异构属性数据的量子聚类方法研究

李志华，王士同
LI Zhi-hua，WANG Shi-tong

江南大学 信息工程学院，江苏 无锡 214122
School of Information and Engineering，Jiangnan University，Wuxi，Jiangsu 214122，China
E-mail：ezhli@yahoo.com.cn

LI Zhi-hua，WANG Shi-tong.Weighted mahalanobis distance-based quantum clustering approach for heterogeneous data. Computer Engineering and Applications，2009，45（23）：63-66.

Abstract： The dissimilarity measure and clustering approach about the heterogeneous dataset are studied，and a Weighted Mahalanobis Distance-based Quantum Clustering（WMDQC）algorithm is presented in this paper.Data often do appear in homogeneous groups，the WMDQC utilizes the structural information to improve the clustering accuracy.Unlike the numeric data，categorical data are often unbalancedly distributed，whose distribution are often unrelated with their distance measure.These characteristics are very similar to the particle world in quantum mechanism，so the WMDQC ascertains the clustering centers by the rewriting quantum potential.Further，a WMDQC-based method WMDQCM is proposed，the WMDQCM mines the structural clue by the agglomerative hierarchical clustering AHC algorithm to construct the weight matrix.By presenting the above to the WMDQC，the final clustering results are obtained.The new WMDQCM exhibits its robustness to initialization and clustering capability to heterogeneous dataset.Experimental results compared with other methods demonstrate that the proposed method has promising performance.

Key words： heterogeneous data；dissimilarity measure；Mahalanobis distance；quantum potential；clustering algorithm

摘 要：研究了异构属性数据的聚类问题。通过挖掘样本中的结构信息，用加权的 Mahalanobis 距离来度量异构样本的相异性；根据分类属性数据的分布与粒子在量子势能场中的分布不平衡的相似性，重写量子势能公式为距离量子势能的形式，提出了一种新的异构属性数据量子聚类 WMDQC 算法。通过进一步集成该算法和 AHC 算法为 WMDQCM 聚类方法，用 AHC 算法更高效地挖掘样本中有利于聚类的结构线索。实验结果表明，方法具有比较优势，显著地改善了聚类性能，具有一定的实用价值。

关键词：异构数据；相异性度量；Mahalanobis 距离；量子势能；聚类算法
DOI：10.3778/j.issn.1002-8331.2009.23.018 **文章编号：**1002-8331（2009）23-0063-04 **文献标识码：**A **中图分类号：**TP181

1 前言

聚类分析的主要任务是：通过挖掘样本中的各种聚类线索，从而有效地选择聚类算法中的合适参数，以达到确定最佳的聚类数目的目的。当前，很多数据样本，如网络检测、网页分类和疾病诊断数据等，这些数据集的组成复杂，存在着大量“构造相似”的同质样本[1]；同时其属性组成呈异构性特点，既有分类属性又有连续属性，而分类属性的数据类型又可能是字符、二值属性和一些具有枚举数据特征的数值等，表现出鲜明的结构特征。通过挖掘样本中蕴含的这些结构信息，能显著提高聚类性能。

研究异构属性样本的聚类，必须首先要能有效地确定聚类中心、并实现对异构属性的相异性的同时度量。由于分类属性数据分布固有的无序性，样本的分布不平衡、分布与空间距离无关[2]，所以对其相异性度量比较困难，造成现今大多数聚类算法不能实现分类属性数据、异构属性数据的聚类，如著名的 FCM 算法就只能实现连续属性数据的聚类。只有少数几个从 k-均值算法演变而来的算法[3-5]能实现分类属性数据的聚类，只有基于混合属性数据（k-prototype）的聚类算法[6]等能实现异构属性数据的聚类。但算法或多或少存在一些不足，如需要处理大量的二进制数据、对初始类中心敏感和需要选取特殊的模式[3-5]等。

虽然量子聚类（Quantum Clustering，QC）算法[7]已经存在，但同样不适宜异构属性数据的聚类，更没有充分挖掘样本中有利于聚类的结构线索。通过挖掘样本中的结构信息给 Mahalanobis 距离加权；用加权的 Mahalanobis 距离来度量异构样本的相异性；并通过重写量子势能公式为距离的形式，用样本的相异性来计算样本的量子势能，提出了基于加权的 Mahalanobis 距离的量子聚类（Weighted Mahalanobis Distance_based

基金项目：2007 年度教育部高等学校创新工程重大培育项目。
作者简介：李志华（1969-），男，副教授，博士研究生，主要研究方向为模式识别、信息与网络安全；王士同（1964-），男，教授，博士，博士生导师，主要研究方向为人工智能、模式识别、模糊系统、生物信息学。
收稿日期：2008-04-28 **修回日期：**2008-06-26

Quantum Clustering，WMDQC）新算法，进一步集成WMDQC算法和凝聚层次聚类（Agglomerative Hierarchical Clustering，AHC）算法，通过AHC算法更高效地挖掘样本中的结构，在此基础上提出了基于WMDQC算法的聚类方法（Mahalanobis Distance_based Quantum Clustering Method，MDQCM），该方法的优势在于解决了结构信息的自动挖掘、自动判断类的数目和对初始化敏感的问题。

2 量子势能简介

波函数是粒子量子态的描述[8]，在量子力学中主要通过薛定锷方程来求解有势场约束的波函数。薛定锷方程有显含时间和不显含时间两种类型，薛定锷方程的物理含义是：在一定的势场中，求解粒子的分布，即波函数，粒子的分布与势场有这样的关系：当粒子的空间分布缩变到一维无限深势阱时，粒子聚集在势能为零的一定宽度的势阱中[8]。因此，量子势能相当于一个抽象的源，随着量子势能趋近于零或比较小时，势阱中粒子的分布密度比较高。基于此，对于粒子分布函数已知的情况，可以通过反向求解薛定锷方程来具体计算量子势能，根据量子势能来确定粒子的分布状态。

已知高斯函数是薛定锷方程的一个解之一，高斯波包如式（1）所示，它代表粒子的分布状态。假定它对应于尺度空间（scale-space）中的一个观测样本集 $\boldsymbol{X}=\{x_1,x_2,\cdots,x_i,\cdots,x_n\}\subset R^d$，$\boldsymbol{x}_i=(x_{i1},x_{i2},\cdots,x_{id})^{\mathrm{T}}\in R^d$ 问题变成使用宽度为 δ（scale parameter）的高斯波包来描述样本点的分布。根据不显含时间的薛定锷方程，此时的量子势能函数的通用表达式如式（2）所示[7-8]：

$$\varphi(x)=\sum_{i=1}^{n}\mathrm{e}^{-\frac{(x-x_i)^2}{2\delta^2}} \tag{1}$$

$$V(x)=E+\frac{(\delta^2/2)\nabla^2\varphi}{\varphi} \tag{2}$$

其中，φ 是波函数，V 是量子势能函数，E 是能量特征值，∇ 是劈形算子，δ 是方程中唯一的一个尺度调节参数，可以看出，不同的 δ 值对应不同的量子势能。

对于一般情况，进一步把式（1）代入式（2），并注意到量子势能函数中的距离特征，根据样本间的距离，重写量子势能函数式（2）如式（3）所示：

$$V(x_i)=E-\frac{d}{2}+\frac{\frac{1}{2\delta^2}\sum_i(\boldsymbol{x}-\boldsymbol{x}_i)^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{x}_i)\exp[-\frac{1}{2\delta^2}(\boldsymbol{x}-\boldsymbol{x}_i)^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{x}_i)]}{\sum_i\exp[-\frac{1}{2\delta^2}(\boldsymbol{x}-\boldsymbol{x}_i)^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{x}_i)]} \tag{3}$$

可见式（3）是一个基于距离的量子势能函数的表达式。其中 d 是薛定锷方程中Hamilton算子的最小的可能特征值，可以用样本的维数表示[7]，当粒子在协振子中振荡时，其能量特征值为 $E=d/2$。对于空间中只有一个粒子 x_1 的情形，即 $n=1$，由式（2）、式（2）求得此时的量子势能函数如式（4）所示。

$$V(x)=\frac{1}{2\delta^2}(\boldsymbol{x}-\boldsymbol{x}_1)^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{x}_1) \tag{4}$$

为了计算的方便，假设 V 是非负且确定的，即 V 的最小值为零，E 可以通过式（2）求得：

$$E=-\min\frac{(\delta^2/2)\nabla^2\varphi}{\varphi} \tag{5}$$

根据量子理论可知，低势能的粒子振动小，相对比较稳定[8]，意味着粒子的分布相对比较集中，分布密度比较高，所以对应于样本分布的研究，可以根据这一粒子的量子分布机制来确定聚类中心。

3 异构属性数据的量子聚类算法

因为Mahalanobis距离考虑了样本的协方差，协方差越大，马氏模越小，包含了更多的结构信息[2]，为了实现异构属性数据相异性的同时度量，以下提供一种基于加权的Mahalanobis距离度量的方法。假设有样本集：$\boldsymbol{x}_i\in\boldsymbol{X},\boldsymbol{x}_i=(x_{i1},x_{i2},\cdots,x_{il},x_{il+1},\cdots,x_{il+m})$，$1\leqslant i\leqslant N$，且前 l 维为分类属性，后 m 维为连续属性，即 $\boldsymbol{X}$ 为异构属性数据集。

3.1 分类属性的相异性度量

对于分类属性数据的聚类算法的研究，最为关键的就是：聚类对象（如样本、聚类中心等）间相异性的度量。采用文献[3]的 σ 函数定义形式来表示分类属性间相异性度量测度 S，如式（6）所示。

$$s(x_i,x_j)=\sum_{p=1}^{l}\sigma(x_{ip},x_{jp})\ (1\leqslant i\leqslant n,1\leqslant j\leqslant n,i\neq j,1\leqslant p\leqslant l) \tag{6}$$

$$\sigma(x_{ip},x_{jp})=\begin{cases}0(x_{ip}=x_{jp})\\1(x_{ip}\neq x_{jp})\end{cases}$$

3.2 异构属性的相异性度量

考虑到样本的属性是异构的，两个异构样本之间的差形式化的表示成：$\boldsymbol{x}_i-\boldsymbol{x}_j=(\sigma(x_{i1},x_{j1}),\sigma(x_{i2},x_{j2}),\cdots,\sigma(x_{il},x_{jl}),x_{il+1}-x_{jl+1},x_{il+2}-x_{jl+2},\cdots,x_{il+m}-x_{jl+m})^{\mathrm{T}}$。算法区分两种情况来分别计算样本之间的距离，当两个样本属于同一类时，其间的Mahalanobis距离用式（7）计算：

$$d_m(\boldsymbol{x}_i,\boldsymbol{x}_j,\boldsymbol{\Sigma})=\sqrt{(\boldsymbol{x}_i-\boldsymbol{x}_j)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}_i-\boldsymbol{x}_j)} \tag{7}$$

当两个样本属于两个不同的类 S_i、S_j 时，其间的距离用加权的Mahalanobis距离如式（8）进行计算[2]：

$$d_{wm}(\boldsymbol{x}_i,\boldsymbol{x}_j)=\sqrt{\frac{|s_i|}{|s_i|+|s_j|}d_M^2(\boldsymbol{x}_i,\boldsymbol{x}_j,\boldsymbol{\Sigma}_i)+\frac{|s_j|}{|s_i|+|s_j|}d_M^2(\boldsymbol{x}_i,\boldsymbol{x}_j,\boldsymbol{\Sigma}_i)} \tag{8}$$

其中 $\boldsymbol{\Sigma}$ 为相应的样本的协方差：$\boldsymbol{\Sigma}=\frac{1}{S}\cdot\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}-\frac{1}{|S|^2}\cdot\boldsymbol{X}^{\mathrm{T}}\bar{1}_{|S|}\bar{1}_{|S|}^{\mathrm{T}}\boldsymbol{X}$，

若 $\boldsymbol{\Sigma}$ 为奇异矩阵，则计算它的伪逆：$\boldsymbol{\Sigma}^{+}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{G}^{-1}\boldsymbol{A}$，其中 $\boldsymbol{G}$ 是一个仅包含 $\boldsymbol{\Sigma}$ 正的特征值的对角矩阵，$\boldsymbol{A}$ 是一个包含 $\boldsymbol{\Sigma}$ 的特征向量的单位正交矩阵；$|S_i|$、$|S_j|$ 可通过训练样本得到，由它们构造Mahalanobis距离的权重矩阵，矩阵的元素即相应的权值。式（8）更好地体现了样本的结构特征，利用这样的聚类线索可显著地改善聚类性能。

另外，为了提高对蕴含于样本中的结构线索的挖掘效率，还可采用Ward's linkage方法的AHC算法来训练样本。因为AHC算法的结果是一棵层次聚类树，可通过截取不同层的子树得到样本的最佳分类数，然后构造Mahalanobis距离的权重矩阵，从而显著改善算法的聚类精度。

3.3 参数 δ 的指导性估计

文献[7]指出，式（2）中 δ 的取值范围为[0，2]，理论上讲有无穷多个取值选择，为了缩短算法的训练时间，采用文献[9]的方法对参数 δ 进行指导性估算。其主要过程如下：通过建立 K-近邻的方差分布的Gamma分布模型，用Gamma分布参数估算的矩方法（moments methods）来估算 δ。首先，估算其 K-近邻方

差。假设δ是一个与局部样本分布的散度有关的参数，并可用样本与其K–近邻之间的欧氏距离来计算。这样对于任意一个抽样样本$\boldsymbol{x}_i$，其K–近邻$x_{(K)}$的方差为：$s_i=(\sum_{k=1}^{K}\|\boldsymbol{x}_{(k)}-\boldsymbol{x}_i\|^2)/(K-1)$，$K<n$，对于$m$个抽样样本，即$i=1,\cdots,m,m\leqslant n$，则这$m$个样本的局部相邻点的方差的平均值为：$\bar{s}=\frac{1}{m}\sum_{i=1}^{m}s_i$；第二，建立$K$–近邻的方差分布的Gamma分布模型。认为样本直方图的方差就相当于抽样样本的局部相邻点方差的分布，用下式估算：$l=(\sum_{i=1}^{m}(s_i-\bar{s})^2)/(m-1)$；第三，用局部相邻点方差的经验分布来估算Gamma分布的两个重要参数。根据矩方法，在仅考虑前面的第1、2个矩时，Gamma分布的两个重要参数：形状参数（shape parameter）α和尺度参数（scale parameter）β，估算成：$\hat{\alpha}=(\bar{s}/l)^2$；$\hat{\beta}=\bar{l}^2/\bar{s}$；最后，用Gamma分布的平均值估算尺度参数$\hat{\delta}$，如式（9）所示：

$$\hat{\delta}=\hat{\alpha}\hat{\beta} \tag{9}$$

这一方法策略，克服了需要反复训练样本来选取参数δ的不足。显著地提高了算法的效率。

3.4　基于加权的Mahalanobis距离的量子聚类算法

在量子力学中，根据粒子在势能场中分布的量子机制，很容易得出如下特征：（1）由于粒子是"跳动"的，所以粒子在能量场中的分布不遵循任何几何形状；（2）势能为零或比较小时，势阱中粒子的分布密度比较大，并且相对比较稳定。所以，对于样本的分布，就可以把势能极小值的样本作为聚类中心，再通过相应的相异性度量方式来度量样本间的差异，进而完成样本的聚类，于是，可以定义这样一个过程为聚类算法，具体描述如下：

算法1　基于加权的Mahalanobis距离的量子聚类（WMDQC）算法

步骤1　初始化δ、$c=0$、md和距离权重矩阵，c是聚类个数、md是度量距离；

步骤2　根据式（3）计算样本的势能V；

步骤3　$c=c+1$；

步骤4　根据样本的势能V，在未被聚类的样本中求最小势能点$V_{\min}=\min_n\{V_n\}$，且令：$V(x_k)=V_{\min}$，$v_c=x_k$，为第c类聚类中心；

步骤5　把满足$d(x_i,v_c)<md$条件的所有样本x_i，假设v_c表示存在的聚类中与x_i最相近的那个聚类，则把x_i聚成第v_c类，并从样本集X中删除x_i；

步骤6　如果X为空，算法结束，否则转步骤3。

以上算法产生c个聚类，聚类中心用$v_i(1\leqslant i\leqslant c)$表示。该算法具有以下明显优点：

（1）算法不依赖于样本的几何形状，充分地挖掘了样本中有利于聚类的结构信息；

（2）聚类中心的确定取决于数据样本的潜在结构；

（3）类中心的选择方法决定该算法对初始类中心的选择不敏感，具有相当的鲁棒性优势；

（4）聚类不受类本身大小的影响，因为样本的分布是不平衡的；

（5）算法具有相当的应用灵活性。

为了更好地构造算法中的距离权重矩阵，缩短算法的训练时间，提高算法的运行效率，改善算法的聚类性能，以下提供一种集成WMDQC算法和AHC算法的聚类方法，见算法2。

算法2　基于WMDQC算法的聚类方法（WMDQCM）

步骤1　用3.3节的方法对参数δ作指导性估算；

步骤2　用AHC算法训练样本，构造距离权重矩阵；

步骤3　用WMDQC算法完成样本聚类，得到聚类结果。

4　仿真实验及分析

本章用两个指标来进行聚类效果分析，其一为错分率，错分率是指某类中被错分的样本与该类中实际样本数的比值；另一个是聚类准确率γ[10]，是指各个聚类中所有被正确聚类的样本数总和与样本集样本总数的比值。表示成公式形式如式（10）所示：

$$\gamma=\frac{\sum_{i=1}^{c}number_i}{n} \tag{10}$$

$number_i$为第i类中被正确聚类的样本个数，n为样本的总数。γ值越大，就说明聚类准确度高，聚类效果好，越小，则刚好相反。

仿真实验采用的实验样本集[12]的组成如表1所示。第1个样本集一共50个样本数据，包含4类2维样本，其中4类样本的数量分别为12、15、13、10；第2个数据样本集作为标准的分类属性样本在多篇文献[3–6]中被采用，一共是35维47个样本，样本集共分为4类，即4种疾病：Phytophthora Rot、Diaporthe Stem Canker、Charcoal Rot、Rhizoctonian Root，4种疾病的样本数量分别为：17、10、10、10，仿真实验仅选用其中的21维分类属；第3个是一个17维的纯字符分类属性样本集，属性的值域{y，n}简单地数值化为{1，2}，样本集共分为2类，分别为democrates、republicans，仿真实验中去除了含有不确定值"？"的样本，共剩余232个样本，其中democrats、republicans的样本数分别为124和108；第4个样本集是通过从标准kddcup99的10%样本子集中随机抽取的，共有998条41维样本，分为9类，其中1类为正常样本，其余8类为异常样本。41维属性中34维为连续属性，7维为分类属性，7维分类属性中又有3维为纯字符，是一个很复杂的异构属性数据集。

对以上样本用WMDQCM方法同其它的算法或其它文献的指标进行比较，结果如表2~表5所示。

表1　实验样本集的组成

样本名称	属性类型	样本个数	维数	类
Synthetic4	numeric	50	2	4
Soybean	categorical	47	21	4
Vote	categorical	232	16	2
kddcup99	heterogeneous	998	41	9

表2　该文算法对三个标准样本的聚类结果（δ=0.707）

算法	md	类1	类2	类3	类4	聚类准确率/（%）
Synthetic4	11.8	0:12	0:15	0:13	0:10	100
Soybean	9.6	0:10	0:10	3:10	0:17	93.62
Vote	11.8	0:108	28:124			87.931

表2是该文方法对3个标准样本集的聚类参数和聚类结果；表4是同文献[4]的比较，WMDQCM算法100次运行的统

表 3 算法 WMDQC 在 kddcup99 样本集上的聚类结果 (md=18.5,δ=0.707)

类	样本数/条	检测率/(%)	大类检测率/(%)
normal	188	86.17	86.17
smurf	578	100	
back	4	50	
neptune	215	100	总检测率:
teardrop	2	100	99.38
ipsweep	3	100	
satan	4	75	
warezclient	4	25	

表 4 该文算法同文献[4]在样本 soybean 上的比较(100 次) (%)

算法	Hard k-modes	Fuzzy k-modes	WMDQCM
聚类准确率	78.2	80.45	93.62

表 5 该文算法同其他算法在样本 synthetic4 上的比较 (%)

算法	k-means	FCM	WMDQCM
聚类准确率	96	96	100

计平均值基本上稳定在 93.62%，这说明算法的聚类稳定性比其它两种算法要好；表 5 说明，当 WMDQCM 退变成纯连续属性样本的聚类算法时，同样具有比较好的聚类性能，这说明通过对样本结构信息的挖掘，能显著改善算法的聚类性能；由于该文算法对 Vote 样本中的不确定值的处理与文献[1]的方法不同，所以没有可比较性，但从表 2 可见取得了比较好的聚类效果；表 3 说明了算法对异构属性数据集的聚类有效性，从入侵检测的角度而言，检测指标比较理想，说明算法具有一定的实用价值。

5 结束语

目前，能实现分类属性数据聚类的算法较少，对于异构属性数据的聚类算法的研究更少，研究、提出具有比较高性能的异构属性数据聚类算法成为一个迫切的的问题。在前人的基础上，提出了一种新的 WMDQC 算法，并通过与著名的 AHC 算法的集成，进一步提出了 WMDQCM 聚类方法，方法更加高效地挖掘了样本中有利于聚类的结构线索，比较显著地提高了算法的聚类性能。文中总结了 WMDQC 算法的优点和 WMDQCM 方法同其它算法的比较优势。但是，算法中同样存在着一些不足，在处理大数据量样本时，容易造成距离矩阵维数过大问题，这也是有待进一步深入研究改进之处。

参考文献：

[1] Kim M,Ramakrishna R S.Projected clustering for categorical datasets[J].Pattern Recognition Letters,2006,27:1405-1417.
[2] Wang De-feng,Yeung D S,Tsang E C C.Weighted mahalanobis distance kernels for support vector machines[J].IEEE Transaction on Neural Networks,2007,18:1453-1462.
[3] Huang Zhe-xue.Extensions to the k-means algorithm for clustering large data sets with categorical values[J].Data Mining and Knowledge Discovery,1998,2:283-304.
[4] Huang Zhe-xue,Ng M K.A fuzzy k-modes algorithm for clustering categorical data[J].IEEE Transactions on Fuzzy System,1999,7(4):446-452.
[5] Ralambondrainy H.A conceptual version of the k-means algorithm[J].Pattern Recognition Lett,1995,16:1147-1157.
[6] Chen Ning,Chen An,Zhou Long-xiang.Fuzzy k-prototypes algorithm for clustering mixed numeric and categorical valued data[J].Journal of Software,2001,12(8):1107-1119.
[7] Horn D,Gottlieb A.Algorithm for data clustering in pattern recognition problems based on quantum mechanics[J].Physical Review Letters,2002,88(1).
[8] Gasiorowicz S.Quantum physics[M].New York:Wiley,1996.
[9] Nasios N,Bors A G.Kernel-based classification using quantum mechanics[J].Pattern Recognition,2007,40:875-889.
[10] Xie X L,Beni G.A validity measure for fuzzy clustering[J].IEEE Transactions on Pattern Analysis and Machine Intelligence,1991,13(8):841-847.
[11] Merz C J,Merphy P.UCI repository of machine learning databases[EB/OL].http://www.ics.uci.edu/mlearn/MLRRepository.html.

(上接 10 页)

$q_5 \vee q_6$，经过计算可知 $\varepsilon_1=\rho(A_1,A_1 \vee B)=3/32$，$\varepsilon_2=\rho(A_2,A_2 \vee B)=3/64$，$\rho(A_1,A_2)=1/16$，此时 $\varepsilon=\varepsilon_1+\varepsilon_2\approx 0.14$，$\eta\approx 0.10$，而 $A_1 \wedge A_2=A_1$，所以 $A_1 \wedge A_2 \in D^1_{3/32}(\Gamma)$，$3/32\approx 0.09$。

参考文献：

[1] Wang Guo-jun,Zhou Hong-jun.Quantatitive logic[J].Information Sciences,2009,179:226-247.
[2] 王国俊.计量逻辑学(I)[J].工程数学学报,2006,23(2):191-215.
[3] 王国俊.数理逻辑引论与归结原理[M].2 版.北京:科学出版社,2006.
[4] Hajek P.Matematics of fuzzy logic[M].[S.l.]:Kluwer Academic Publishers,1998.
[5] 王国俊.数理逻辑引论与归结原理[M].北京:科学出版社,2003.
[6] Hamilton A G.Logic for mathematicians[M].New York:Cambridge University Press,1978.
[7] 王国俊,傅丽,宋建社.二值命题逻辑中命题的真度理论[J].中国科学:A 辑,2001,31(11):998-1008.
[8] 王国俊,李璧镜.Lukasiewicz n 值命题逻辑中公式的真度理论和极限定理[J].中国科学:E 辑,2005,35(6):561-569.
[9] 王国俊.计算智能:第二册[M].北京:高等教育出版社,2005.
[10] Zhou H J,Wang G J.A new consistency index based on deduction theorems in several logic systems[J].Fuzzy Sets and Systems,2006,157(3):427-443.
[11] Wang Guo-jun,Zhou Hong-jun.Quantatitive logic[J].Information Sciences,2009,179:226-247.
[12] 惠小静.三种近似推理模式的等价性[J].计算机工程与应用,2008,44(27):56-57.