

分段光滑的半监督支持向量分类机

范旭慧 张捷 班登科

(西北工业大学电子信息学院 西安 710072)

摘要 为了解决半监督支持向量分类优化模型中的非凸非光滑问题,基于分段逼近的思想提出了一个分段函数,并以此逼近非凸非光滑的目标函数。给出的分段函数可以根据不同的精度要求选择不同的逼近参数,同时构造出基于上述分段函数的光滑半监督支持向量机模型。采用了 LDS(Low Density Separation)算法求解模型,分析了其对对称铰链损失函数的逼近精度。理论分析和数值实验结果都证明分段光滑的半监督支持向量机的分类性能和效率优于以往提出的光滑模型。

关键词 算法,分类器,优化,半监督支持向量机,分段函数,光滑技术

中图分类号 TP391.4 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2016.6.054

Piecewise Smooth Semi-supervised Support Vector Machine for Classification

FAN Xu-hui ZHANG Jie BAN Deng-ke

(School of Electronics and Information, Northwestern Polytechnical University, Xi'an 710072, China)

Abstract In order to focus on the non-smooth and non-convex problems of the semi-supervised support vector machine, a piecewise function based on piecewise ideas was proposed to approach the non-convex and non-smooth objective function. The approach degree of the piecewise function to objective function can be chosen according to the accuracy demand. A new piecewise smooth semi-supervised support vector machine (PWSS³VM) model based on piecewise function was constructed. LDS algorithm was applied to solve the model and its approximation performance to the symmetric hinge loss function was analyzed. Theoretical analysis and numerical experiments confirm that PWSS³VM model has better classification performance and higher classification efficiency than previous smooth models.

Keywords Algorithms, Classifiers, Optimization, Semi-supervised support vector machine, Segment function, Smooth technologies

1 引言

随着通信网络技术的快速发展,各行各业都积累了大量的数据,而这些数据无任何标记,对这些无标记的样本进行标记的代价高、效率低。因此根据实际需要,研究人员提出了半监督支持向量机^[1-3],它可以利用少量的标号数据和大量的未标号数据训练一个性能良好的分类器^[4,5]。为了解决 S³VM (Semi-supervised Support Vector Machine)的非凸非光滑问题,许多研究者致力于其算法的研究并取得了较大的进步。针对此问题,2005年,Chapelle和Zien采用指数函数逼近无标记样本的对称铰链损失函数,建立了一个无约束的光滑半监督支持向量机模型^[6]。2009年,刘三阳等人也提出了一族多项式光滑函数^[7],得到了多项式光滑的半监督支持向量机模型。2015年,张晓丹等人提出一个广义三次样条光滑半监督支持向量机^[5]。这些模型的性能和分类效率还有待提高。因此本文提出了分段光滑的半监督支持向量机模型,理论分析和数值实验都证明它在数据分类的性能和效率上优于以往提出的模型。

2 分段光滑的半监督支持向量机模型

2.1 分段函数的构造

定理 设 $L(|t|) = \max(0, 1 - |t|)$ 为对称铰链损失函数, $0 < k < 1$, $t_0 = -k$, $t_1 = 0$, $t_2 = k$ 是节点,则存在唯一一个逼近对称铰链损失函数的分段函数,并且这个分段函数在 t_0 , t_1 , t_2 处满足二阶光滑条件。形式如下:

$$f(t, k) = \begin{cases} L(|t|), & |t| > k \\ 1 - 2k + \sqrt{2k^2 - t^2}, & |t| \leq k \end{cases}$$

证明:如图 1 所示,设 $f_0(t, k)$ 是圆的一部分,圆心为 O , $L(|t|)$ 在区间 $|t| > k$ 的部分是 $f_0(t, k)$ 的切线,切点为 $P(-k, 1-k)$ 和 $O(k, 1-k)$,由几何知识知, OP, OQ 都垂直于 $L(|t|)$ 且圆心 O 的坐标为 $(0, 1-2k)$,半径为 $\sqrt{2k}$ 。则圆的标准方程为:

$$t^2 + [y - (1 - 2k)]^2 = (\sqrt{2k})^2$$

求解得

$$y = 1 - 2k + \sqrt{2k^2 - t^2}, |t| \leq k, 0 < y < 1$$

验证 $f(t, k)$ 在 $t_0 = -k, t_1 = 0, t_2 = k$ 的二阶光滑性。

到稿日期:2015-05-25 返修日期:2015-08-07

范旭慧(1989-),女,硕士生,主要研究方向为模式识别、机器学习,E-mail:f769397603@163.com;张捷(1964-),男,教授,硕士生导师,主要研究方向为模式识别、移动与无线通信、通信信号处理。

1) 由于 $f_0(t, k) = 1 - 2k + \sqrt{2k^2 - t^2}$ 在区间 $[-k, k]$ 是一个连续函数, 因此 $f(t_1 + 0, k) = f(t_1 - 0, k) > 0$, $f^{(d)}(t_1 + 0, k) = f^{(d)}(t_1 - 0, k)$, $d = 1, 2$;

$$2) f(t_0, k) = L(|t_0|), f'(t_0, k) = \frac{-t_0}{\sqrt{2k^2 - t_0^2}} = 1 = L'(|t_0|),$$

$$f^{(2)}(t_0, k) = -\frac{2k^2 - 2t_0^2}{(2k^2 - t_0^2)^{3/2}} = 0 = L^{(2)}(|t_0|);$$

$$3) f(t_2, k) = L(|t_2|), f'(t_2, k) = -1 = L'(|t_2|), f^{(2)}(t_2, k) = 0 = L^{(2)}(|t_2|).$$

根据 1)–3), 此定理得证。

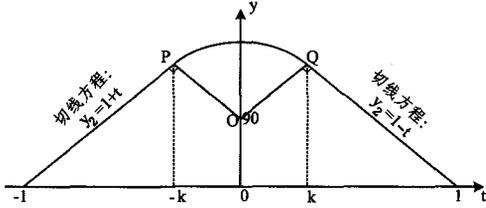


图 1 分段函数的推导示意图

2.2 新光滑半监督支持向量机模型

考虑如下的二分类问题: 假设数据集包含 n 个标记样本 X_i ($1 \leq i \leq n$) 和 m 个未标记样本 X_i ($n+1 \leq i \leq n+m$)。假设有标记样本的标签 y_i ($1 \leq i \leq n$) 是二元的, 即 $y_i = \pm 1$ 。由文献 [8, 9] 知, SVM 的目标是找到一组标记向量 $y_m = [y_{n+1}, \dots, y_{n+m}]$, 使 SVM 在数据集上训练能得到最大间隔。即让下面的目标函数最小化:

$$\min_{\omega, \gamma, \xi} \frac{1}{2} \|\omega\|_2^2 + C \sum_{i=1}^n \xi_i + C^* \sum_{i=n+1}^{n+m} \xi_i$$

$$\text{s. t. } y_i(\omega^T x_i + b) \geq 1 - \xi_i, 1 \leq i \leq n \quad (1)$$

$$|\omega^T x_i + b| \geq 1 - \xi_i, n+1 \leq i \leq n+m$$

式(1)是一个有约束条件的优化模型。修正该模型的目标函数^[6], 可得到其无约束的最优化问题:

$$\min_{\omega, \gamma, \xi} \frac{1}{2} \|\omega\|_2^2 + C \sum_{i=1}^n L(y_i(\omega^T x_i + b)) + C^* \sum_{i=n+1}^{n+m} L(|\omega^T x_i + b|) \quad (2)$$

其中, $L(|t|) = \max(0, 1 - |t|)$ 是对称铰链损失函数。模型(2)的目标函数是非凸非光滑的, 所以求解这个模型很难^[8, 9]。为了能利用快速优化算法^[6]求解, 引入了光滑技术^[6, 7], 对上述模型(2)的非光滑函数进行替换。

首先, 用光滑函数 $L^2(t) = \max(0, 1 - t^2)$ 去逼近非对称铰链损失函数 $L(t) = \max(0, 1 - t)$, 其次, 用任一光滑函数去逼近对称铰链损失函数 $L(|t|) = \max(0, 1 - |t|)$ 。文献^[6]中用指数函数 $E(t) = e^{-3t^2}$ 来逼近 $L(|t|)$, 文献^[7]则用多项式函数 $P_n(t) = \frac{1-t^2}{2} + \sum_{i=2}^n \frac{(2i-3)!!}{(2i)!!} (1-t^2)^i$ 进行逼近。虽然指数函数能够达到任一阶光滑, 但它的逼近精度不高; 多项式函数逼近的精度虽然高, 但效率不高。

本文使用分段函数来逼近对称铰链损失函数, 得到分段光滑的半监督支持向量机模型。

$$\min_{\omega, \gamma, \xi} \frac{1}{2} \|\omega\|_2^2 + C \sum_{i=1}^n L^2(y_i(\omega^T x_i + b)) + C^* \sum_{i=n+1}^{n+m} L^2(|\omega^T x_i + b|) \quad (3)$$

3 逼近误差分析

令 $1 - |t| = f(t, k) + \Gamma(t, k)$, 则误差函数为

$$\Gamma(t, k) = \begin{cases} 0 \leq 2k, & |t| > k \\ 2k - \sqrt{2k^2 - t^2} - |t| \leq 2k, & |t| \leq k \end{cases} \quad (4)$$

因此, $f(t, k)$ 对 $1 - |t|$ 的逼近误差为

$$\Gamma(k) \leq 2k, 0 < k < 1 \quad (5)$$

令 $\Gamma(t) = |e^{-3t^2} - (1 - |t|)|$, 因为 $\Gamma(t)$ 是一个只与 t 相关的函数, 因此令

$$\Gamma = \max(\Gamma(t)) = \max(|e^{-3t^2} - (1 - |t|)|) \quad (6)$$

则 Γ 就是指数函数 $E(t)$ 对 $1 - |t|$ 的逼近误差。

因为 Γ 是一个不随 t 变化的常数, 所以当 $\Gamma > \alpha$ (α 是要求达到的逼近精度) 时, $E(t)$ 满足对 $1 - |t|$ 的逼近精度要求; 而当 $\Gamma \leq \alpha$ 时, $E(t)$ 就无法满足对 $1 - |t|$ 的逼近精度要求, 此时就不能用 $E(t)$ 去逼近原函数。但 $f(t, k)$ 的逼近误差 $\Gamma(t, k)$ 则可避免这样的问题, 因为 $\Gamma(k)$ 随 k 的减小而减小, 对于任意小的 α , 只要 k 取得适度小就能满足 $\Gamma \leq \alpha$, 因此 $f(t, k)$ 对 $1 - |t|$ 可达到任意逼近精度。从文献^[7]知, 多项式光滑函数的逼近误差也可以达到任意逼近精度的要求, 但其分类效率没有分段函数的高。

各光滑函数对对称铰链损失函数的逼近精度如图 2 所示。从图中看出, 分段函数的逼近效果最好, 并且 k 越小, 它的逼近效果越好。

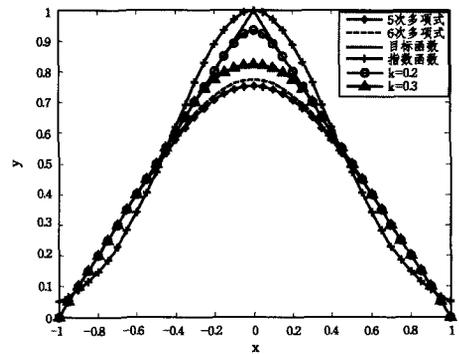


图 2 不同光滑函数的对比图

4 算法

令目标函数为

$$\min_{\omega, b} \phi(\omega, b) = \min_{\omega, \gamma, \xi} \frac{1}{2} \|\omega\|_2^2 + C \sum_{i=1}^n L^2(y_i(\omega^T x_i + b)) + C^* \sum_{i=n+1}^{n+m} f(|\omega^T x_i + b|) \quad (7)$$

采用文献^[6]中的 LDS (Low Density Separation) 算法求解式(7), 算法描述如下。

要求参数 ρ 和 C

计算 ρ -距离

1. 用 $\omega_{ij} = e^{d(i,j)} - 1$ 的边长建立一个全连图, 其中 $d(i, j)$ 代表欧几里得距离

2. 利用 Dijkstra 算法^[10]去计算所有成对点的最短路径长度 $d_{sp}(i, j)$

3. 根据 $D_{ij} = [\frac{1}{\rho} \log(1 + d_{sp}(i, j))]^2$ 来计算平方 ρ -路径距离矩阵 D ,

$$\text{其中 } D_{ij} = e^{-\frac{D_{ij}^{\rho}}{2\sigma^2}}$$

执行多维标度法(MDS)^[11]

4. $U \Delta U^T = -H D H$, 其中 $H_{ij} = \delta_{ij} - \frac{1}{n+m}$

5. 找一个合适的阈值 p 使得 $\sum_{i=1}^n \lambda_i \geq (1 - \delta) \sum \max(0, \lambda_i)$ 和 $\lambda_p \leq \delta \lambda_1$ 成立

6. x_i 的表达式可以重写为 $\bar{x}_{ik} = U \sqrt{\lambda_k}$, $1 \leq k \leq p$

训练 TSVM

7. for $i=0$ to 10 do

8. 设置 $C^* = 2^{i-10}C$

9. 在约束条件 $\frac{1}{m} \sum_{i=1}^m \omega^T x_i + b = \frac{1}{n} \sum_{i=1}^n y_i$ 下, 由梯度下降法最小化目标函数(7)

10. end for

5 数值实验

所有的实验都是基于 1.9GHZ AMD APU 和 4GB RAM 的个人计算机上的 Matlab R2011b 平台。每组数据都进行 10 次实验, 结果取 10 次实验的平均值。

为了更好地比较不同光滑半监督支持向量机的计算效率和分类性能, 用未标记样本的分类错误率和训练时间来衡量。本文采用模型(3), 其中分段函数 $f(x)$ 可替换为指数函数和多项式函数。为了方便, 将不同的光滑半监督支持向量机模型分别进行标记: 采用指数函数的光滑半监督支持向量机模型记为 ESS³ VM; 采用多项式函数的光滑半监督支持向量机模型记为 PSS³ VM; 采用分段函数的光滑半监督支持向量机模型记为 PWSS³ VM。

5.1 数据集

数据集 g50c 和 g10n 是根据文献[12]进行构造的人工数据集。Coil20 数据集^[13]是从不同角度拍摄的 20 个物体的灰度图像。Uspst 数据集是包括 USPS 数据集中关于手写体数字识别的测试部分。表 1 是所选数据集的信息。

表 1 数据集信息

数据集	种类	样本属性个数	样本总数	标记样本个数
g50c	2	50	550	50
g10n	2	10	550	50
Coil20	20	1024	1440	40
Uspst	10	256	2007	40

5.2 人工数据集上的实验

首先在 g50c 数据集上实验, 利用十折交叉验证法, 为以下 3 个参数选择合适的值, 即 $\delta=0.1, \rho=1, \sigma=\infty$ 。取 $C=0.1$, 对不同光滑参数 k 的实验结果进行比较, 结果如图 3 所示。

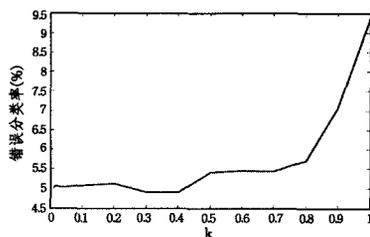


图 3 当 k 取不同值时 PWSS³ VM 的分类错误率

由图 3 知, C 值固定, k 取值越小, PWSS³ VM 模型分类错误率越低; 当 k 不断减小时, 分类器的性能趋于稳定, 这是由于舍入误差等原因使得分类器性能没有继续提高。因此, 针对不同的实际分类问题, 对于不同的取值 C , 对 k 的选取适当即可, 即惩罚参数 C 无论取何值, PWSS³ VM 模型总能找到最优的参数 k 值, 从而获得最好的分类性能。

同样利用十折交叉验证法为 g10n 的 3 个参数选择合适的值, 即 $\delta=0.1, \rho=1, \sigma=\infty$ 。对不同的光滑模型进行实验比较, 具体结果如表 2 和表 3 所列。

表 2 用 3 种光滑模型测试 g50c 的分类错误率和 CPU 时间

	PWSS ³ VM($k=0.4$)	PSS ³ VM	ESS ³ VM
错误率%($C=0.1$)	4.90	5.06	5.40
CPU_TIME(s)	2.090	2.235	1.972
错误率%($C=1$)	6.26	6.54	6.90
CPU_TIME(s)	2.018	2.163	1.900

表 3 用 3 种光滑模型测试 g10n 的分类错误率和 CPU 时间

	PWSS ³ VM($k=0.2$)	PSS ³ VM	ESS ³ VM
错误率%($C=10$)	11.2	11.2	11.3
CPU_TIME(s)	1.954	2.082	1.761
错误率%($C=100$)	9.92	10.42	10.48
CPU_TIME(s)	1.599	1.719	1.500

比较表 2 与表 3 中的 CPU 时间可以看出, PWSS³ VM 的时间相比 ESS³ VM 要高一点。但从分类样本错误率可看出: 当惩罚参数 C 取不同的值时, PWSS³ VM 都有最好的分类性能。因此可以得出结论, PWSS³ VM 模型以牺牲较少的 CPU 时间获得了最优的分类性能。

5.3 多分类数据集上的实验

对于 Coil20 和 Uspst 这样的多分类问题, 训练时采取一对多法。利用十折交叉验证法为 Uspst 和 Coil20 数据集选择参数设置。在 Uspst 数据集上, 参数设置为 $\delta=-50, \rho=4, \sigma=\infty, C=100$; 在 Coil20 数据集上, $\delta=-50, \rho=16, \sigma=\infty, C=100$ 。具体实验结果如表 4 和表 5 所列。

表 4 用 3 种光滑模型测试 Uspst 的分类错误率和 CPU 时间

	PWSS ³ VM($k=0.2$)	PSS ³ VM	ESS ³ VM
错误率%($C=100$)	21.6	22.2	21.2
CPU_TIME(s)	52.42	59.91	57.23

表 5 用 3 种光滑模型测试 Coil20 的分类错误率和 CPU 时间

	PWSS ³ VM($k=0.2$)	PSS ³ VM	ESS ³ VM
错误率%($C=100$)	5.82	5.91	6.0
CPU_TIME(s)	41.60	51.37	44.88

从上表可看出, 针对大规模的多分类数据, 采用分段函数逼近对称较链损失函数不仅能保证分类器的分类正确率, 而且执行时间较短。

结束语 综上所述, 本文从理论上证明了所构造的分段函数对对称较链损失函数的逼近存在唯一性; 分析了该函数的逼近精度。最后在 4 个数据集上进行了数值实验, 实验结果显示: 所提出的 PWSS³ VM 模型不仅对小规模的二分类数据集有较好的分类性能, 对大规模的多分类数据集也同样适用; 而且相对于其他光滑模型, 本模型具有更高的分类正确率和更快的处理速度。

参考文献

- [1] Burges C. A tutorial on support vector machines for pattern recognition[J]. Data Mining and Knowledge Discovery, 1998, 2(2):127-167
- [2] Lu S X, Tian R N. Structural Weighted Least Squares Support Vector Machine Classifier[J]. Computer Science, 2013, 40(12): 52-54(in Chinese)
鲁淑霞, 田如娜. 结构化加权最小二乘支持向量机[J]. 计算机科学, 2013, 40(12): 52-54
- [3] Sun L, Jing L, Xia X D. A new Proximal Support vector Machine for Semi-supervised Classification [M]. Heidelberg: Springer Berlin, 2006: 23-34

- [4] Bennett K P, Demiriz A. Semi-supervised support vector machines[M]. Cambridge: MIT Press, 1998; 10-20
- [5] Zhang X D, Ma J. A general cubic spline smooth semi-supervised support vector machine[J]. Chinese Journal of Engineering, 2015, 37(3): 386-389 (in Chinese)
张晓丹, 马菁. 一个广义三次样条光滑半监督支持向量机[J]. 工程科学学报, 2015, 37(3): 386-389
- [6] Chapelle O, Zien A. Semi-supervised classification by low density separation[C]//10th International Workshop on Artificial Intelligence and Statistics. Barbados, 2005; 57-68
- [7] Liu Y Q, Lin S Y. Polynomial Smooth Semi-supervised Support Vector Classifier[J]. Computer Science, 2009, 36(7): 179-181 (in Chinese)
刘叶青, 刘三阳. 一种多项式光滑的半监督支持向量机[J]. 计算机科学, 2009, 36(7): 179-181
- [8] Bennett K, Demiriz A. Semi-supervised support vector machines [C]//Proceedings of the 1998 Conference Advances in Neural Information Processing Systems. 1998; 368-374
- [9] Joachims T. Transductive inference for text classification using support vector machines[C]//International Conference on Machine Learning, Pittsburgh. 1999; 200-209
- [10] Dijkstra E W. A note on two problems in connection with graphs [J]. Numerische Mathematics, 1959, 1(1): 269-271
- [11] Cox T F, Cox M A. Multidimensional Scaling[J]. Journal of the Royal Statistical Society Series A (Statistics in Society), 1994, 45(2): 271-281
- [12] Bengio Y, Grandvalet Y. Semi-supervised learning by entropy minimization[J]. Neural Information Processing Systems, 2005, 5(2): 529-536
- [13] Nene S A, Nayar S K, Murase H. Columbia object image library (coil-20) [R]. CUCS-005-96, Columbia Univ., USA, February 1996

(上接第 253 页)

- [3] Liu X, Aberer K. SoCo: A social network aided context-aware recommender system[C] //Proc of the 22nd Int Conf on World Wide Web. Brazil, IW3C2, 2013; 781-802
- [4] Herlocker J, Konstan J, Borchers A, et al. An algorithmic framework for performing collaborative filtering [C] // Proc of the 22nd Annual Int ACM SIGIR Conf on Research and Development in Information Retrieval. New York; ACM, 1999; 230-237
- [5] Hu Y, Koren Y, Volinsky C. Collaborative filtering for implicit feedback datasets[C]//Proc of the 8th IEEE Int Conf on Data Mining. Pisa, Italy: ICDM, 2008; 263-272
- [6] Sarwar B, Karypis G, Konstan J, et al. Item-Based collaborative filtering recommendation algorithms[C]//Proc of the 10th Int Conf on World Wide Web. Hong Kong, China; ACM Press, 2001; 285-295
- [7] Yang Xing-yao, Yu Jiong, Ibrahim T, et al. Collaborative filtering model fusing singularity and diffusion process [J]. Journal of Software, 2013, 24(8): 1868-1884 (in Chinese)
杨兴耀, 于炯, 吐尔根·依布拉音, 等. 融合奇异性和扩散过程的协同过滤模型[J]. 软件学报, 2013, 24(8): 1868-1884
- [8] Adomavicius G, Tuzhilin A. Context-Aware recommender systems[M]. Recommender Systems Handbook. Berlin: Springer-Verlag, 2011; 217-253
- [9] Adomavicius G, Tuzhilin A. Context-Aware recommender systems[C]//Proc of the RecSys 2008. New York: ACM Press, 2008; 335-336
- [10] Adomavicius G, Ricci F. Workshop on context-aware recommender systems[C]//Proc of the RecSys 2009 (CARS 2009). New York; ACM Press, 2009; 423-424
- [11] Karatzoglou A, Amatriain X, Baltrunas L, et al. Multiverse recommendation: n-dimensional tensor factorization for context-aware collaborative filtering[C]//Proc of the fourth ACM Conference on Recommender systems. Barcelona, Spain; ACM, 2010; 74-86
- [12] Rendle S, Gantner Z, Freudenthaler C, et al. Fast context-aware recommendations with factorization machines[C]//Proc of the 34th Int ACM SIGIR Conference on Research and Development In Information Retrieval. Beijing, China; SIGIR, 2011; 635-644
- [13] Zhong E, Fan W, Wang L, et al. Comsoc: adaptive transfer of user behaviors over composite social network[C]//Proc of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2012; 696-704
- [14] Liu F, Lee H. Use of social network information to enhance collaborative filtering performance [J]. Expert Systems with Applications, 2010, 37(7): 4772-4778
- [15] Zhou K, Yang S, Zha H. Functional matrix factorizations for cold-start recommendation[C]//Proc of the 34th Int ACM SIGIR Conference on Research and Development In Information Retrieval. Beijing, China; SIGIR, 2011; 69-78
- [16] Zhang M, Tang J, Zhang X, et al. Addressing cold start in recommender systems: A semi-supervised co-training algorithm [C]//Proc of the 37th Int ACM SIGIR Conference on Research & Development in Information Retrieval. New York; ACM, 2014; 73-82
- [17] Jiang Sheng, Wang Zhong-qun, Xiu Yu, et al. Collaborative Filtering Recommendation Method Based on Dynamic Social Behavior and Users' Background Information [J]. Computer Science, 2015, 42(3): 252-255, 265 (in Chinese)
蒋胜, 王忠群, 修宇, 等. 基于动态社会行为和用户背景的协同推荐方法[J]. 计算机科学, 2015, 42(3): 252-255, 265
- [18] Wang Peng, Wang Jing-jing, Yu Neng-hai. A kernel and user-based collaborative filtering recommendation algorithm [J]. Journal of Computer Research and Development, 2013, 50(7): 1444-1451 (in Chinese)
王鹏, 王晶晶, 俞能海. 基于核方法的 User-Based 协同过滤推荐算法[J]. 计算机研究与发展, 2013, 50(7): 1444-1451
- [19] Hu Xun, Meng Xiang-wu, Zhang Yu-jie, et al. Recommendation algorithm combing item features and trust relationship of mobile users [J]. Journal of Software, 2014, 24(8): 1817-1830 (in Chinese)
胡勋, 孟祥武, 张玉洁, 等. 一种融合项目特征和移动用户信任关系的推荐算法[J]. 软件学报, 2014, 25(8): 1817-1830
- [20] Liu H, Setiono R. Chi2: Feature selection and discretization of numeric attributes[C]//ICTAI. 1995; 388-391
- [21] Yang X, Steck H, Liu Y. Circle-based recommendation in online social networks[C]//Proc of the 18th ACM SIGKDD Int Conference on Knowledge Discovery and Data Mining. 2012; 1267-1275