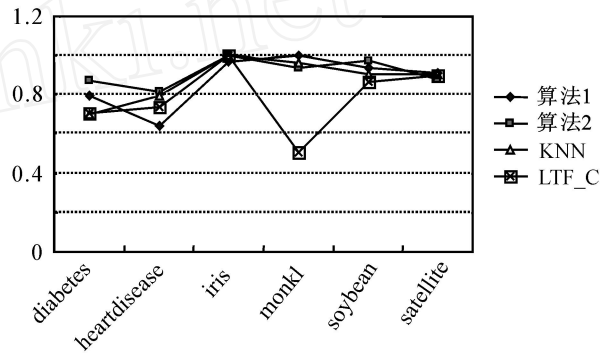


图3 分类算法精确度比较

## 6 结论

对目标类型的识别问题一直是工业和军事部门研究的热点问题. DS 方法和粗糙集理论是两种有效解决目标识别中不确定性问题的方法, DS 方法中采用信任函数和似然函数描述不确定性, 粗糙集理论采用上下近似集描述不确定性. 本文提出一种融合粗糙集和 DS 方法的空中目标类型识别算法, 能够通过先验知识平台数据库 PDB 推导得出基本概率分配赋值, 克服了人为经验的主观性. 仿真实验表明, 该融合算法与 DS 方法具有同样的识别准确度和更快的识别时间. 在与其他分类算法的比较中, 本文提出的融合算法具有较高的分类精确度.

### 参考文献(References)

[1] 徐从富, 耿卫东, 潘云鹤. 面向数据融合的 DS 方法综述[J]. 电子学报, 2001, 29(3): 393-396.  
 (Xu C F, Geng W D, Pan Y H. Review of Dempster-Shafer method for data fusion [J]. Chinese J of Electronics, 2001, 29(3): 393-396.)  
 [2] 徐从富. 基于多 Agent 信息融合技术研究[D]. 杭州: 浙江大学, 2000.  
 (XU C F. Study on information fusion technology based on multi-agents [D]. Hangzhou: Zhejiang University, 2000.)

(下转第 928 页)

## 5 结 论

本文将遗传规划用于作业车间调度问题,并对其作了改进,使之适合解决作业车间调度问题.层次结构表达和染色体长度可变是遗传规划的特点.利用 Read 线性编码和基于工序的编码相结合进行有效的编码,使复杂的树型结构变成对一串数组的遗传操作,提高了运算效率,而且解码过程产生的主动调度减小了搜索空间.同时,对变长染色体在交叉运算时进行了改进,以保证新产生的子代是可行解,防止了无效解的产生.利用典型实例进行了实验,与传统的简单遗传算法进行比较的结果表明,遗传规划可有效求解作业车间调度问题,是一种解决该问题的新途径.当然,仍需进一步改进遗传操作方法,以提高解的质量.

### 参考文献(References)

- [1] Koza J R. Genetic programming: On the programming of computers by natural selection[M]. Cambridge: MIT Press, 1992.
- [2] Koza J R. Genetic programming : Automatic discovery of reusable programs [M]. Cambridge: MIT Press, 1994.
- [3] Koza J R, Bennet F, Andre D. Genetic programming : Darwinian invention and problem solving[M]. San Francisco: Morgan Kaufmann Publishers Inc, 1999.
- [4] Yin W J, Liu M, Wu C. Learning single-machine scheduling heuristics subject to machine [J]. Evolutionary Computation, 2003, 2: 1050-1055.
- [5] 轩建平,史铁林,廖广兰,等.利用遗传编程提取齿轮多重故障分类特征[J].振动工程学报,2006,19(1): 70-74.  
(Xuan J P, Shi T L, Liao G L, et al. Classification feature extraction of multiple gear faults using genetic programming[J]. J of Vibration Engineering, 2006, 19(1): 70-74.)
- [6] 何霆,刘飞,马玉林,等.车间生产调度问题研究[J].机械工程学报,2000,36(5): 97-102.  
(He T, Liu F, Ma Y L, et al. Study on shop-floor scheduling [J]. Chinese J of Mechanical Engineering, 2000, 36(5): 97-102.)
- [7] Alaa F S, Ahmed M. Forecasting using genetic programming [C]. Proc of the 33rd Southeastern Symposium on System Theory. Athens, 2001: 343-347.
- [8] Goncalves, Jose Fernando, De Magalhaes Mendes, et al. A hybrid genetic algorithm for the job shop scheduling problem[J]. European J of Operational Research, 2005, 167(1): 77-95.
- [3] 胡春海,王晓丽,邹晓红. DS 证据理论和粗集理论在数据融合中的应用[J]. 现代雷达, 2004, 26(9): 53-55.  
(Hu C H, Wang X L, Zou X H. Application of DS evidential theory and rough set theory in data fusion[J]. Modern Radar, 2004, 26(9): 53-55.)
- [4] 杨万海. 多传感器数据融合及其应用[M]. 西安: 西安电子科技大学出版社, 2004.  
(Yang W H. Multi-sensor data fusion and its applications[M]. Xi'an: Publishing House of University of Electronic Science and Technology, 2004.)
- [5] Pawlak Z. Rough sets theory and its applications to data analysis[J]. Cybernetics and Systems, 1998, 29(7): 661-688.
- [6] 张文修,吴伟志,梁吉业. 粗糙集理论与方法[M]. 北京: 科学出版社, 2001.  
(Zhang W X, Wu W Z, Liang J Y. The theory and method of rough set[M]. Beijing: Publishing House of Science, 2001.)
- [7] 刘清. Rough 集及 Rough 推理[M]. 北京: 科学出版社, 2001.  
(Liu Q. Rough set and rough reasoning[M]. Beijing: Publishing House of Science, 2001.)
- [8] 宋元. 目标战术意图推理理论及其应用研究[D]. 大连: 海军大连舰艇学院, 2005.  
(Song Y. Research on target's tactical intention inference theory and its applications[D]. Dalian: Dalian Naval Academy, 2005.)
- [9] 李进军,丛蓉,熊吉光. 融合目标战术特征的空中目标类型识别模型[J]. 舰船科学技术, 2005, 27(3): 53-55.  
(Li J J, Cong R, Xiong J G. Differentiating model of aerial targets type integrated with targets tactical characteristic[J]. Ship Science and Technology, 2005, 27(3): 53-55.)
- [10] Newman D J, Hettich S, Blake C L. UCI repository of machine learning databases [BD/OL]. (1998-03-20). <http://www.ics.uci.edu/~mllearn/MLRepository.html>.

(上接第 918 页)