用于多标记学习的局部顺序分类器链算法*

邱继钊[1,2]，计　华[1,2]，张化祥[1,2]
(1. 山东师范大学 信息科学与工程学院，济南 250014；2. 山东省分布式计算机软件新技术重点实验室，济南 250014)

摘　要： 标记间的相关性在分类问题中具有重要作用，目前有研究将标记相关性引入多标记学习，通过分类器链的形式将标记结果引入属性空间，为学习其他标记提供有用信息。分类器链中标记的预测顺序具有随机性，分类结果存在着很大的不确定性与不稳定性，且容易造成错误信息的传播。为此充分考虑标记的局部分布特性，提出了一种局部顺序分类器链算法，解决分类器链中分类器顺序问题。实验表明，该算法性能优于其他常用多标记学习算法。

关键词： 多标记学习；标记相关性；分类器链；K-近邻
中图分类号： TP391　　**文献标志码：** A　　**文章编号：** 1001-3695(2013)09-2606-04
doi:10.3969/j.issn.1001-3695.2013.09.011

Locally ordinal classifier chains for multi-label learning

QIU JI-zhao[1,2], JI Hua[1,2], ZHANG Hua-xiang[1,2]
(1. *School of Information Science & Engineering*, *Shandong Normal University*, *Jinan* 250014, *China*; 2. *Shandong Provincial Key Laboratory for Novel Distributed Computer Software Technology*, *Jinan* 250014, *China*)

Abstract: The correlation among different labels plays an important role in classification problems, and recent studies have taken into account label correlation during multi-label learning. The label information is marked into the attribute space through the classifier chains and provides useful information for the other labels during the classification process. The classification results are indeterminate and instable because of the random classifier order in the classifier chain. Besides, it may cause to propagate the error label information. This paper fully considerd the local distribution of instance labels, and proposed a locally ordinal classifier chain algorithm. Experimental results show that, the new algorithm outperforms the other commonly used multi-label algorithms most of the time.

Key words: multi-label learning; label correlation; classifier chains; K-NN

0　引言

传统监督学习中，一个样例仅与一个类别标记相关联，这类问题称为单标记分类问题。但在现实世界的分类问题中，一个样例可能同时具有多个不同的类别标记。例如，一幅图片可能因为含有多个语义概念而被标记上不同的关键词，一个文档可能由于涉及不同的内容而具有多个不同的主题[1]。本文称这类问题为多标记分类问题。

多标记学习已逐渐成为国际机器学习界一个新的研究热点。在多标记学习中，一种最直接也是最常用的方法就是问题转换方法，通过将多标记分类问题转换为一个或者多个单标记分类问题，采用现有的单标记分类算法对多标记数据进行分类。BR(binary relevance)[2~4]方法是多标记学习中最常见的问题转换方法，它将每个标记的预测看做是一个独立的单标记分类问题，并为每个标记训练一个独立的分类器，用全部训练数据对每个分类器进行训练，但该算法忽略了标记间的相互关系，往往无法获得令人满意的分类效果。PW(pairwise)[5]方法同样可以作为一种问题转换方法用于多标记分类问题，通过对比标记集合中任意两个标记间的关系，建立 $m(m-1)/2$ 个分类器(m 为标记集合中标记的总数)。每个分类器在两个标记间进行投票，然后组合所有的投票结果作为最终的多标记分类结果。若基础分类器的时间复杂度是 $O(n)$，那么该算法的时间复杂度为 $O(m(m-1)/2\cdot n)$，所以该算法在处理标记数目较多的数据时会变得比较困难。LP(label powerset)[6]是另一种常用的问题转换方法。它将训练数据中的每种标记组合作为一种新标记，通过这种方式将多标记数据转换为单标记数据。尽管该算法在一定程度上考虑了标记间的相关性，但该算法仍有不足，即时间复杂度问题(指数型)和过拟合训练数据问题(只能为训练数据中存在的标记组合建立模型，不能预测新的标记组合)。

问题转换方法将多标记问题转换为单标记问题，这种方式忽略了标记间的相互关系。在分类过程中，一个标记的信息可能会对学习另一个标记提供有用信息，特别是那些含有少量训练样例的标记，考虑标记相关性能够减小因数据不足造成的误差，所以在分类过程中考虑标记间的相关性非常重要[7~10]。近年来，从标记相关性出发处理多标记分类问题已逐渐成为多

收稿日期： 2012-12-21；**修回日期：** 2013-02-01　　**基金项目：** 国家自然科学基金资助项目(61170145)；国家教育部高等学校博士点专项基金资助项目(20113704110001)；山东省自然科学基金和科技攻关计划资助项目(ZR2010FM021，2008B0026，2010G0020115)

作者简介： 邱继钊(1989-)，男，硕士，主要研究方向为机器学习、数据挖掘(qiuji.zhao@163.com)；计华，副教授，硕导，主要研究方向为机器学习、模式识别及 Web 挖掘等；张化祥，教授，博导，主要研究方向为机器学习、模式识别及 Web 挖掘等.

标记学习中一个新的研究方向。Read 等人[11]将标记相关性成功引入多标记学习问题,提出了分类器链(classifier chains, CC)算法,Zhang 等人也相继提出了 LEAD[12]、ML-LOC[13]等标记相关性算法。

1　CC 算法介绍

在本文中,多标记学习描述如下:已知一个定义在实数域 R 上的 d 维输入空间 $X=R^d$,每个样例由一个含有 d 维属性值的向量表示 $\boldsymbol{x}=[x_1,x_2,\cdots,x_d]$;一个有限标记集合 $L=\{l_1,l_2,\cdots,l_m\}$,每个样例与该集合的一个子集相关联,即每个样例仅具有该集合中的部分标记;这里用一个 m 维向量表示该子集 $\boldsymbol{y}=[y_1,y_2,\cdots,y_m]$(若样例含有标记 l_i,则 $y_i=1$,否则 $y_i=0$);训练样例集合表示为 $D=\{(x_i,y_i)\mid i=1,2,\cdots,n\}$。多标记分类问题的任务就是根据训练样例集合学得分类器,当输入一个未分类数据时,通过分类器预测样例可能具有的标记。

CC 算法实质上仍是 BR 方法,与 BR 不同的是,它将先前分类器得到的标记结果(通过 0 和 1 表示)加入到属性空间,为其他标记的预测提供信息,这样就形成了一个分类器链。CC 算法的基本思想:对于任意样例 x,首先预测 x 是否具有第一个标记 l_1,将结果(通过 0 和 1 表示)加入到 x 的属性中,得到属性扩展后的新 x;然后再预测是否具有第二个标记 l_2,此时采用的是属性扩展后的新 x,将结果加入到 x 的属性中,之后依次对所有标记进行预测,形成一个分类器链结构。在预测一个样例是否具有某个标记时,CC 算法会将之前得到的所有标记结果加入到样例属性中,为预测该标记提供信息。表 1 给出了 BR 与 CC 算法执行过程的对比,为表示方便,假设属性值由 0/1 表示,样例 $\boldsymbol{x}=[0,1,0,1,0,0,1,1,0]$。

表 1　BR 与 CC 算法比较

(1)BR's transformation		(2)CC's transformation	
h:　$x\rightarrow$	y	h:$x\rightarrow$	y
h_1:[0,1,0,1,0,0,1,1,0]	1	h_1:[0,1,0,1,0,0,1,1,0]	1
h_2:[0,1,0,1,0,0,1,1,0]	0	h_2:[0,1,0,1,0,0,1,1,0,1]	0
h_3:[0,1,0,1,0,0,1,1,0]	0	h_3:[0,1,0,1,0,0,1,1,0,1,0]	0
h_4:[0,1,0,1,0,0,1,1,0]	1	h_4:[0,1,0,1,0,0,1,1,0,1,0,0]	1
h_5:[0,1,0,1,0,0,1,1,0]	0	h_5:[0,1,0,1,0,0,1,1,0,1,0,0,1]	0

Cheng 等人[14]在 CC 算法的基础上进行改进,提出了 PCC(probabilistic classifier chains)算法,对于给定的待测样例 $\boldsymbol{x}$,每种标记组合的概率可以通过概率的乘积规则计算得到

$$P_{\boldsymbol{x}}(\boldsymbol{y})=P_{\boldsymbol{x}}(y_1)\cdot\prod_{j=2}^{L}P_{\boldsymbol{x}}(y_j\mid y_1,\cdots,y_{j-1})\tag{1}$$

当 $h(\cdot)$是一个概率性的分类器时,该公式可改写为

$$P_{\boldsymbol{x}}(\boldsymbol{y})=h_1(\boldsymbol{x})\cdot\prod_{j=2}^{L}h_j(\boldsymbol{x},y_1,\cdots,y_{j-1})\tag{2}$$

但是,CC 方法中存在分类器链中分类器的顺序问题,即先预测哪个标记后预测那个标记的问题。如果一个标记的结果是错误的,那么这个错误信息就会传到之后所有分类器,对分类结果造成干扰。所以标记预测的顺序问题对分类结果有着很大的影响。

CC 分类器链中标记的预测顺序是随机的,具有很大的不稳定性,Read 等人[15]提出了 ECC(ensembles of classifier chains)算法来解决该问题,即取多个随机顺序并投票的方式,但分类效果仍存在着很大的不确定性与不稳定性。基于以上认识,本文提出了一种基于局部标记顺序的分类器链算法,解决了分类器链中的分类器顺序问题。

2　局部顺序分类器链算法 LOCC

CC 算法中标记的信息是顺序传递的。一般情况下,如果一个标记的预测结果是错误的,那么这个错误信息就会一直传下去,而且在接下来的预测过程中可能会导致多个错误信息,之后所有标记的预测结果是错误的可能性就越大,使得分类结果与真实结果相差甚大;如果一个标记的预测结果是正确的,那么之后所有标记的预测结果是正确的可能性也就越大。所以标记预测顺序的选择至关重要。不同标记间具有一定的相关性,样例所具有的标记间的相关性则更强,彼此间提供的信息也就更加有效和准确;在分类过程中优先考虑这些标记。

本文充分考虑标记的局部分布特性,从概率角度给出样例可能具有的标记,并根据可能性大小对分类器进行排序。对一个待测样例,首先找到它在训练集中的 k 个最近邻,然后对 k 个近邻中的所有标记结果进行统计,得到一个标记的排序序列。因为样例具有该序列中排在首位的标记的可能性最大,所以误分的可能性也就越小;同样地,序列中排在前面的标记比排在后面的标记的可能性要大。标记预测的顺序按照序列中的先后顺序进行,可以在很大程度上避免错误信息的出现和传播,降低误分概率,有利于提高算法分类效果。

假设待测样例 $\boldsymbol{x}$ 的 k 个近邻为$\{x_1,x_2,\cdots,x_k\}$,k 个近邻相应的标记为$\{y_1,y_2,\cdots,y_k\}$,首先进行标记统计:

$$w_j=\sum_{i=1}^{k}y_{i,j}\quad j\in 1\cdots m\tag{3}$$

$$\boldsymbol{w}=[\boldsymbol{w}_1,\boldsymbol{w}_2,\cdots,\boldsymbol{w}_m]\tag{4}$$

其中:$y_{i,j}$表示在 y_i 中标记 l_j 的取值(0/1 表示);$\boldsymbol{w}_i$ 表示 k 个近邻中标记 l_i 出现的次数;$\boldsymbol{w}$ 为所有标记结果的统计向量。

向量 $\boldsymbol{w}_i$ 表示第 i 个待测样例按标记出现的概率进行排序,得到一个标记的排序序列 $S(\boldsymbol{w}_i)$。在预测标记时根据该序列对分类器依次进行排序,即得到有序分类器链 $h=h(\boldsymbol{w}_i)$。LOCC 算法描述如下:

输入:训练集 $D=\{(x_1,y_1),(x_2,y_2),\cdots,(x_n,y_n)\}$,K 近邻的参数 k,待标记样例 x。
输出:预测的标记集 $\boldsymbol{y}$。
过程:
　确定未见样例 $\boldsymbol{x}$ 的 k 个近邻;
　for $i=1$ to k
　　for $j=1$ to m
　　do 根据式(3)和(4)对标记进行统计得到有序分类器链$h=h(\boldsymbol{w}_i)$
　　end
　end
　$\boldsymbol{x}=[x_1,x_2,\cdots,x_d]$
　$\boldsymbol{y}=[\,]$
　　for $i=S(\boldsymbol{w}_i)$
　　do $\boldsymbol{x}\leftarrow[\boldsymbol{x},y_i]$
　　$y_i=h_i(\boldsymbol{x})$
　　$\boldsymbol{y}=[\boldsymbol{y},y_i]$
　end
　return $\boldsymbol{y}$

可以看出,k 个近邻具有的标记仅仅是标记集合中的部分标记,对于没有出现的标记,其顺序是随机的。为了避免结果不稳定,采用多个随机顺序的投票方式处理这些标记。如表 2 所示,假设近邻中仅具有 l_1、l_2、l_3 三个标记,对于未出现的标记 l_4、l_5、l_6,选取了五个随机顺序进行分类,得到五个不同分类结果$[y_1,y_2,y_3,y_4,y_5]$并对标记进行投票,得到最终结果 $\boldsymbol{v}$,最后根据给定的阈值 t 确定样例所具有的标记。表 2 中 t 取值 0.5,如果结果大于 0.5,就认为样例具有该标记。

表 2　LOCC 算法中对近邻中不存在标记的投票过程

	y_1	y_2	y_3	y_4	y_5	y_6
$y_1=$	1	0	1	1	0	0
$y_2=$	1	0	1	0	0	1
$y_3=$	1	0	1	0	1	0
$y_4=$	1	0	1	1	0	1
$y_5=$	1	0	1	0	0	1
$v=$	1.0	0.0	1.0	0.4	0.2	0.6
$y=f_{t=0.5}(v)=$	1	0	1	0	0	1

这样会产生一个置信度向量 $v=[v_1,v_2,\cdots,v_m]$，其中 v_j 表示标记 l_j 的置信度。对于给定的 q 个分类器链 $h_1,h_2,\cdots,h_q$，标记 l_i 置信度由以下公式计算得到：

$$v_j=\frac{1}{q}\sum_{i=1}^{q}y_{i,j}\quad j\in 1,\cdots,m \tag{5}$$

其中：$y_{i,j}$ 表示在分类器链 h_i 中第 j 个标记的预测结果。本文采用一个阈值函数将标记分成相关标记与非相关标记两部分。

$$y=f_t(v)$$

$$t=\arg\min_t \left\| LC-\left(\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m}y_j^{v_j\geqslant t}\right)\right\| \tag{6}$$

其中：$y_j=\begin{cases}1 & \text{if } v_j\geqslant t\\ 0 & \text{if } v_j<t\end{cases}$；$LC$ 表示数据集中平均每个样例所具有的标记数目，即 $LC=\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m}y_{i,j}$。

3　实验结果及分析

本文所有实验均在 MATLAB 平台上完成，选用 image 和 scene 两个数据集。实验中采用的对比算法有 ECC、LEAD、ML-KNN[16]、BSVM[1]，其中 ECC 和 LEAD 考虑了标记相关性，ML-KNN 和 BSVM 为传统多标记分类算法（未考虑标记相关性）。使用 Hamming loss、one-error、coverage、ranking loss、average precision 等指标对实验结果进行评估，采用十折交叉验证。实验结果如表 3 ~ 5 所示（表中准确率高的数据用粗体标出）。

表 3　LOCC 与其他算法在 image 数据集上的实验结果比较

evaluation criterion	算法				
	LOCC	ECC	LEAD	ML-KNN	BSVM
Hamming loss ↓	**0.170 ±0.006**	0.180 ±0.015	0.173 ±0.011	0.170 ±0.008	0.176 ±0.007
one-error ↓	**0.278 ±0.017**	0.289 ±0.026	0.313 ±0.026	0.330 ±0.026	0.314 ±0.021
coverage ↓	**0.184 ±0.007**	0.199 ±0.020	0.184 ±0.007	0.194 ±0.020	0.189 ±0.021
ranking loss ↓	0.178 ±0.017	0.225 ±0.024	**0.164 ±0.018**	0.175 ±0.019	0.169 ±0.019
average precision ↑	**0.805 ±0.014**	0.794 ±0.016	0.799 ±0.017	0.792 ±0.017	0.796 ±0.015

表 4　LOCC 与其他算法在 scene 数据集上的实验结果比较

evaluation criterion	算法				
	LOCC	ECC	LEAD	ML-KNN	BSVM
Hamming loss ↓	0.085 ±0.005	0.096 ±0.010	0.098 ±0.005	**0.084 ±0.008**	0.104 ±0.006
one-error ↓	**0.208 ±0.027**	0.226 ±0.034	0.264 ±0.024	0.219 ±0.029	0.250 ±0.027
coverage ↓	0.087 ±0.007	0.091 ±0.008	0.087 ±0.007	**0.078 ±0.010**	0.089 ±0.009
ranking loss ↓	**0.074 ±0.010**	0.088 ±0.013	0.087 ±0.009	0.076 ±0.012	0.089 ±0.011
average precision ↑	**0.871 ±0.012**	0.852 ±0.016	0.848 ±0.014	0.869 ±0.017	0.849 ±0.016

表 5　LOCC 在不同近邻取值下的实验结果比较（image 数据集）

evaluation criterion	number of nearest neighbors considered					
	$k=20$	$k=21$	$k=22$	$k=23$	$k=24$	$k=25$
Hamming loss ↓	0.172 ±0.004	0.171 ±0.003	0.170 ±0.006	0.170 ±0.007	0.171 ±0.007	0.173 ±0.011
one-error ↓	0.279 ±0.009	0.280 ±0.012	0.278 ±0.017	0.276 ±0.014	0.275 ±0.020	0.276 ±0.010
coverage ↓	0.187 ±0.007	0.185 ±0.008	0.184 ±0.007	0.183 ±0.003	0.182 ±0.010	0.183 ±0.004
ranking loss ↓	0.183 ±0.010	0.181 ±0.013	0.180 ±0.017	0.178 ±0.013	0.179 ±0.010	0.177 ±0.013
average precision ↑	0.802 ±0.012	0.804 ±0.016	0.805 ±0.014	0.807 ±0.016	0.806 ±0.012	0.808 ±0.011

根据以上实验数据，对 LOCC 算法作了以下分析：

a）如表 3 所示，在 image 数据集上，LOCC 算法在 Hamming loss、one-error、coverage、average precision 四项评价指标上要优于其他分类算法，其中 one-error 表现最为突出。LOCC 算法充分考虑了标记的局部分布特性，优先判断可能性较大的那些标记，在很大程度上避免了错误信息的出现，同时也加强了标记间的关联性，使得其他分类器继承的标记信息在最大程度上是正确的，提高了算法的分类效果。

b）表 4 结果表明，LOCC 算法在 one-error、ranking loss、average precision 三项评价指标上好于其他算法，其中 one-error 与 ranking loss 表现较为突出；Hamming loss、coverage 与表现最好的 ML-KNN 算法几乎持平；与同样考虑了标记相关性的算法相比，其五项评价指标均明显好于 ECC 与 LEAD，其中 ranking loss 的表现最为突出，该指标表明了样本对其所属标记的隶属度低于对其非属标记的隶属度的可能性。总体看来，LOCC 算法在 scene 数据集上的分类效果较好。

c）表 5 与图 1 ~ 5（image 数据集）结果表明，随着 k 值的不断增加，标记的统计结果逐渐趋于稳定，LOCC 的分类效果也逐渐趋于稳定；而 ECC 因为具有较强的随机性，其分类效果呈现了一定的不稳定性。总体来看，引入标记相关性的算法要优于其他传统分类算法，LOCC 算法要优于其他标记相关性算法。

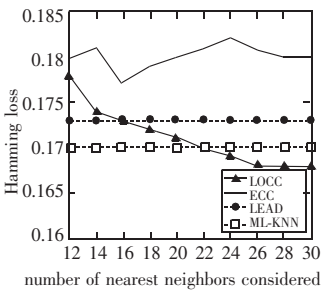

图 1　k 取值对 Hamming loss 的影响

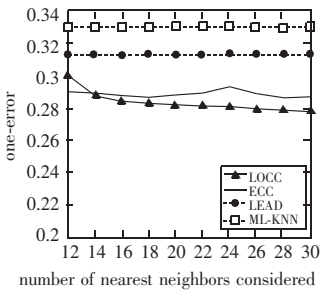

图 2　k 取值对 one-error 的影响

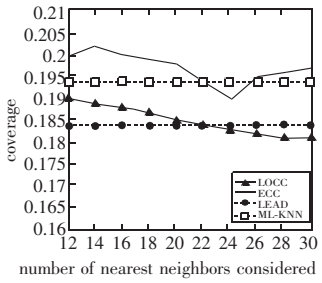

图 3　k 取值对 coverge 的影响

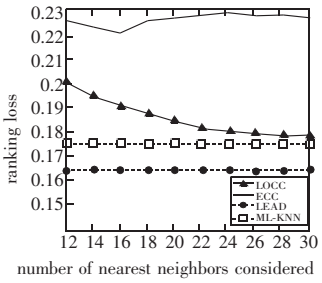

图 4　k 取值对 ranking loss 的影响

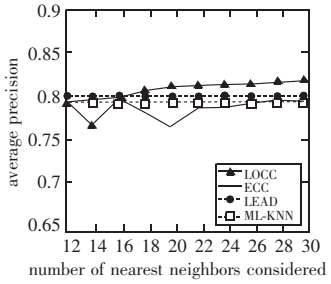

图 5　k 取值对 average precision 的影响

4　结束语

由于标记相关性的存在,一个标记的信息可以通过分类器链形式传递给其他标记,为学习其他标记提供有用信息。由于分类器链中错误信息传递问题的存在,所以分类器链中分类器的顺序选择至关重要。本文充分考虑标记的局部分布特性,提出了局部顺序分类器链算法,通过 K-NN 找出样例最可能具有的标记,并根据可能性大小对分类器进行排序。对于近邻中不存在的标记采用多个随机顺序的处理方式,这样在很大程度上降低了错误信息的传播,提高了算法的分类效果。下一步笔者将在标记相关性的基础上引入样例间的相关性,并寻求更好的标记排序方法。

参考文献:

[1] BOUTELL M R, LUO Jie-bo, SHEN Xi-peng, *et al*. Learning multi-label scene classification [J]. Pattern Recognition, 2004, 37(9): 1757-1771.

[2] ZHANG Min-ling, ZHOU Zhi-hua. A k-nearest neighbor based algorithm for multi-label classification[C]//Proc of IEEE International Conference on Granular Computing. New York: IEEE Press, 2005: 718-721.

[3] GODBOLE S, SARAWAGI S. Discriminative methods for multi-labeled classification[C]//Proc of the 8th Pacific-Asia Conference on Knowledge Discovery and Data Mining. Berlin: Springer, 2004: 22-30.

[4] TSOUMAKAS G, KATAKIS I. Multi label classification: an overview [J]. International Journal of Data Warehousing and Mining, 2007, 3(3): 1-13.

[5] HULLERMEIER E, FURNKRANZ J, CHENG W, *et al*. Label ranking by learning pairwise preferences [J]. Artificial Intelligence, 2008, 172(16): 1897-1916.

[6] DIMOU A, TSOUMAKAS G, MEZARIS V, *et al*. An empirical study of multi-label learning methods for video annotation[C]//Proc of the 7th International Workshop on Content-Based Multimedia Indexing. Washington DC: IEEE Computer Society, 2009: 19-24.

[7] JI Sui-wang, TANG Lei, YU Shi-peng, *et al*. Extracting shared subspace for multi-label classification[C]//Proc of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM Press, 2008: 381-389.

[8] READ J, PFAHRINGER B, HOLMES G. Multi-label classification using ensembles of pruned sets[C]//Proc of the 8th IEEE International Conference on Data Mining. New York: IEEE Press, 2008: 995-1000.

[9] LOZA M E, FÜMKRANZ J. Efficient pairwise multi-label classification for large-scale problems in the legal domain[C]//Proc of ECML-PKDD '08 European Conference on Machine Learning and Knowledge Discovery in Databases. Berlin: Springer, 2008: 50-65.

[10] SUN Liang, JI Shui-wang, YE Jie-ping. Hyper-graph spectral learning for multi-label classification [C]//Proc of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM Press, 2008: 668-676.

[11] READ J, PFAHRINGER B, HOLMES G, *et al*. Classifier chains for multi-label classification [C]//Proc of ECML/PKDD. 2009: 254-269.

[12] ZHANG Min-ling, ZHANG K. Multi-label learning by exploiting label dependency [C]//Proc of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM Press, 2010: 999-1007.

[13] HUANG Sheng-jun, ZHOU Zhi-hua. Multi-label learning by exploiting label correlations locally[C]//Proc of the 26th AAAI Conference on Artificial Intelligence. 2012: 949-955.

[14] CHENG W, HÜLLERMEIER E. Bayes optimal multi-label classification via probabilistic classifier chains[C]//Proc of the 27th International Conference on Machine Learning. Haifa: Omni Press, 2010: 279-286.

[15] READ J, PFAHRINGER B, HOLMES G, *et al*. Classifier chains for multi-label classification[J]. Machine Learning, 2011, 85(3): 333-359.

[16] ZHANG Min-ling, ZHOU Zhi-hua. ML-KNN: a lazy learning approach to multi-label learning [J]. Pattern Recognition, 2007, 40(7): 2038-2048.

(上接第 2593 页)

[2] ADOMAVICIUS G, KWON Y. Improving aggregate recommendation diversity using ranking-based techniques[J]. IEEE Trans on Knowledge and Data Engineering, 2012, 24(5): 896 -911.

[3] ZIEGLER C, MCNEE S M, KONSTAN J A, *et al*. Improving recommendation lists through topic diversification[C]//Proc of the 14th International Conference on World Wide Web. New York: ACM Press, 2005: 22-32.

[4] JAMBOR T, WANG Jun. Optimizing multiple objectives in collaborative filtering[C]//Proc of the 4th ACM Conference on Recommender Systems. New York: ACM Press, 2010: 55-62.

[5] VARGAS S, CASTELLS P. Rank and relevance in novelty and diversity metrics for recommender systems[C]//Proc of the 5th ACM Conference on RecSys. New York: ACM Press, 2011: 109-116.

[6] SARWAR B, KARYPIS G, KONSTAN J, *et al*. Item-based collaborative filtering recommendation algorithms[C]//Proc of the 10th International Conference on World Wide Web. New York: ACM Press, 2001: 285-295.

[7] 张爱君,秦新强,龚春琼. 求解 0-1 二次规划问题的迭代禁忌搜索算法[J]. 计算机工程, 2012, 38(1): 140-142.

[8] MILLER B, ALBERT I, LAM S, *et al*. Movielens unplugged: experiences with an occasionally connected recommder system[C]//Proc of the 8th International Conference on Intelligence User Interfaces. New York: ACM Press, 2003: 263-266.

[9] CREMONESI P, KOREN Y, TURRIN R. Performance of recommender algorithms on top-*N* recommendation tasks[C]//Proc of the 4th ACM Conference on RecSys. New York: ACM Press, 2011: 39-46.

[10] HERLOCKER J L, KONSTAN J A, TERVEEN L, *et al*. Evaluating collaborative filtering recommender systems[J]. ACM Trans on Information Systems, 2004, 22(1): 5-53.