

引用格式:魏煜宁,张劲东,李勇,等.雷达干扰信号识别决策树的自动化设计方法[J].电光与控制,2020,27(4):82-86. WEI Y N, ZHANG J D, LI Y, et al. A method for automatic design of decision tree in radar jamming signal recognition[J]. Electronics Optics & Control, 2020, 27(4):82-86.

## 雷达干扰信号识别决策树的自动化设计方法

魏煜宁, 张劲东, 李勇, 苟立婷  
(南京航空航天大学, 南京 211100)

**摘要:**针对雷达干扰识别决策树分类器设计需要人工介入的问题,提出了一种基于模糊聚类、Xie-Beni 指标和信息增益的决策树自动化设计方法。该方法首先对干扰信号在时域、频域和脉压后时域等维度建立参数特征集,接着在决策树建立过程中引入模糊 C 均值聚类(FCM),从而解决传统决策树需要先验知识设置判决门限的问题;然后通过 Xie-Beni 指标动态确定决策树节点分支数,优化决策树复杂程度;最后使用基于信息增益的 ID3 算法建立模糊聚类决策树。该方法解决了干扰识别决策树的自动化设计问题,且优化了决策树性能。计算机仿真及某雷达对抗实验数据验证了所提方法的有效性。

**关键词:**模糊聚类; Xie-Beni 指标; 决策树; 自动化设计

**中图分类号:** TN974 **文献标志码:** A **doi:**10.3969/j.issn.1671-637X.2020.04.016

## A Method for Automatic Design of Decision Tree in Radar Jamming Signal Recognition

WEI Yu'ning, ZHANG Jindong, LI Yong, GOU Liting  
(Nanjing University of Aeronautics and Astronautics, Nanjing 211100, China)

**Abstract:** To solve the problem that manual intervention is required in the design of the classifier of the decision tree in radar jamming signal recognition, a method for automatic design of the decision tree based on fuzzy clustering, Xie-Beni index and information gain is proposed. Firstly, the feature set of the parameters of the jamming signal in such dimensions as the time domain, the frequency domain and the time domain after pulse compression is established. Then, Fuzzy C-Means (FCM) clustering is introduced in the process of decision tree establishment, so as to solve the problem that the traditional decision tree needs prior knowledge to set the decision threshold. Then, Xie-Beni index is used to dynamically determine the number of branches of decision tree nodes and lower the complexity of the decision tree. Finally, the ID3 algorithm based on information gain is used to establish the fuzzy-clustering decision tree. This method realizes the automatic design of the decision tree in jamming signal recognition and optimizes the performance of the decision tree. Computer simulation and the experimental data of radar countermeasure have verified the effectiveness of this method.

**Key words:** fuzzy clustering; Xie-Beni index; decision tree; automatic design

### 0 引言

随着雷达干扰技术和干扰设备不断更新发展,对雷达的抗干扰能力提出了越来越高的要求。有源干扰技术,特别是基于数字射频存储器(DRFM)的欺骗式干扰技术的发展,使雷达的抗干扰技术面临着极大的挑战。能否正确识别干扰信号是雷达选择抗干扰措施的前提条件。对雷达干扰信号类型快速准确的判别,

是提高雷达在日益复杂的电磁环境下的工作能力和生存能力的关键。

雷达干扰识别的方式主要分为两种,即基于极大似然准则的干扰识别算法和基于特征提取的干扰识别算法<sup>[1]</sup>,其中基于特征提取的干扰识别算法应用最为广泛。国内外学者对基于特征提取的干扰识别算法进行了很多研究:文献[2]使用了广义似然比检验和自适应相干估计进行分类,对于分类阈值较小的情况效果较差;文献[3-5]通过对信号时频分析,将分析结果转换到图像处理领域进行识别,这种方法实现较为复杂,无法满足雷达实时性的要求;文献[6]对有源干

收稿日期:2019-05-10

修回日期:2019-05-28

作者简介:魏煜宁(1995—),男,江苏睢宁人,硕士生,研究方向为雷达信号处理。

扰提取3个熵的参数信息,通过朴素贝叶斯分类器确定识别类型,其对所选参数要求敏感,也无法区分多类型混合干扰;文献[7-8]在时域、频域、双谱域等提取干扰信号特征参数,并通过GA-BP神经网络分类器进行识别,但是需要较长时间进行训练,且灵活性较差;文献[9]使用决策树(Decision Tree)和SVM向量机进行识别,其决策树的门限值需要经验判定或提前分析,而SVM向量机的核函数和部分常数同样需要测试才能选取最佳值,无法实现整个流程的自动化。

针对现有的雷达干扰识别方法存在的主要问题,本文提出一种基于决策树的雷达干扰信号识别的自动化设计方法。决策树易于理解和实现,可以在短时间内处理大量数据,同时对准备的数据要求较简单,但决策树需要对数据进行预处理,其门限的选择需要经验或分析得出。为了解决这个问题,在决策树生成过程中加入了模糊聚类,实现数据自动预处理分类,同时使用Xie-Beni指标,确定决策树分支数,从而实现了雷达干扰信号识别决策树整体流程的自动化设计,随后通过仿真对其性能进行分析,最后通过某雷达对抗实验数据验证了该方法的有效性。

## 1 干扰信号特征参数集

雷达干扰信号特征参数集是识别的关键。通过对雷达干扰信号多个变换域上特征参数的提取,建立特征参数集,从多个角度描述干扰信号在各个变换域上的差异,从而根据这些差异将其区分开来。考虑到雷达工作时电磁环境的复杂性,所选特征参数不仅要有较大区分性,同时要对噪声不敏感。

在各个变换域上都有一些常用的特征参数<sup>[10-13]</sup>。

1) 在时域上,时域矩峰度系数描述信号在顶峰处的陡峭程度,时域矩偏度系数描述信号分布偏离对称性程度,归一化瞬时幅频最大值包含信号起伏信息,归一化瞬时频率分布方差反映信号多普勒频率的变化,相邻波形走动度量相邻波形重叠部分的抖动变化。

2) 在频域上,频域矩峰度考察信号功率谱分布的陡峭程度,频域矩偏度考察信号功率谱分布偏离对称性程度;载波因子考察信号在频域的分布是否存在突变。

3) 在脉压后的时域上,脉压后峰均比反映了信号匹配的情况,脉压后面积比反映了信号对应脉压后的能量大小。

4) 在图形域上,回波灰度图反映了由于天线方向图调制导致的信号回波幅值变化情况。还有分形域特征参数,如盒维数、信息维数等。

## 2 模糊聚类决策树

决策树算法是一种归类学习算法,通过不同的属

性值将样本划分为不同的类,然后利用递归算法对所有的属性划分,最后得到整个样本分类规则。随着模糊集理论在人工智能领域的应用,人们将模糊集同决策树算法结合,提出了模糊决策树算法。模糊决策树算法使得决策树算法的应用范围从处理精确的数据扩大到能处理不精确的数据<sup>[14]</sup>。与传统的决策树算法相比,模糊决策树算法拥有更强的分类能力及稳健性,使得知识表示的方式更为自然<sup>[15]</sup>。使用模糊决策树算法对雷达干扰信号进行分类时,同样也需要对干扰信号特征参数进行分析,确定其划分的隶属函数,若干干扰信号类型较多,同时其类型较为相似时,分析会变得复杂,无法快速选出有用特征参数;另一个问题是生成决策树时每个属性上的聚类数目,其影响着决策树的规模以及节点学习收敛的时间,过多会增加收敛时间,过少则会增加决策树规模。本文提出模糊C均值聚类(Fuzzy C-Means, FCM)决策树算法,该算法通过模糊聚类的方法自动给出决策树的隶属函数,避免了对干扰信号特征参数的先验分析;同时引入Xie-Beni指标,动态衡量聚类数目的平衡。

### 2.1 模糊C均值聚类(FCM)算法

针对干扰信号特征参数集,选定一个特征参数 $\alpha$ ,所有干扰信号的特征参数 $\alpha$ 的值就组成了特征参数集 $A$ ,在特征参数集 $A$ 中, $N$ 为样本总数,即 $a_j \in A, j=1, 2, \dots, N, a_j$ 为每个干扰信号的特征参数的值,FCM算法采用模糊分组,将特征参数集 $A$ 划分为 $C$ 个模糊组,求出每个模糊组的聚类中心 $\bar{a}_i (i=1, 2, \dots, C)$ ,使其价值函数 $J$ 的值最小,所得结果即为最终划分结果。FCM算法的初始化隶属矩阵约束条件为

$$\sum_{i=1}^C u_{ij} = 1 \quad \forall j = 1, 2, \dots, N \quad (1)$$

式中, $0 \leq u_{ij} \leq 1$ ,隶属矩阵 $U$ 描述了干扰信号同所划分类的隶属度。那么可得FCM算法的价值函数为

$$J(U, \bar{a}_1, \bar{a}_2, \dots, \bar{a}_C) = \sum_{i=1}^C J_i = \sum_{i=1}^C \sum_{j=1}^N u_{ij}^m d_{ij}^2 \quad (2)$$

式中: $u_{ij}$ 是模糊隶属度且 $0 \leq u_{ij} \leq 1$ ;  $m$ 表示权重系数且 $m \geq 2$ ;  $d_{ij} = \|\bar{a}_i - a_j\|$ ,为第 $i$ 个聚类中心到第 $j$ 个数据点间的欧氏距离,雷达干扰识别中使用的是单一属性的聚类,这时 $d_{ij} = |\bar{a}_i - a_j|$ 。

构造如下新的目标函数,可求得使式(2)达到最小值的必要条件为

$$\begin{aligned} \bar{J}(U, \bar{a}_1, \dots, \bar{a}_C, \lambda_1, \dots, \lambda_N) = \\ J(U, \bar{a}_1, \dots, \bar{a}_C) + \sum_{j=1}^N \lambda_j \left( \sum_{i=1}^C u_{ij} - 1 \right) = \\ \sum_{i=1}^C \sum_{j=1}^N u_{ij}^m d_{ij}^2 + \sum_{j=1}^N \lambda_j \left( \sum_{i=1}^C u_{ij} - 1 \right) \end{aligned} \quad (3)$$

式中,  $\lambda_j (j=1, 2, \dots, N)$  为拉格朗日乘子。对上式所有的输入参量求导, 可得使价值函数  $J$  最小的两个必要条件分别为

$$\bar{a}_i = \frac{\sum_{j=1}^N u_{ij}^m a_j}{\sum_{j=1}^N u_{ij}^m} \quad (4)$$

$$u_{ij} = \frac{1}{\sum_{k=1}^c \left( \frac{d_{ij}}{d_{ik}} \right)^{2/(m-1)}} \quad (5)$$

式(4)与式(5)就是通过迭代求出了聚类中心  $\bar{a}_i$  和模糊隶属度  $u_{ij}$  的值。最终隶属矩阵  $U$  是一个大小为  $C \times N$  的矩阵, 每一列代表了一种干扰信号, 其对应的  $C$  个隶属度的值表示该干扰信号对每个聚类结果的隶属程度, 隶属程度最高代表该干扰信号隶属于其代表的类。

由上述两个必要条件可以看出该算法是一个迭代算法, 其步骤如下:

- 1) 使用随机数初始化隶属矩阵  $U$ , 根据式(1)的约束条件对其归一化;
- 2) 根据式(4)计算各个聚类中心  $\bar{a}_i, i=1, 2, \dots, C$ ;
- 3) 根据式(2)计算价值函数的值, 同上一次求出的价值函数的值做比较, 若其改变量小于某个阈值  $\varepsilon$ , 则算法停止, 否则继续执行步骤4);
- 4) 根据式(5)计算出新的隶属矩阵  $U$ , 返回步骤2)。

图1是FCM算法的流程图。

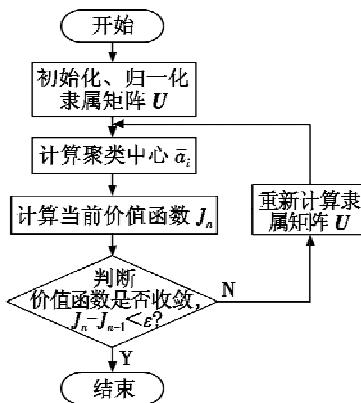


图1 FCM算法流程图

Fig.1 Flow chart of FCM algorithm

### 2.2 Xie-Beni 指标

采用模糊聚类虽然解决了构建传统决策树需要先验知识的问题, 但同时也带来新的问题, 即模糊聚类需要具体的聚类数目  $C$ 。为了选取合适的聚类数目, 引入 Xie-Beni 指标, 其定义为

$$V_{XB} = \frac{\sum_{i=1}^c \sum_{j=1}^N u_{ij}^m \| \bar{a}_i - a_j \|^2}{N(\min_{j \neq k} \| \bar{a}_j - \bar{a}_k \|^2)} \quad (6)$$

式中:  $\frac{1}{N} \sum_{i=1}^c \sum_{j=1}^N u_{ij}^m \| \bar{a}_i - a_j \|^2$  用于评价聚类的紧致性;

$\min_{j \neq k} \| \bar{a}_i - \bar{a}_k \|^2$  用于评价类间的分离性。好的聚类划分应使得类间的间隔尽可能大, 而类内的分散性尽可能小, Xie-Beni 指标的目的在于寻找类内紧凑度和类间分离度之间的某个平衡点, 当其取极小值的时候对应最佳划分<sup>[16]</sup>。表1所示是根据 Xie-Beni 指标  $V_{XB}$  所求最佳划分同实际最佳划分比较结果。其中, 每个数字代表对应的特征参数的分类数目。在决策树生成过程中其分类数目会随着分支不同出现不同的变化。

表1  $V_{XB}$  结果与实际最佳结果比较表

Table 1 Comparison of  $V_{XB}$  result with the actual best result

特征参数	时域矩峰度	时域矩偏度	时域峰均比	频域包络起伏度	频域矩偏度	脉压后时域峰均比
$V_{XB}$ 结果	4	3	3	2	2	3
实际最佳划分	4	3	3	2	2	3

### 2.3 模糊聚类决策树算法的构建

使用 FCM 算法完成了干扰信号数据的聚类后, 通过划分后的信息, 计算各个属性的信息增益值, 作为生成决策树的指标。设在所有干扰信号特征参数集  $D$  中,  $|D|$  为所有干扰信号特征参数集的样本总数,  $K$  为干扰信号类别总数, 类别  $k$  的样本集合用  $C_k$  表示, 则有  $\sum_k |C_k| = |D|$ ,  $D$  的信息熵  $H(D)$  为

$$H(D) = - \sum_k \frac{|C_k|}{|D|} \lg \frac{|C_k|}{|D|} \quad (7)$$

假设特征参数  $A$  有  $N$  个不同的取值, 分别为  $a_1, a_2, a_3, \dots, a_N$ , 特征参数  $A$  通过 FCM 算法将  $D$  聚类划分为  $C$  个子集  $D_1, D_2, D_3, \dots, D_C$ ,  $|D_i|$  为  $D_i$  的样本个数, 且有  $\sum_i |D_i| = |D|$ ,  $C_{ik}$  为子集  $D_i$  中属于类别  $C_k$  的样本集合,  $|C_{ik}|$  表示  $C_{ik}$  集合样本的个数, 则特征  $A$  对数据集  $D$  的条件熵  $H(D|A)$  为

$$H(D|A) = - \sum_i \frac{|D_i|}{|D|} \sum_k \frac{|C_{ik}|}{|D_i|} \lg \frac{|C_{ik}|}{|D_i|} \quad (8)$$

则特征  $A$  的信息增益为

$$G(D, A) = H(D) - H(D|A) \quad (9)$$

决策树算法在每个节点选择当前节点中信息增益最大的特征参数作为该节点的分裂属性, 并以此分裂生长下去, 满足下列条件时, 决策树停止继续生长:

- 1) 每个分支中干扰信号类别已经使用完, 即所有的特征参数在该分支中都被使用一遍, 没有继续生长的分裂属性可供选择;
- 2) 当前节点中所包含的样本都属于同一类别, 即分类已经完成, 没有继续分裂生长的必要。

图2为本文提出的雷达干扰信号识别决策树算法流程图。

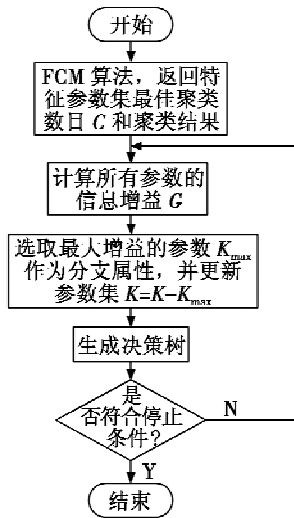


图2 雷达干扰信号识别决策树算法流程图  
Fig.2 Flow chart of the decision tree algorithm for radar jamming signal recognition

### 3 实验数据验证

为验证所提方法的有效性,本文基于 Matlab 仿真了 9 种干扰信号(脉冲宽度  $T_p = 20 \mu s$ ,发射信号带宽

$B = 20 \text{ MHz}$ ),并计算了 6 种特征参数,为了验证这 6 种参数是否能够将 9 种干扰信号区分开来,根据 FCM 算法的同类距离最近的原则将各参数处理后的结果示于表 2,每个参数下的数字代表干扰信号在当前参数下所属的分类标签。

表 2 仿真数据 FCM 处理结果  
Table 2 Result of FCM processing

干扰编号	干扰类型	时域矩峰度	时域矩偏度	时域峰均比	频域包络起伏度	频域矩偏度	脉压后时域峰均比
1	噪声调幅	3	3	2	2	1	1
2	噪声调频	3	3	3	1	4	3
3	梳状谱	2	2	2	1	4	1
4	相参压制	4	3	1	1	4	2
5	前沿复制	1	3	1	1	4	3
6	中间复制	1	3	3	1	4	3
7	后沿复制	4	3	1	1	3	3
8	密集假目标	4	1	2	1	2	2
9	间歇采样	1	2	1	1	4	3

从表 2 可以看出,FCM 算法完成了传统决策树算法中需要提前对数据设置门限然后进行标签分类等过程;将表 2 中的数据以信息增益为标准选择节点进行分支,生成的决策树如图 3 所示。

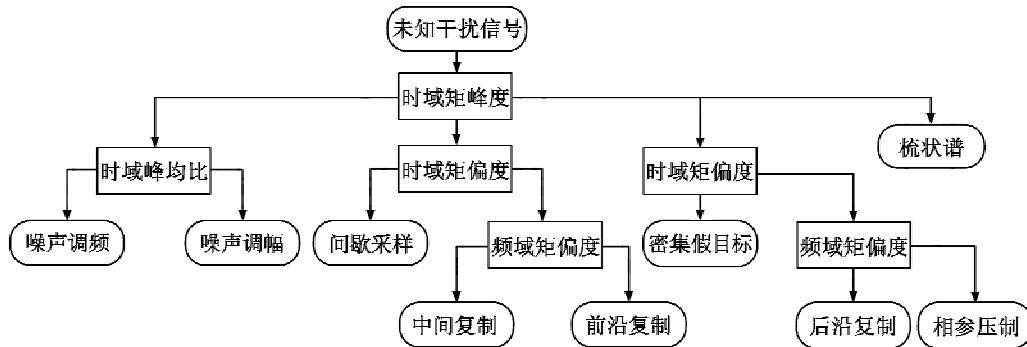


图 3 基于本文算法生成的干扰信号识别决策树

Fig.3 The decision tree of jamming signal recognition based on the algorithm in this paper

从图 3 中可以看出,所仿 9 种干扰信号只用了 4 种特征参数就被完全区分开来,整个过程没有像传统决策树一样人工挑选特征参数以及确定决策树分类门限,实现了整个流程的自动化设计。为了验证该算法生成决策树分类器的识别率,采用上述仿真数据,同传统 ID3 决策树算法进行比较,结果如图 4 和表 3 所示,图 4 是两种算法下干扰信号的整体识别率的比较,表 3 是在  $R_{J,N} = 12 \text{ dB}$  的情况下各个干扰信号的识别率。

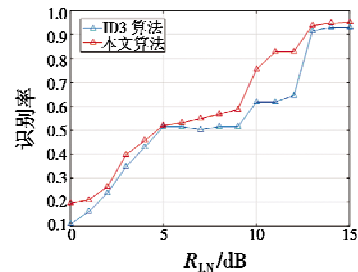


图 4 两种算法在不同干燥比下的识别率

Fig.4 The recognition rates of the two algorithms under different  $R_{J,N}$

表 3 两种决策树算法下几种干扰信号的识别率对比 ( $R_{J,N} = 12 \text{ dB}$ )

Table 3 Comparison of recognition rates of several jamming signals under two decision tree algorithms ( $R_{J,N} = 12 \text{ dB}$ )

	噪声调幅	噪声调频	梳状谱	相参压制	前沿复制	中间复制	后沿复制	密集假目标	间歇采样
ID3 算法	0.8974	0.9275	0.6248	0.9059	0.9159	0.8984	0.5915	0.4165	0.9175
本文算法	0.9794	0.9813	0.8243	0.9836	0.9366	0.9265	0.7042	0.9589	0.9817

从图 4 和表 3 中可以看出,本文算法的识别率在同样条件下比采用传统 ID3 决策树算法的识别率有明显的提高,这是因为传统 ID3 算法需要人为设置判决门限,而这个门限在干噪比较低时受噪声影响较大,不能保证干扰识别的识别率,特别是表 3 中梳状谱、后沿复制和密集假目标等干扰信号的 ID3 算法识别率较低。本文算法在解决建立决策树需要人工筛选特征参数的问题的同时,无需大量数据学习和设置判决门限,降低了噪声对于干扰识别的影响,在较低的干噪比下也能取得较高的识别率。

最后,将本文算法应用于某雷达对抗实验,取得了较为理想的效果,平均识别率可达到 96% 以上。其各干扰信号识别率如表 4 所示。

表 4 某雷达对抗实验干扰信号识别率  
Table 4 Radar recognition rate of jamming signals in experiment

		输出				
干扰信号类型		前沿复制	密集假目标	密集卷积	相参压制	引导噪声
输 入	前沿复制	0.9776	0	0.0037	0.0149	0.0037
	密集假目标	0	0.9554	0.0037	0.0409	0
	密集卷积	0	0.034	0.9170	0.0491	0
	相参压制	0	0	0.0036	0.9927	0.0036
	引导噪声	0	0	0	0.0038	0.9962

#### 4 总结

本文针对原有雷达干扰信号识别决策树的设计过程需要人工介入解决门限及特征需要筛选的缺点,提出了一种雷达干扰信号识别决策树的自动化设计方法,该方法通过模糊 C 均值聚类解决传统决策树的判决门限需要人工先验知识介入的问题;同时引入 Xie-Beni 指标,决定决策树最佳分支数,降低决策树复杂性;最后使用信息增益作为选择分裂条件的标准。该方法训练所需样本少、速度快。仿真表明该方法比传统的决策树方法识别率高,设计过程较简单,最后某雷达对抗实验数据验证了本方法的可行性。

#### 参 考 文 献

[1] 朱清祎,尚朝轩,董健. 雷达干扰模式识别技术综述

[J]. 飞航导弹,2018,1(8):71-75.

- [2] GRECO M S, GINI F, FARINA A. Radar detection and classification of jamming signals belonging to a cone class [J]. IEEE Transactions on Signal Processing, 2008, 56(5):1984-1993.
- [3] 杨兴宇,阮怀林. 基于时频图像 Zernike 矩特征的欺骗干扰识别[J]. 现代雷达,2018,40(2):91-95.
- [4] 刘建洋. 基于时频域分析的电子干扰识别方法研究[D]. 成都:电子科技大学,2018.
- [5] 杨兴宇,阮怀林. 基于 Zernike 矩和双谱特征的新型干扰识别算法[J]. 空军工程大学学报:自然科学版,2018,19(2):56-61.
- [6] 李紫航,宋万杰. 有源干扰感知的熵理论方法[J]. 信号处理,2017,33(12):1652-1656.
- [7] 史忠亚,吴华,沈文迪,等. 基于双域特征的雷达欺骗干扰样式识别方法[J]. 火力与指挥控制,2018,43(1):136-140.
- [8] 李娜. 雷达有源干扰分类与识别方法研究[D]. 西安:西安电子科技大学,2017.
- [9] 梁金弟,程郁凡,杜越,等. 联合多维特征的干扰识别技术研究[J]. 信号处理,2017,33(12):1609-1615.
- [10] 檀鹏超. 雷达有源欺骗干扰多维特征提取与识别技术研究[D]. 成都:电子科技大学,2016.
- [11] 常成. 雷达欺骗式干扰检测与实现[D]. 上海:上海交通大学,2009.
- [12] 郭红敬. 无人机测控链路干扰感知与识别技术研究[D]. 成都:电子科技大学,2017.
- [13] 范伟. 雷达有源干扰信号特征分析与识别算法研究[D]. 成都:电子科技大学,2007.
- [14] NAYAK N R, DASH P K, BISOI R. A hybrid time frequency response and fuzzy decision tree for non-stationary signal analysis and pattern recognition [J]. International Journal of Automation and Computing, 2019, 16(3):398-412.
- [15] 张琳,陈燕,李桃迎,等. 决策树分类算法研究[J]. 计算机工程,2011,37(13):66-67,70.
- [16] 赵娜娜,钱雪忠,冯振华. 新的模糊聚类有效性指标[J]. 微电子学与计算机,2016,33(8):121-125.