

基于相容关系的基因选择方法

焦 娜^{1,2} 苗夺谦¹

(同济大学计算机科学与技术系 上海 201804)¹ (华东政法大学信息科学与技术系 上海 201620)²

摘 要 有效的基因选择是对基因表达数据进行分析的重要内容。粗糙集作为一种软计算方法能够保持在数据集分类能力不变的基础上,对属性进行约简。由于基因表达数据的连续性,为了避免运用粗糙集方法所必需的离散化过程带来的信息丢失,将相容粗糙集应用于基因的特征选取,提出了基于相容关系的基因选择方法。首先,通过 t-检验对基因表达数据进行排列,选择评分靠前的若干基因;然后,通过相容粗糙集对这些基因进一步约简。在两个标准的基因表达数据上进行了实验,结果表明该方法是可行性和有效性的。

关键词 基因表达数据,基因选择,属性约简,粗糙集,相容关系

Gene Selection with Tolerance Rough Set Theory from Gene Expression Data

JIAO Na^{1,2} MIAO Duo-qian¹

(Department of Computer Science and Technology, Tongji University, Shanghai 201804, China)¹

(Department of Information Science and Technology, East China University of Political Science and Law, Shanghai 201620, China)²

Abstract Efficient gene selection is a key procedure of the discriminant analysis of microarray data. Rough set theory is an efficient tool for further reducing redundancy. One limitation of rough set theory is the lack of effective methods for processing real-valued data. However, gene expression data sets are always continuous. Discretization methods can result in information loss. This paper investigated an approach combining feature ranking together with features selection based on tolerance rough set theory. To evaluate the performance of the proposed approach, we applied it to two benchmark gene expression data sets and compared our results with those obtained by conventional method. Experimental results illustrate that our algorithm is more effective for selecting high discriminative genes in cancer classification task.

Keywords Gene expression data, Gene selection, Attribute reduction, Rough set theory, Tolerance relation

1 引言

微阵列技术的发展为癌症的深入研究带来了曙光。癌症基因表达数据分析成为基因表达数据分析的重要研究内容之一。癌症基因表达数据分析使人们能够进行正常和疾病两种状态下的基因表达状况的比较研究,从而帮助人们更好地进行癌症诊断与分类和癌症防治等相关课题的研究。基因表达数据主要有两个特点:1)基因数量大(从几千个到几万个);2)样本数量少(从几个到几十个,最多几百个)。

在 DNA 微阵列实验中,虽然基因的数目成千上万,但实际上影响样本分类的,往往只有很少一部分的关键基因,其他的基因往往是不相关的或是冗余的。过多的基因会导致噪声的增加,加大分类算法的复杂性,并影响到分类的效果,而且与分类无关的基因将会干扰相关基因的信息。因此通过微阵列数据找到在不同类别中有显著不同表达的基因子集。最常用的差异表达基因识别方法主要有两类:基于 filter 的选择方法和基于 wrapper 的选择方法。基于 filter 的选择方法^[1]常常使用数据本身的内在特性作为评价基因的准则。t-检验和

F-检验是两种有效的选择方法。基于 wrapper 的选择方法^[2]根据分类器的某种性能来评价基因或基因子集的重要性,但 wrapper 方法的主要缺点是在基因的选择过程中需要反复地调用分类算法来评估基因的分类性能,往往造成极高的计算量。

通过过滤算法(filter)选择出来的若干个基因可能具有较强的相关性,粗糙集的属性约简能删除这种冗余,能够用更少的属性集合来有效区分决策表中的对象。粗糙集^[3-15]1982年由 Pawlak 提出以来,已经被广泛应用在特征选择和分类领域中^[3,4]。很多研究工作者致力于使用粗糙集理论进行属性约简的研究^[5-15]。然而,粗糙集处理的是离散化的数据,基因表达数据却往往都是连续的。一种方法是将基因表达数据先进行离散化^[16,17],但离散化过程必定会造成某种程度的信息损失。而相容粗糙集则放宽了等价类的严格划分条件,从而避免了一定程度的信息丢失。利用相容粗糙集方法对属性值为连续值的基因进行选取,能最大限度地保持数据集的分类能力。

本文根据相容度、综合相容度、相容关系和新的属性依赖

到稿日期:2009-11-10 返修日期:2010-01-27 本文受国家自然科学基金(60475019,60775036),教育部博士学科点专项科研基金(20060247039)资助。

焦 娜(1977-),女,博士生,主要研究方向为模式识别与智能系统、粗糙集等,E-mail:zdx_jn@163.com;苗夺谦(1964-),男,教授,博士生导师,主要研究方向为人工智能、模式识别、粗糙集等。

度提出基于相容关系的基因选择方法。该方法主要由两个部分组成:1)用 t-检验的过滤算法计算每个基因的分值,选择评分靠前的若干个基因组成条件属性集合;2)用相容粗糙集对选择出的基因进行属性约简。然后,在两个标准的基因表达数据上进行实验。最后,采用 KNN 和 C5.0 作为分类器分别进行分类测试。结果表明,本方法提取的基因组比粗糙集提取的基因组更能准确分类。

2 基于粗糙集的基因选择方法

2.1 基本概念

基因表达数据中有一个类别属性,因此可以把基因表达数据看成粗糙集里面的决策表。

定义 1 决策表 $T=(U, C \cup D, V, f)$, 其中 U 为非空有限论域; C 为条件属性集, D 为决策属性集; $V = \bigcup_{a \in C \cup D} V_a$ 为属性值域, 其中 V_a 为属性 a 的值域; $f: U \times (C \cup D) \rightarrow V$ 为信息函数。

定义 2 给定决策表 $T=(U, C \cup D, V, f)$, $B \subseteq C \cup D$, 论域 U 关于不可分辨关系 B 定义为:

$$IND(B) = \{(x, y) \in U \times U \mid \forall b \in B, b(x) = b(y)\} \quad (1)$$

显然, $IND(B)$ 是一个等价关系。

定义 3 给定决策表 $T=(U, C \cup D, V, f)$, $B \subseteq C \cup D$, 论域 U 关于不可分辨关系 B 的划分定义为:

$$U/IND(B) = \{[x_i]_B \mid x_i \in U\} \quad (2)$$

式中, $[x_i]_B$ 表示 U 中所有与 x_i 在关系 $IND(B)$ 下是等价的元素构成的集合。

定义 4 给定决策表 $T=(U, C \cup D, V, f)$, $X \subseteq U$, $B \subseteq C$, X 关于 B 的上、下近似集分别记为 $\overline{BX}$, $\underline{BX}$, 即:

$$\overline{BX} = \bigcup \{[x_i]_B \mid [x_i]_B \subseteq X\} \quad (3)$$

$$\underline{BX} = \bigcup \{[x_i]_B \mid [x_i]_B \cap X \neq \emptyset\} \quad (4)$$

定义 5 给定决策表 $T=(U, C \cup D, V, f)$, $X \subseteq U$, $B \subseteq C$, 决策属性集 D 相对条件属性集 B 的正区域定义为:

$$POS_B(D) = \bigcup_{X \in U/IND(D)} \underline{BX} \quad (5)$$

定义 6 给定决策表 $T=(U, C \cup D, V, f)$, 决策属性集 D 相对条件属性集 B 的分类质量, 即依赖度定义为:

$$\gamma_B(D) = |POS_B(D)| / |U| \quad (6)$$

2.2 t-检验

基因表达数据的特点是样本少、维数高。每个样本都记录了组织细胞中所有可测基因的表达水平。然而只有少数基因才包含了样本具体的类别信息, 大部分基因与样本类别并不相关。与分类无关的基因, 称为“无关基因”或“噪声基因”。

因此, 一般可以先对基因表达数据进行过滤。在两个类别中, 分布差异较大的样本数据能够提供样本的分类信息。所以本文采用 t-检验^[18,19], 先选取分布差异较大的若干个基因。对于二元分类问题, 有如下的 t-统计。

定义 7 给定基因 a , t-检验定义为:

$$t(a) = \frac{\mu_1 - \mu_2}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} \quad (7)$$

式中, n_i , μ_i 和 σ_i 分别为第 i ($i=1,2$) 类的样本数、基因 a 的均值和标准方差。

2.3 基于粗糙集的基因选择方法

基于粗糙集的基因选择方法是由 t-检验及粗糙集属性约简^[15]两部分组成的。首先, 通过 t-检验方法选出评分靠前的

n 个基因组成条件属性集合; 然后对所有选出的属性进行离散化; 最后, 对每一个属性进行判断, 删除掉冗余的属性, 直到满足终止条件, 即得到一个约简。

基于粗糙集的基因选择方法(GSRS)的具体步骤如下。

算法 1 GSRS 算法

1. 计算每个基因的 t-检验值, 选出评分靠前的 n 个基因组成条件属性集合 C 。
2. 离散化属性集 C
3. Set $P=C$
4. do
5. for each $a \in P$
6. if $\gamma_{P-\{a\}}(D) = \gamma_C(D)$
7. $P = P - \{a\}$
8. until $\gamma_{(P-\{a\})}(D) < \gamma_C(D)$
9. return P

值得提出的是, GSRS 算法的复杂度为 $O(n^2)$ (n 是属性个数), 而且它并不是一个完备的属性约简算法, 但是在大多数情况下确实能够有效地得到决策表的约简。

3 基于相容关系的基因选择方法

为了避免因粗糙集离散化所带来的信息损失, 本文放松了等价关系的严格定义。首先, 给出在特定属性上对象间相容度的定义; 然后, 扩展到属性子集上对象间综合相容度的定义; 最后, 在此基础上给出了相容关系定义。同时, 上、下近似、正区域及依赖度都被重新定义。

3.1 相容度

定义 8^[7,9] 给定决策表 $T=(U, C \cup D, V, f)$, $a \in C \cup D$, $x, y \in U$, 任意两个对象 x, y 关于属性 a 的相容度定义为:

$$S_a(x, y) = 1 - \frac{|a(x) - a(y)|}{|a_{\max} - a_{\min}|} \quad (8)$$

式中, $a_{\max}$ 和 $a_{\min}$ 分别是属性 a 的最大值和最小值。对于属性子集, 综合属性相容度有如下定义。

定义 9 给定决策表 $T=(U, C \cup D, V, f)$, $B \subseteq C \cup D$, $x, y \in U$, 任意两个对象 x, y 关于属性子集 B 的综合相容度定义为:

$$(x, y) \in S_{B,\tau} \text{ iff } \frac{\sum_{a \in B} S_a(x, y)}{|B|} \geq \tau \quad (9)$$

式中, τ ($0 < \tau \leq 1$) 是一个阈值。

定义 10 给定决策表 $T=(U, C \cup D, V, f)$, $B \subseteq C \cup D$, $0 < \tau \leq 1$, 论域 U 关于 B 的 τ 相容关系定义为:

$$S_{B,\tau}(x) = \{x \in U \mid (x, y) \in S_{B,\tau}\} \quad (10)$$

定义 11 给定决策表 $T=(U, C \cup D, V, f)$, $X \subseteq U$, $B \subseteq C$, $0 < \tau \leq 1$, X 关于 B 的 τ 上、下近似集分别记为 $\overline{B_\tau X}$, $\underline{B_\tau X}$, 即:

$$\underline{B_\tau X} = \{x \mid S_{B,\tau}(x) \subseteq X\} \quad (11)$$

$$\overline{B_\tau X} = \{x \mid S_{B,\tau}(x) \cap X \neq \emptyset\} \quad (12)$$

$\langle \underline{B_\tau X}, \overline{B_\tau X} \rangle$ 为相容粗糙集, 相应的正区域及依赖度有如下定义。

定义 12 给定决策表 $T=(U, C \cup D, V, f)$, $X \subseteq U$, $B \subseteq C$, $0 < \tau \leq 1$, 决策属性集 D 相对条件属性集 B 的 τ 正区域定义为:

$$POS_{B,\tau}(D) = \bigcup_{X \in U/IND(D)} \underline{B_\tau X} \quad (13)$$

定义 13 给定决策表 $T=(U, C \cup D, V, f)$, $B \subseteq C$, $0 <$

$\tau \leq 1$, 决策属性集 D 相对条件属性集 B 的 τ 分类质量, 即 τ 信赖度定义为:

$$\gamma_{B,\tau}(D) = |\text{POS}_{B,\tau}(D)| / |U| \quad (14)$$

相容粗糙集属性约简方法通过 $\gamma_{B,\tau}(D)$ 来度量属性子集的重要性。

3.2 基于相容关系的基因选择方法

基于相容关系的基因选择方法同样是由 t -检验及相容粗糙集属性约简两部分组成的。从 t -检验过滤出评分靠前的 n 个基因组成条件属性集合。然后, 从所有条件属性集合开始逐一删除掉冗余的属性, 直到满足终止条件, 即得到一个约简。

基于相容关系的基因选择方法(GSTRS)的具体步骤如下。

算法2 GSTRS 算法

1. 计算每个基因的 t -检验值, 选出评分靠前的 n 个基因组成条件属性集合 C 。
2. Set $P=C$
3. do
4. for each $a \in P$
5. if $\gamma_{P-\{a\},\tau}(D) = \gamma_{C,\tau}(D)$
6. $P = P - \{a\}$
7. until $\gamma_{P-\{a\},\tau}(D) < \gamma_{C,\tau}(D)$
8. return P

3.3 实例分析

为说明相容粗糙集属性约简的方法, 本文对表1所列的决策表进行约简。表1中前3个属性是连续值的条件属性, 最后一个为离散值的决策属性。设 $\tau=0.8$, $C=\{a,b,c\}$, $D=\{d\}$ 。

表1 数据集

Object	a	b	c	d
1	0.3	0.4	0.2	R
2	0.3	1	0.6	A
3	0.4	0.3	0.4	R
4	0.9	0.4	0.7	R
5	0.9	0.7	0.7	A
6	1	0.4	0.7	A

相容类分别是:

$$U/S_{D,\tau} = \{\{1,3,4\}, \{2,5,6\}\}$$

$$U/S_{C,\tau} = \{\{1\}, \{2\}, \{3\}, \{5\}, \{4,6\}\}$$

$$U/S_{C-\{a\},\tau} = U/S_{\{b,c\},\tau} = \{\{1\}, \{2\}, \{3\}, \{5\}, \{4,6\}\}$$

$$U/S_{C-\{b\},\tau} = U/S_{\{a,c\},\tau} = \{\{1\}, \{2\}, \{3\}, \{5\}, \{4,6\}\}$$

$$U/S_{C-\{c\},\tau} = U/S_{\{a,b\},\tau} = \{\{1,3\}, \{4,6\}, \{2\}, \{5\}\}$$

$$U/S_{C-\{a,b\},\tau} = U/S_{\{c\},\tau} = \{\{1\}, \{2\}, \{3\}, \{4,5,6\}\}$$

$$U/S_{C-\{a,c\},\tau} = U/S_{\{b\},\tau} = \{\{1,3,4,6\}, \{2\}, \{5\}\}$$

$$U/S_{C-\{b,c\},\tau} = U/S_{\{a\},\tau} = \{\{1,2,3\}, \{4,5,6\}\}$$

对于属性集 C 的 τ 下近似为:

$$\underline{C}_\tau\{1,3,4\} = \underline{a,b,c}_\tau\{1,3,4\} = \{x | S_{\{a,b,c\},\tau}(x) \subseteq \{1,3,4\}\} = \{1,3\}$$

$$\underline{C}_\tau\{2,5,6\} = \underline{a,b,c}_\tau\{2,5,6\} = \{x | S_{\{a,b,c\},\tau}(x) \subseteq \{2,5,6\}\} = \{2,5\}$$

因此, 决策属性集 D 相对条件属性集 C 的 τ 正区域为:

$$\text{POS}_{C,\tau}(D) = \bigcup_{x \in U/S_{D,\tau}} \underline{C}_\tau X = \underline{C}_\tau\{1,3,4\} \cup \underline{C}_\tau\{2,5,6\} = \{1,2,3,5\}$$

决策属性集 D 相对条件属性集 C 的 τ 信赖度为:

$$\gamma_{C,\tau}(D) = \frac{|\text{POS}_{C,\tau}(D)|}{|U|} = \frac{|\{1,2,3,5\}|}{|\{1,2,3,4,5,6\}|} = \frac{4}{6}$$

对于属性集 $C-\{a\}$, 决策属性集 D 相对条件属性集 $C-\{a\}$ 的 τ 信赖度为:

$$\gamma_{C-\{a\},\tau}(D) = \frac{|\text{POS}_{C-\{a\},\tau}(D)|}{|U|} = \frac{|\{1,2,3,5\}|}{|\{1,2,3,4,5,6\}|} = \frac{4}{6}$$

$$\gamma_{C-\{a\},\tau}(D) = \gamma_{\{b,c\},\tau}(D) = \gamma_{C,\tau}(D) = \frac{4}{6}$$

因此, 可以在属性集 C 中删除属性 a 。类似地, 决策属性集 D 相对条件属性集 $\{b,c\}-\{b\}$ 的 τ 信赖度为:

$$\begin{aligned} \gamma_{\{b,c\}-\{b\},\tau}(D) &= \frac{|\text{POS}_{\{b,c\}-\{b\},\tau}(D)|}{|U|} = \frac{|\{1,2,3\}|}{|\{1,2,3,4,5,6\}|} \\ &= \frac{3}{6} \end{aligned}$$

$$\gamma_{\{b,c\}-\{b\},\tau}(D) = \frac{3}{6} < \gamma_{C,\tau}(D) = \frac{4}{6}$$

最后, 算法终止, 约简结果为 $\{b,c\}$ 。

4 实验结果与分析

为了验证 GSTRS 算法的有效性, 选用两个公开的基因表达数据集。Lymphoma 数据集 (<http://llmpp.nih.gov/lymphoma>) 包含了 4026 个基因、96 个样本, 其中 54 个 Other type 子类和 42 个 B-cell lymphoma 子类。Liver cancer 数据集 (<http://genome-www.stanford.edu/hcc/>) 包含有 1648 个基因、156 个样本, 其中 82 个 HCCs 子类和 74 个 nontumor livers 子类。实验数据集如表2所列。

表2 两个基因表达数据集

Data sets	Genes	Classes	Samples
Lymphoma	4026	B-cell	42
		Other type	54
Liver cancer	1648	HCCs	82
		nontumor livers	74

根据 GSRS 和 GSTRS 算法, 首先, 在 Lymphoma 和 Liver cancer 基因表达数据中, 通过 t -检验选择出 50 个基因。对于有缺失值的数据用文献[20]方法进行完备化。GSRS 算法需要对连续的属性值用文献[17]的方法进行离散化。对于 GSTRS 算法, 设 $\tau=0.9$ 。约简结果如表3所列。

表3 Lymphoma 和 Liver cancer 数据集的约简结果

Data sets	Genes	Samples	GSRS	GSTRS
Lymphoma	4026	96	7	7
Liver cancer	1648	156	6	6

实验结果需要比较两方面: 一是选择的基因个数; 二是分类能力。由表3可知, GSRS 和 GSTRS 算法在 Lymphoma 和 Liver cancer 数据集中提取出的特征基因个数均相同。

接下来再比较两组基因的分类能力, 分别用 KNN, C5.0 作分类器进行实验, 并用留一交叉检验, 实验结果如表4所列。

表4 Lymphoma 和 Liver cancer 数据集的特征基因分类结果

Data sets	Lymphoma		Liver cancer	
	GSRS	GSTRS	GSRS	GSTRS
KNN	93.5%	94.8%	89.6%	92.5%
C5.0	95.2%	97.4%	91.3%	94.3%

上述实验结果表明, 无论基于粗糙集的基因选择方法还是基于相容关系的基因选择方法提取的基因都能够保持整个

基因数据集的分类能力,并且基于相容关系的基因选择方法由于避免了粗糙集离散化过程的信息丢失,提取的特征基因分类精度优于基于粗糙集的基因选择方法提取的基因。尤其选用 C5.0 作分类器时,基于相容关系的基因选择方法提取的特征基因能得到更高的分类准确率。

结束语 粗糙集不能够处理连续性数据的局限性成为它在基因表达数据研究中的主要障碍。本文给出相容度、综合相容度、相容关系和相容关系下的依赖度的定义,提出了基于相容关系的基因选择方法,并通过一个实例来加以说明。该方法由于避开了离散化过程,因此减少了信息损失,从而相对于基于粗糙集理论的基因选择方法选择的基因有更好的准确率。该方法为基因组的基因选择研究提供了一种新的尝试。

参考文献

- [1] Tibshirani R, Hastie T, Narashiman B, et al. Diagnosis of multiple cancer types by shrunken centroids of gene expression[C]// Nat'l Academy of Sciences, USA, 2002; 6567-6572
- [2] Kohavi R, John G H. Wrappers for feature subset selection[J]. Artificial Intelligence, 1997; 273-324
- [3] Banerjee M, Mitra S, Banka H. Evolutionary-rough feature selection in gene expression Data[J]. IEEE Transaction on Systems, Man, and Cybernetics, Part C: Application and Reviews, 2007, 37; 622-632
- [4] Momin B F, Mitra S, Datta Gupta R. Reduct generation and classification of gene expression data[C]// Proceeding of First International Conference on Hybrid Information Technology (ICHICT06). New York, 2006; 699-708
- [5] Pawlak Z. Rough sets[J]. International Journal of Information Computer Science, 1982, 11(5); 341-356
- [6] Dubois D, Prade H. Putting rough sets and fuzzy sets together [J]. Intelligent Decision Support, 1992; 203-232
- [7] Jensen R, Shen Q. Tolerance-based and fuzzy-rough feature selection[C]// Proceedings of the 16th International Conference on Fuzzy Systems (FUZZ-IEEE'07). 2007; 877-882
- [8] Yang M, Yang P. A novel condensing tree structure for rough set feature selection[J]. Neurocomputing, 2008, 71; 1092-1100
- [9] Parthaldin N M, Shen Q. Exploring the boundary region of tolerance rough sets for feature selection[J]. Pattern Recognition, 2009, 42; 655-667
- [10] Yao Y Y, Zhao Y. Discernibility matrix simplification for constructing attribute reducts[J]. Information Sciences, 2009, 179; 867-882
- [11] 苗夺谦, 胡桂荣. 知识约简的一种启发式算法[J]. 计算机研究与发展, 1999, 36(6); 681-684
- [12] Yang X B, Xie J, Song X N, et al. Credible rules in incomplete decision system based on descriptors[J]. Knowledge-Based Systems, 2009, 22; 8-17
- [13] Shen Q, Chouchoulas A. A rough-fuzzy approach for generating classification rules[J]. Pattern Recognition, 2002, 5; 2425-2438
- [14] Qian Y H, Dang C Y, Liang J Y, et al. On the evaluation of the decision performance of an incomplete decision table[J]. Data & Knowledge Engineering, 2008, 65; 373-400
- [15] 王国胤. 粗糙集理论与知识获取[M]. 西安: 西安交通大学出版社, 2001
- [16] 苗夺谦. 粗糙集理论中连续属性的离散化方法[J]. 自动化学报, 2001, 27(3); 296-302
- [17] Grzymala-Busse J W. Discretization of numerical attributes[M]. Klösgen W, Zytkow J, eds. In Handbook of Data Mining and Knowledge Discovery, Oxford University Press, 2002; 218-225
- [18] Golub T R, Slonim D K, Tamayo P, et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring[J]. Science, 1999, 286; 531-537
- [19] Wang L P, Feng C, Xie X. Accurate cancer classification using expressions of very few genes[J]. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2007, 4; 40-53
- [20] Grzymala-Busse J W, Grzymala-Busse W J. Handling missing attribute values[M]. by Maimon O, Rokach L, ed. In Handbook of Data Mining and Knowledge Discovery, 2005; 37-57

(上接第 183 页)

入更多的长方体来提高空间的利用率。

参考文献

- [1] Bischoff E E, Ratcliff M S W. Issues in the development of approaches to container loading [J]. International Journal of Management Science, 1995, 23(4); 377-390
- [2] Lim A, Rodrigues B, Wang Y. A multi-faced buildup algorithm for three-dimensional packing problems [J]. International Journal of Management Science, 2003, 31(6); 471-481
- [3] Gehring H, Bortfeldt A. A genetic algorithm for solving the container loading problem [J]. International Transactions in Operational Research, 1997, 4; 401-418
- [4] Bortfeldt A, Gehring H. A tabu search algorithm for weakly heterogeneous container loading problems [J]. OR Spectrum, 1998, 20(4); 237-250
- [5] Bortfeldt A, Gehring H. A hybrid genetic algorithm for the container loading problem [J]. European Journal of Operational Research, 2001, 131; 143-161
- [6] Gehring H, Bortfeldt A. A parallel genetic algorithm for solving the container loading problem [J]. International Transactions in Operational Research, 2002, 9(4); 497-511
- [7] Loh T H, Nee A Y C. A packing algorithm for hexahedral boxes [C]// Proceedings of the Conference of Industrial Automation. Singapore, 1992; 115-126
- [8] 张德富, 彭煜, 朱文兴, 等. 求解三维装箱问题的混合模拟退火算法 [J]. 计算机学报, 2009, 32(11); 2147-2156
- [9] Huang Wen-Qi, He Kun. A caving degree approach for the single container loading problem [J]. European Journal of Operational Research, 2009, 196(1); 93-101
- [10] Parreño F, Alvarez-Valdes R, Oliveira J F, et al. A maximal-space algorithm for the container loading problem [J]. INFORMS Journal on Computing, 2008, 20(3); 412-422
- [11] Beasley J E. OR-Library: Distributing test problems by electronic mail [J]. Journal of the Operational Research Society, 1990, 41(11); 1069-1072