

# 复杂细微差异化网络数据特征的语义优化提取算法

杨伟杰

(北京工商大学计算机与信息工程学院 北京 100048)

**摘要** 对网络数据的复杂、细微、差异化特征进行语义提取,是实现 Web 网络数据准确识别和检索的关键技术。复杂、细微、差异化的网络数据语义特征具有非线性和随机散布性的特点,其主题分布广、更新频率大,从而造成语义特征提取困难。传统方法采用小波基函数投影算法进行语义特征的提取,性能不好。提出了一种基于 Dopplerlet 变换匹配投影的网络数据特征语义优化提取算法。首先构建语义高斯边缘化矩形窗函数进行融合滤波处理,通过文本切分把大量的信息熵数据进行小波基函数投影,有效剔除簇内异常数据;然后利用 Dopplerlet 变换匹配投影的自相似特性,自适应匹配语义的非线性谱特征,在 Hilbert 张成子空间中,实现对语义特征的提取和优化表达,再完成提取。仿真实验表明,该算法提高了对网络数据特征语义的表达能力,能有效区分差异网络数据中的冗余数据和残差数据,提高对杂细微差异化网络数据的检测识别和检索能力。

**关键词** Dopplerlet 变换,语义,特征提取

**中图分类号** TP391 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2015.8.055

## Optimized Semantic Extraction Algorithm of Complex Subtle Differentiated Network Data Characteristics

YANG Wei-jie

(School of Computer and Information Engineering, Beijing Technology and Business University, Beijing 100048, China)

**Abstract** The complex subtle differentiated network data need to be extracted for semantic, and it is a key technology to realize the accurate identification and retrieval of Web network data. The network data have nonlinear and random distribution, subject distribution is wide, and the update frequency is fast, so it is difficult to extract. A semantic optimization feature extraction algorithm for network data was proposed based on the Dopplerlet transform projection. The semantic Gauss edge rectangle window function is given, and data fusion system of difference fusion filtering is constructed. The text segmentation is taken, and massive data of information entropy are constructed, which can effectively eliminate the abnormal data in the cluster. Self similar characteristics of Dopplerlet transform are used for matching the projection. Nonlinear adaptive matching semantic spectrum feature is extracted, and maximal linearly independent group is searched out at the Hilbert subspace. Simulation results show that the algorithm can increase the feature semantic expression ability, effectively distinguish the differences between redundant data and residual data in network data, and improve the subtle error detection and retrieval capabilities for heterozygous network.

**Keywords** Dopplerlet transform, Semantics, Feature extraction

## 1 引言

在计算机网络时代,各种类型的网络传输数据需要进行融合和特征提取处理,以提高数据传输的精度和 Web 数据库的访问效率,节省网络节点的能量开销,延长网络生命周期,进而达到提高数据收集效率和准确度的目的。Web 网络数据通常为复杂、细微且具有差异化特性的大数据,对复杂细微差异化网络数据的语义特征提取和语义表达是实现网络数据挖掘和智能分析的基础,也是实现 Web 网络数据准确识别和检索的关键。研究网络数据的语义特征提取算法,在数据挖掘和网络 Web 数据库访问等领域具有重要的意义和价值。

传统方法中,语义特征提取算法主要有 K-Means 法、FCM 法、遗传算法、最大似然法、Fisher 判别法、神经网络

等,通常是一个算法或者几个算法的综合来实现对网络数据特征的聚类分析和语义特征提取<sup>[1,2]</sup>。很多专家进行了类似算法的研究并取得了一定的成果,其中,文献[3]采用模糊 FCM 算法理论,集合遗传算法,利用模糊集理论进行大型 Web 数据库的异常数据的语义特征挖掘,能在一定程度上克服硬划分的缺点,提高数据特征识别能力,但算法收敛性不好,容易导致局部收敛;文献[4]利用样本数据属性差异实现对网络数据的特征分区和语义表达,但是算法不能有效识别异常数据的非线性特征,特征识别性能不好;文献[5]提出一种基于关联维特征提取的网络异常数据的语义分析算法,取得了一定的特征识别性能,提高了对异常数据的拦截能力,但该算法无法有效识别差异化网络数据的语义特征;文献[6]提出一种基于最小二乘算法的混合蛙跳优化聚类的数据挖掘算

到稿日期:2014-10-24 返修日期:2014-12-27 本文受国家自然科学基金(61170112),中央财政支持地方高校发展专项资金:人才培养和创新创业团队建设项目(19005323132)资助。

杨伟杰(1980—),女,博士,讲师,主要研究领域为信息检索与智能信息处理,E-mail:yangwj@th.btbu.edu.cn.

法来进行语义特征挖掘;文献[7]提出一种基于语义高斯边缘化的数据库索引路径选择方法,该算法把语义相似度分解为用户查询意图的相关性指向函数,但精度不高,实现困难。复杂细微差异化网络数据的语义特征具有非线性、随机散布性,其主题分布广,更新频率大,提取困难<sup>[8-10]</sup>。

针对上述问题,本文提出一种基于 Dopplerlet 变换匹配投影的网络数据特征的语义优化提取算法。首先构建复杂细微差异化网络数据的信息融合模型。实现信号模型的构建;以此为基础,进行 Dopplerlet 变换匹配投影算法的设计,实现对语义特征提取算法的改进。通过仿真实验进行性能验证,表明了其优越性能,最后得出有效性结论。

## 2 差异化网络数据的信息融合模型与信号系统构建

### 2.1 复杂细微差异化网络数据的产生和融合模型

复杂细微差异化网络数据产生于 Web 数据库交互和网络信息交互过程中,对 Web 数据库进行采样,提取 Web 数据库中的一些有效特征(可为主题分布、更新频率等特征值),以减少查询的代价等。这些交互数据生成了复杂细微差异化网络数据海量系统模型,构建大数据库,通过特征提取采样技术获得一份针对客户的可行性样本并保存在本地客户端。客户端为了实现对数据库的方位和数据调度,就需要进行网络数据的语义特征提取。而复杂细微差异化网络数据的语义特征具有非线性和随机散布性,和主题分布广、更新频率大等特性,而且中间夹杂大量的异常数据和冗余数据,需要有效地对异常数据和冗余数据进行滤波处理,实现数据的融合,最终达到数据库访问和数据信息检索的目的。

根据上述分析,构建复杂细微差异化网络数据系统生成结构模型,假设大型网络数据库数据集  $X = \{x_1, x_2, \dots, x_n\}$ ,  $n$  是数据集  $X$  的数目,  $X$  中的每个元素都是一个  $P$  维矢量,  $X$  含有  $c$  个类别,第  $i$  个类的中心为  $v_i = \{v_{i1}, v_{i2}, \dots, v_{ip}\}$ 。构建语义自相关特征,建立表象关联知识库,实现数据查询过程中的纠错、消重、整合。

采用有向图的形式表示语义信息。假设  $G_c$  是数据库两个语义特征信息转换而成的有向图  $G_1$  和  $G_2$  的交集,得到第  $i$  时刻获得的 Web 网络数据库数据输入变量为  $\{x_{i1}, x_{i2}, \dots, x_{in}\}$ ,其对应的数据类型为  $y_i$ ,  $y_i$  的值为 1 或 -1,其中 1 代表正常, -1 代表异常。得到复杂细微差异化网络数据的异常与正常数据的判决式为:

$$z_{\max} = \max_{y=n_1}^{n_2} \{ \max_{x=m_1}^{m_2} \{ z_{xy} \} - \min_{x=m_1}^{m_2} \{ z_{xy} \} \} \quad (1)$$

融入了  $\lambda$  水平特征量聚点的概念,用于描述网络数据的差异性特征。对于两个标量时间序列  $y_1$  和  $y_2$ ,其联合概率密度函数为  $f(y_1, y_2)$ ,构建复杂细微差异化网络数据的差异化特征识别数学模型为:

$$y_i = f(x_1, x_2, \dots, x_m) \quad (2)$$

提取大型数据库中的异常数据训练样本,得到异常数据的频域模型表达式为:

$$\begin{aligned} z(t) &= s(t) + js(t) \otimes h(t) = s(t) + j \int_{-\infty}^{+\infty} \frac{s(u)}{t-u} du \\ &= s(t) + jH[s(t)] \end{aligned} \quad (3)$$

根据上述复杂细微差异化网络数据的生成系统,计算有向图  $G_1, G_2$  的语义关系高斯边缘化程度  $S_r$ ;最后综合  $S_c$  和  $S_r$  组成统一的语义高斯边缘化的矩形窗函数  $h(t)$ ,窗函数的

宽度为  $T = (2d+1)T_s, F_s = 1/T_s$ 。构建语义高斯边缘化矩形窗函数数据融合系统,进行差异化网络数据的融合滤波处理。设数据集测试数据待测试窗口的特征向量表示为:

$$E_{i,j} = \langle e_1, e_2, \dots, e_m \rangle \quad (4)$$

差异化网络数据的融合模型构建采用了自顶向下和分治的方法,其构造方式如下:令  $A = \{a_1, a_2, \dots, a_n\}$  为网络数据序列训练集的属性集,  $B = \{b_1, b_2, \dots, b_m\}$  为信息增益的类别集,  $a_i$  的语义属性值为  $\{c_1, c_2, \dots, c_k\}$ ,得到融合后的网络数据语义表达信息增益的表达式为:

$$Info(B) = - \sum_{i=1}^m p_i \times \log_2 p_i \quad (5)$$

$$Info_A(B) = \sum_{j=1}^v \frac{|B_j|}{|B|} \times Info(B_j) \quad (6)$$

$$Gain(A) = Info(B) - Info_A(B) \quad (7)$$

其中,  $B_j$  表示在训练集中含有  $a_x$  属性中的  $c_v$  值的元素集合。通过上述处理,实现了差异化网络数据的信息融合模型的构建,为下一步进行语义特征优化提取奠定系统模型基础。

### 2.2 网络数据的复杂细微差异化特征表达及信号模型设计

利用上述数据融合结果,对网络数据的复杂细微差异化特性进行信号模型表达,以此作为构建后续语义特征提取系统模型的输入向量。对网络数据进行文本切分,将文本切分为若干数据块 Chunk,将每个完整的小文本当作一个 Chunk 来进行分块,得到完整数据信息流的小波基函数为:

$$\begin{aligned} x_{id}(t+1) &= wx_{id}(t) + c_1 r_1 [r_1^{t_0 > T_0} p_{id} - x_{id}(t)] + \\ & c_2 r_2 [r_2^{t_g > T_g} p_{gd} - x_{id}(t)] \end{aligned} \quad (8)$$

式中,  $t_0$  和  $t_g$  分别表示数据块边界偏移的个体极值和全局极值进化停滞步数;  $T_0$  和  $T_g$  分别表示个体极值和全局极值需要扰动的停滞步数阈值。通过文本切分,把大量的信息熵数据进行小波基函数投影,可有效剔除簇内异常数据,提高数据融合效率。这一过程将产生复杂细微差异化数据,数据的差异化随机概率分布函数为:

$$w(e_p k_q) = \alpha \times w(s_p k_q) \quad (9)$$

令  $H$  为 Hilbert 空间,定义  $D = (d_\gamma)_{\gamma \in \Gamma}$  为  $H$  中的向量组成的基函数集。在 Dopplerlet 空间中构建网络数据的复杂细微差异化特征的信号模型,信号在 Dopplerlet 空间中相干点积功率累积尺度坐标描述为:

$$(x, v)x = \frac{t}{S}, v = f * S \quad (10)$$

进而得到网络数据的复杂细微差异化特征信号的状态空间更新迭代:

$$\hat{\mu}_{k+1} = \hat{\alpha} \hat{\mu} + (1 - \hat{\alpha}) \hat{\mu}_k \quad (11)$$

$$\hat{\sigma}_{k+1} = \hat{\beta} \hat{\sigma} + (1 - \hat{\beta}) \hat{\mu}_k \quad (12)$$

对网络数据特征进行信号尺度建模和时延耦合分解,实现信号模型构建,为实现网络数据的语义表达提供信号源和数据基础。

## 3 Dopplerlet 变换匹配投影的提出和数据特征语义提取算法优化

### 3.1 Dopplerlet 变换匹配投影的提出

在上述信号模型构建的基础上,采用 Dopplerlet 变换匹配投影算法实现网络数据的语义优化表达。复杂细微差异化网络数据的语义特征具有非线性和随机散布性,其主题分布广,更新频率大,提取困难。分析传统算法可知,传统方法采

用小波基函数投影算法进行特征提取,对复杂细微差异化网络数据特征的语义提取性能不好。本文利用 Dopplerlet 变换匹配投影的自相似特性能自适应匹配语义的非线性谱特征的特点,进行算法改进设计。Dopplerlet 变换匹配投影描述如下:令  $V$  是基函数集  $D$  中的向量的闭线性张成,  $D$  中向量的有限线性展开在空间  $V$  中是稠密的,对于网络数据的复杂细微差异化特征,其差异化信号  $f$  可被分解成:

$$f = \langle f, d_{\gamma_0} \rangle d_{\gamma_0} + R_f \quad (13)$$

其中,  $\langle f, d_{\gamma_0} \rangle d_{\gamma_0}$  是信号  $f$  在  $d_{\gamma_0}$  方向上的投影,  $R_f$  是投影后的残差信号。  $d_{\gamma_0}$  与  $R_f$  是正交的, 又因  $\|d_{\gamma_0}\| = 1$ , 所以有:

$$\|f\|^2 = |\langle f, d_{\gamma_0} \rangle d_{\gamma_0}|^2 + \|R_f\|^2 \quad (14)$$

采用 Dopplerlet 变换匹配投影, 搜索出来的复杂细微差异化网络数据基函数即为一组极大线性无关组, 其线性张成成为  $L^2(\mathbb{R})$ , 采用  $K$  次循环 Dopplerlet 变换匹配投影, 在张成子空间中更有利于实现对语义特征的提取。由此得到基于 Dopplerlet 的自适应匹配投影特征分解算法, 其实现框图如图 1 所示。

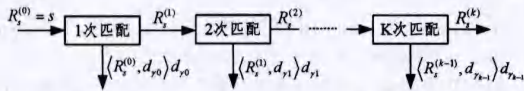


图 1 基于 Dopplerlet 的自适应匹配投影特征分解算法

### 3.2 语义特征优化提取

基于 Dopplerlet 变换匹配投影算法, 实现对复杂细微差异化网络数据特征的语义特征优化提取。改进算法实现的关键技术描述如下:

令  $H$  为 Hilbert 空间, 定义  $D = (d_{\gamma})_{\gamma \in \Gamma}$  为  $H$  中的向量组成的基函数集。在 Hilbert 空间, 设  $A$  是  $m \times n$  的实矩阵, 有  $m$  阶正交矩阵  $U$  和  $n$  阶正交矩阵  $V$ , 在有限的线性空间中展开, 则  $V$  是稠密的, 使得:

$$A = USV^T = U \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} V^T, U * U^T = I, V * V^T = I \quad (15)$$

$$\Sigma = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r), \sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0 \quad (16)$$

其中,  $A$  为  $m \times n$  矩阵,  $A * A'$  和  $A' * A$  为语义特征提取的特征向量平方根。考虑在网络数据的语义特征信号分解中采用的时频关系为非线性的原子特性, 定义相应的离散 Dopplerlet 变换为:

$$DDT_{s(n)} = \sum_{n=0}^{N-1} s(n) d_{t_0, \Delta t, t, v, c}^*(n) \quad (17)$$

语义自相关特征的最优值函数  $J^*(\cdot)$  为有界函数, 对 Doppler 信号进行窗函数调制, 经尺度伸缩, 得到语义特征提取的 Dopplerlet 原子为:

$$d(t) = \frac{1}{\sqrt{\sigma}} g\left(\frac{t-t_c}{\sigma}\right) \exp\left\{j2\pi c \left[ c + \frac{v^2(t-t_c)}{\sqrt{t^2 + v^2(t-t_c)^2}} \right]^{-1} f_c\left(\frac{t-t_c}{\sigma}\right)\right\} \quad (18)$$

改进的算法设计克服了中心频率变化对算法性能的影响, 适应了复杂细微差异化网络数据的非线性和随机散布性特征, 在更新频率大的情况下也能保证较好的语义特征性能。

## 4 仿真结果与实验分析

为了测试本文算法在实现对复杂细微差异化网络数据的语义特征优化提取中的性能, 进行仿真实验, 在 NS-2 仿真平台上进行了算法模拟。仿真实验的硬件环境为 Intel Core i3

530 1GB 内存, 操作系统为 Windows 7, 网络数据由 100 个模拟网络用户线程发出的搜索请求产生, 进行网络数据的检索和记录, 每个搜索请求对目标数据在各属性上的限制是随机产生的。数据集分为 3 类。网络设计中, 共有 10 个簇, 每个簇有 5 个节点。Dopplerlet 变换匹配投影的中心频率测试为  $f_0 = 1000\text{Hz}$ , 离散采样率为  $f_s = 10 * f_0 \text{Hz} = 10\text{kHz}$ , 带宽  $B = 1000\text{Hz}$ 。网络数据语义特征提取计算过程中, 全局迭代次数为 500, 变量的维数  $n$  为 30, 算法独立运行 30 次, 进行循环迭代, 实现算法仿真和性能对比分析。首先进行网络数据采集, 采集到的网络数据时域波形如图 2 所示。

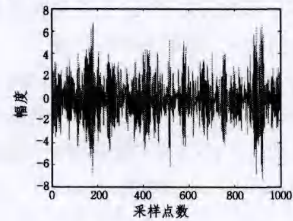


图 2 网络数据采集时域波形

从图 2 可知, 原始的网络数据含有大量的噪声干扰, 由海量复杂细微差异化特征组成, 具有非线性和随机散布性, 由于主题分布广、更新频率大的特点, 无法直接区分异常数据和冗余数据, 语义特征提取困难。采用本文算法, 进行网络数据的复杂细微差异化特征表达及信号模型设计, 采用 Dopplerlet 变换匹配投影算法, 自适应匹配语义的非线性谱特征, 得到的归一化 Dopplerlet 匹配投影谱分析结果如图 3 所示。

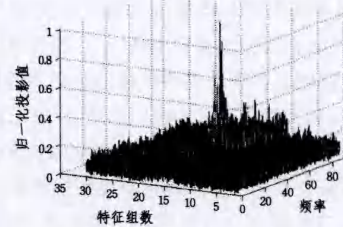


图 3 归一化 Dopplerlet 匹配投影谱分析

基于 Dopplerlet 变换匹配投影算法, 实现对复杂细微差异化网络数据特征的语义特征优化提取, 在时频坐标空间上对语义特征进行表达和分析。为对比性能, 采用本文算法和传统的波基函数投影算法进行性能测试, 得到网络数据语义特征提取表达结果, 如图 4 所示。

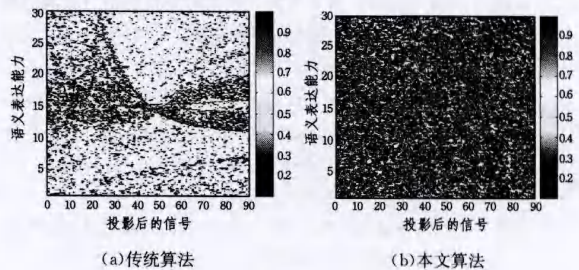


图 4 网络数据语义特征提取表达结果

分析图 4 的结果可知, 采用本文算法, 通过 Dopplerlet 变换匹配投影, 搜索出来的复杂细微差异化网络数据基函数即为一组极大线性无关组, 在张成子空间中更有利于实现对语义特征的提取。改进的算法设计克服了中心频率变化对算法

性能的影响,适合表达复杂细微差异化网络数据的非线性和随机散布性语义特征;采用本文算法,通过 Dopplerlet 变换匹配投影,计算投影后的残差信号,对复杂细微差异化网络数据的语义特征提取结果准确,语义表达能力提高明显,能有效区分差异网络数据中的冗余数据和残差数据。图 4 的语义特征表达结果直观地展示了本文算法的优越性能。

为定量分析本文算法的特征提取性能,通过 1000 次 Monte Carlo 实验,在不同的干扰数据信噪比环境下,对网络数据进行语义特征检测的性能分析,得到如图 5 所示的结果,其中,FRFT 表示傅里叶变换,FRFT-FOMCS 表示分数阶傅里叶变换。从图 5 可知,89.5% 以上的语义的相似程度值大于 0.1,随着信噪比的增大,查准率不断下降,当阈值为 0.9 时,查准率为 98.7%,采用本文算法能有效提高对复杂细微差异化网络数据的检测识别性能,结果优越于传统算法。本文算法对差异化网络数据的语义关系表达清晰,性能优越。

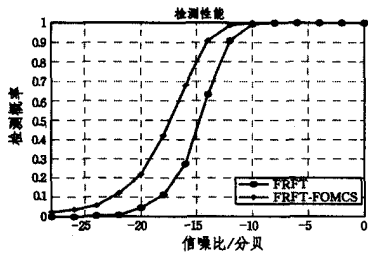


图 5 语义特征提取检测性能对比

**结束语** 对复杂细微差异化网络数据的语义特征提取和语义表达是实现网络数据挖掘和智能分析的基础,是实现 Web 网络数据准确识别和检索的关键。研究网络数据的语义特征提取算法,在数据挖掘和网络 Web 数据库访问等领域意义重大。复杂细微差异化网络数据的语义特征具有非线性和随机散布性,其主题分布广、更新频率大,提取困难。本文提出一种基于 Dopplerlet 变换匹配投影的网络数据特征的语义优化提取算法。利用 Dopplerlet 变换匹配投影的自相似特性以及能自适应匹配语义的非线性谱特征的特点,进行算法改进。研究表明,改进算法能克服中心频率变化对算法性能的影响,明显提高语义表达能力,能有效区分差异网络数据中的冗余数据和残差数据,提高对复杂细微差异化网络数据的检测识别和检索能力。

## 参考文献

- [1] Bimal K M, Gholam M A. Differential epidemic model of virus and worms in computer network [J]. International Journal of Network Security, 2012, 14(3): 149-155
- [2] Zhu Q Y, Yang X F, Yang L X, et al. Optimal control of computer virus under a delayed model [J]. Applied Mathematics and Computation, 2012, 218(23): 11613-11619
- [3] Zhu Q Y, Yang X F, Ren J, et al. Modeling and analysis of the spread of computer virus [J]. Communications in Nonlinear Science and Numerical Simulation, 2012, 17(12): 5117-5124
- [4] Zahran B M, Kanaan G. Text feature selection using particle swarm algorithm [J]. World Applied Sciences Journal, 2009, 25(7): 69-74
- [5] 姜大庆, 周勇, 夏士雄. 基于语义描述与优化的网络性能数据聚类方法 [J]. 计算机应用, 2012, 32(6): 1522-1525  
Jiang Da-qing, Zhou Yong, Xia Shi-xiong. Network performance data clustering method based on semantic description and optimization [J]. Journal of Computer Applications, 2012, 32(6): 1522-1525
- [6] Hu W M, Hu M, Maybank S. Adaboost based algorithm for network intrusion detection [J]. IEEE Transactions on Systems, Man and Cybernetics, Part B: Cybernetics, 2008, 38(2): 577-583
- [7] 何永强, 谷春英. 基于子任务区域分片下的分布式空间查询处理与并行调度方法 [J]. 科技通报, 2014, 30(1): 110-116  
He Yong-qiang, Gu Chun-ying. Distributed Spatial Query Processing and Parallel Schedule Based on Zonal Fragmentation [J]. Bulletin of Science and Technology, 2014, 30(1): 110-116
- [8] 邓兵, 陶然, 平殿发, 等. 基于分数阶傅里叶变换补偿多普勒徙动的动目标检测算法 [J]. 兵工学报, 2009, 30(10): 1034-1039  
Deng Bing, Tao Ran, Ping Dian-fa, et al. Moving-Target-Detection Algorithm with Compensation for Doppler Migration Based on FRFT [J]. Acta Armentarii, 2009, 30(10): 1034-1039
- [9] 蒋芸, 陈娜, 明利特, 等. 基于 Bagging 的概率神经网络集成分类算法 [J]. 计算机科学, 2013, 40(5): 242-246  
Jiang Yun, Chen Na, Ming Li-te, et al. Bagging-based probability Neural Network Ensemble Classification Algorithm [J]. Computer Science, 2013, 40(5): 242-246
- [10] 陈昊, 杨俊安, 庄镇泉. 变精度粗糙集的属性核和最小属性约简算法 [J]. 计算机学报, 2012, 35(5): 1011-1017  
Chen Hao, Yang Jun-an, Zhuang Zhen-quan. The Core of Attributes and Minimal Attributes reduction in Variable Precision Rough Set [J]. Chinese Journal of Computers, 2012, 35(5): 1011-1017
- [11] Han Xiao-hai, Zhang Yao-hui, Sun Fu-jun, et al. A method of determining index weights based on PCA [J]. Journal of Sichuan Ordnance, 2012, 33(10): 124-126
- [12] 钱玲飞, 杨建林, 张莉. 基于灰色关联度的学科创新力影响因素权重分析——以情报学为例 [J]. 图书情报工作, 2011, 55(16): 37-40  
Qian Ling-fei, Yang Jian-lin, Zhang Li. Weight Analysis on Influence Factors of Disciplinary Creativity based on Grey Relation-Take Information Science as an Example [J]. Library and Information Service, 2011, 55(16): 37-40
- [13] 张明卫, 王波, 张斌, 等. 基于相关系数的加权朴素贝叶斯分类算法 [J]. 东北大学学报(自然科学版), 2008, 29(7): 952-955  
Zhang Ming-wei, Wang Bo, Zhang Bin, et al. Weighted Naive Bayes Classification Algorithm Based on Correlation Coefficients [J]. Journal of Northeastern University (Natural Science), 2008, 29(7): 952-955
- [14] 韩小孩, 张耀辉, 孙福军, 等. 基于主成分分析的指标权重确定方法 [J]. 四川兵工学报, 2012, 33(10): 124-126
- [15] 刘盾, 胡培, 蒋朝哲. 一种基于粗糙集理论的属性权重构造方法 [J]. 系统工程与电子技术, 2008, 30(8): 1482-1484  
Liu Dun, Hu Pei, Jiang Chao-zhe. New methodology of attribute weights acquisition based on rough sets theory [J]. Systems Engineering and Electronics, 2008, 30(8): 1482-1484
- [16] 张文修, 吴伟志. 粗糙集理论介绍和研究综述 [J]. 模糊系统与数学, 2000, 14(4): 1-12  
Zhang Wen-xiu, Wu Wei-zhi. An Introduction and a Survey for the Studies of Rough Set Theory [J]. Fuzzy Systems and Mathematics, 2000, 14(4): 1-12
- [17] 王国胤. Rough 集理论与知识获取 [M]. 西安: 西安交通大学出版社, 2001  
Wang Guo-ying. Rough set theory and knowledge acquisition [M]. Xi'an: Xi'an Jiaotong University Press, 2001

(上接第 268 页)