

基于 ART1 用户聚类的 Web 预取模型研究

姚 瑶,张 慧

(中州大学 信息工程学院,河南 郑州 450044)

摘 要:传统的 Web 预取机制所采用的预测算法主要针对用户个人预取,随着用户数量激增会加重网络负载,降低网络资源利用率。针对这一问题,在综合分析经典 ART1 神经网络模型和预取系统的基础上,提出一种基于用户聚类的 UCPM 模型预取新方法。首先,基于改进的 ART1 算法对用户访问序列特征向量进行聚类,挖掘兴趣相似的用户集合;然后,针对自底向上权重最大值所关联用户群的兴趣进行预取;最后,在 Web 预取系统上验证该方法的有效性及其可靠性。实验结果表明,UCPM 模型表现出较好的聚类效果,同时应用在预取系统保持了较高的预测准确率,降低了延迟比和流量开销比。

关键词:ART1 神经网络;用户聚类;Web 预取;延迟比;流量开销比

中图分类号:TP393

文献标识码:A

文章编号:1673-629X(2015)09-0106-05

doi:10.3969/j.issn.1673-629X.2015.09.023

Research on Web Prefetching Model Based on ART1 User Clustering

YAO Yao,ZHANG Hui

(School of Information Engineering,Zhongzhou University,Zhengzhou 450044,China)

Abstract:Prediction algorithm of traditional prefetching mechanism adopted by the invisible network load will increase the utilization rate,reduce the cyber source rate. In view of this,through the comprehensive analysis of the classical ART1 neural network model and the prefetching system,a new method of UCPM prefetching model is proposed based on user clustering. Firstly,based on improved traditional ART1 algorithm for clustering user access sequence feature vector,mine the similar interests of a user set,and then prefetching technique predicts the destination pages for user set on the basis of bottom-up maximum weight value. Finally,the experiments on Web prefetching system verify the effectiveness and reliability of the proposed algorithm. Experimental results show that the model exhibits better clustering effect,at the same time,the prediction accuracy in prefetching system has been retained highly,lowering the ratio of latency and traffic.

Key words:ART1 NN;user clustering;Web prefetching;ratio of latency;ratio of traffic

0 引 言

随着网络的飞速发展,如何提高网络服务质量、降低访问延迟始终是人们研究的热点。诱发网络延迟的主要原因包括 Web 服务器负载过重、网络拥塞、带宽限制和传播延迟等。预取机制在用户请求网页之前提前取出页面从而减少用户期望延迟,是一项有效的 Web 加速方法之一^[1]。

典型的预取方法有基于概率的预取(Markov 树、双关联图 DDG、多维矩阵)、基于聚类的预取(SVM、基于图的聚类^[2]、PCM、基于文本聚类^[3])和基于加权函数的预取(PPM、Web 日志挖掘)。目前对预取机制的研究主要集中在针对个人用户的预测算法改进,但随着用户数量的激增,个人用户预取机制容易导致负

荷过重,缺少对用户群进行预测的有效方法。

基于聚类的预取方法中,文献[4]提出一种基于两阶段的用户聚类算法 GCCR,主要研究了基于图摘要和内容相似的混合聚类,该方法侧重内容特征的相似度来对用户聚类,算法开销较大。文献[5]提出 clustWeb 机制聚类跨站点的网页,采用基于关联规则和页面间的连通性分割网络导航图从而获得聚类。该方法虽然保证较高的字节命中率,但是预取出聚类中的所有请求对象,影响了预取效率。文献[6]提出 PCM 模型,采用一阶 Markov 模型模拟用户的请求序列。为了节约内存,提出预分配方案,运行时,可用内存的总大小可以适当改变。该预取方法适用于较低内存的设备,例如 PDA 或者手机等。文献[7]提出一种

收稿日期:2014-11-06

修回日期:2015-02-06

网络出版时间:2015-08-26

基金项目:国家自然科学基金资助项目(U1304603);2012 年郑州市科技计划项目(121PPTGG364)

作者简介:姚 瑶(1982-),女,讲师,硕士,研究方向为网络信息技术和 Web 挖掘。

网络出版地址:http://www.cnki.net/kcms/detail/61.1450.TP.20150826.1603.064.html

基于向量空间模型的随机索引方法,主要用来挖掘 Web 用户行为的潜在特征,并没有得到较高的聚类效果,仍然局限在对用户个人的行为挖掘上。

以上研究对降低访问延迟起到一定作用,但没有考虑到用户聚类算法和预取机制的协同问题,从而导致相应研究结果对于预取系统而言不甚理想。为解决上述问题,文中首先采用改进的 ART1 聚类算法对用户进行聚类;其次,提出基于用户聚类的 UCMP(User Clustering Prefetching Model)策略,数据模型向量为每一个用户聚类推荐最大众化的访问请求代表该聚类所有成员的广义需求;最后,将该模型应用于预取系统,验证方法的有效性。

1 问题描述与分析

衡量 Web 预取模型性能的关键指标之一是延迟时间,亦称系统的响应时间,即从发出请求开始到请求得到响应为止的时间间隔。影响延迟时间的因素有很多,诸如 Web 流量(在线时间、离线时间和交互时间)、页面大小、预取对象的数量、用户数量和访问链接的带宽^[8]等。影响因子的详细指标归纳如表 1 所示。其中 Web 流量可用开关(ON-OFF)过程模拟。ON 状态用户可以请求和下载对象,OFF 状态即交互时刻,用户可浏览网页。

表 1 响应时间影响因子一览表

影响因子	名称	服从分布	形状参数	尺度参数
ON Time	在线时间	韦伯分布	0.77	$e^{0.44}$
Inter-arrival Time	交互时间	韦伯分布	0.5	1.5
OFF Time	离线时间	二八分布	0.58	60
Page Size	页面大小	二八分布	0.3	300
Prefetched objects	预取对象数量	韦伯分布	0.5	0.5
non-Prefetched objects	无预取对象的数量	韦伯分布	0.9	4
m	用户数量	-	-	-
B	带宽	-	-	-

如表 1 所示,参数服从韦伯分布的概率密度函数可描述为:

$$f(x) = \frac{\beta}{\alpha} \left(\frac{x}{\alpha}\right)^{\beta-1} \exp\left[-\left(\frac{x}{\alpha}\right)^\beta\right] \quad (x > 0) \tag{1}$$

其中,β 是形状参数;α 是尺度参数。

参数服从帕累托分布(二八分布)的概率密度函数可描述为:

$$f(x) = \delta k^{\delta-1} x^{-\delta-1} \tag{2}$$

其中,δ 是形状参数;k 是尺度参数。

Web 预取模型虽然在一定程度上能改善延迟时间,但随着大数据时代的到来,网络用户和资源对象的激增,导致对延迟时间的改善能力逐渐下降。对预取

模型的改进迫在眉睫。

2 基于用户聚类的 Web 预取模型

基于用户聚类的 Web 预取模型(UCMP)增加聚类模块,对初始数据进行行为兴趣分析,聚类模块为客户端用户进行分组并返回每一组的原型向量,为后续的预取工作节省开销。

2.1 改进的 ART1 的用户聚类方法

聚类模块采取的聚类方法源于自适应共振理论(Adaptive Resonance Theory, ART)。该神经网络模型主要应用于数据挖掘的聚类领域,它在学习过程中能够实现稳定性与可塑性的统一^[9]。其基本模型如图 1 所示。 F_1 为输入层, F_2 为聚类层, $b_{i,j}$ 和 $t_{j,i}$ 分别代表 bottom-top 和 top-down 权重,依据计算机输入二进制向量的相似度以及复位机制进行调整迭代,警戒参数 ρ 用以控制聚类过程中的序列相似程度, G_1, G_2 为增益控制单元。

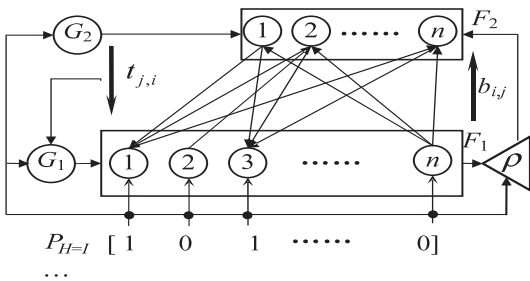


图 1 ART1 系统结构图

为了将 ART1 应用于预取系统,首要工作需从第 h 个主机节点的访问序列提取输入二进制向量。向量 $S = \{URL_1, URL_2, \dots, URL_n\}$ 代表主机的访问序列,对每一个主机 H, P_H 代表其二进制序列向量, P_H 为每一个 $URL_i (i = 1, 2, \dots, n)$ 的访问频率到二进制值的映射。 $P_H = \{p_1, p_2, \dots, p_n\}, 1 \leq H \leq h$ 。每一个 $i \in \{1, 2, \dots, n\}$, 如果主机 H 请求了 URL_i 两次或两次以上,那么 p_i 取值 1, 否则取值 0。尽管如此,每一个主机发出的请求序列不尽相同,访问次数阈值两次并不能总是作为未来预测的最佳值。

文中对访问次数的临界点进行改进,所选择的阈值需要兼顾主机和 URL。设 r_{H_i} 为主机访问 URL_i 的数量, $H \in \{1, 2, \dots, h\}$, p_i 取值更新为:

$$p_i = \begin{cases} 1 & r_{H_i} > \frac{1}{n} \sum_{j=1}^n r_{H_j} \\ 0 & \text{其他} \end{cases} \tag{3}$$

一般地,二进制向量 X 和 S 的相似度为其矩阵范数的比值,即 $\|X\| / \|S\|$, 比值越大相似度越高。在 Web 预取系统,该比值并不能反映出用户会话的相似性^[10-11]。为解决该问题,文中提出向量 X 和向量 $S = (s_1, s_2, \dots, s_n)$ 之间的距离:

定义 1: $\mathbf{X} - \mathbf{S} = (d_1, d_2, \dots, d_n)$,

$$\text{其中 } d_i = \begin{cases} 0 & x_i = s_i \\ 1 & x_i \neq s_i \end{cases}.$$

定义 2: 向量的相似度

$$\text{Sim}_{ij} = \|\mathbf{X} - \mathbf{S}\| / n \quad (4)$$

当 Sim_{ij} 在警戒参数的 $(0, 1)$ 范围以内, $b_{i,j}$ 和 $t_{j,i}$ 权重会根据聚类单元进行调整。当相似度超过警戒参数时, 重新选择候选聚类单元。当前聚类单元被抑制, 所以在当前学习测试阶段不能再次选择该候选单元, 并且 F_1 单元的激励因子重置为 0。相同的输入向量再次将其信号输入给接口单元, 接口单元将其 $b_{i,j}$ 信号传给 F_2 层, 这种竞争机制迭代进行。

算法 1: 改进的 ART1 用户聚类算法。

Input: 代表用户访问序列的特征向量 \mathbf{P} , 警戒参数 $\rho \in (0.3, 0.5)$;

Output: 用户聚类。

Step1: 初始化, F_1 和 F_2 层节点设置为 0; top-down

($t_{j,i}$) = 1; bottom-up($b_{i,j}$) = $\frac{1}{n+1}$; 设置 ρ ;

Step2: 重复 Step3 至 Step10 直至输入向量均在 F_1 层出现;

Step3: F_1 层随机选择输入向量;

Step4: 计算 F_2 层的每一个节点的 y_i 值: $y_i = \sum_{i=1}^{200} P_i \times b_{ij}$;

Step5: F_2 层具有最大 y_k 值的节点决定 k 值: $y_k =$

$$\sum_{j=1}^m \max(y_j), m \text{ 为 } F_2 \text{ 层节点数量};$$

Step6: F_1 层节点 k 计算激励函数 $\mathbf{X}_k^* = (X_1^*, X_2^*, \dots, X_{200}^*)$, 其中 $\mathbf{X}_i^* = t_{ki} \times P_i$;

Step7: 计算 \mathbf{X}_k^* 和输入向量 \mathbf{P}_H 的相似度: $\frac{\|\mathbf{X}_k^*\|}{\|\mathbf{P}_H\|}$

$$= \frac{\sum_{i=1}^{200} X_k^*}{\sum_{i=1}^H P_i};$$

Step8: if ($\frac{\|\mathbf{X}_k^*\|}{\|\mathbf{P}_H\|} > \rho$)

{ 节点 k 的激励函数设置为 0;

更新 k 的 $t_{ki}(\text{new}) = t_{ki} \times P_i$, 其中 $i = 1, 2, \dots, 200$;

else 转至 Step9;

Step9: 为 F_2 层创建一个新节点 q

{ 为当前输入序列初始化权重 t_{li} ;

为节点 q 初始化权重 $b_{il}(\text{new}) = \frac{X_i^*}{0.5 + \sum_{i=1}^{200} X_i^*}$, 其

中 $i = 1, 2, \dots, 200$;

Step10: 转至 Step2。

2.2 构建 UCPM 预取模型

每一个客户端发出的请求代表一组用户的请求, 对客户端的聚类即实现了对用户的聚类。每一个聚类群的原型向量指出了该聚类所有成员访问最频繁的 URL 的广义代表。该策略应用在文中提出的预取方案中, 当一个主机连接到服务器或代理服务器时, 原型向量中与聚类相关联的代表客户端的 URL 被预取。文中提取的预取方案有两个优点: 一是由于是针对用户群而不是某一个用户进行预取, 确保了网络资源的合理有效利用; 二是预取准确率较高, 能达到 90% 以上。预取机制中对用户的聚类采用改进的 ATR1 聚类算法, 每一个 URL 和聚类依据 $b_{i,j}$ 和 $t_{j,i}$ 的权重合并在一起。 $b_{i,j}$ 权重的最大值的 URL 关联的聚类将会被预取。选取聚类后, 算法将根据 $t_{j,i}$ 权重决策需要预取的页面。算法描述如下:

算法 2: 生成预取 URL。

// C_1, C_2, \dots, C_n 是由原型向量生成的聚类; 第 k 个聚类的原型向量是 $T_k = (t_{k1}, t_{k2}, \dots, t_{km})$, 其中 $t_{kj} = 1, 2, \dots, 200$ 代表节点 k 在 ART1 第 F_2 层的 top-down 权重。

Input: 主机 ID 号

// 使用 ART 神经网络聚类算法聚类提出访问请求的主机

Output: Prefetched_URLs[]

Step1: 初始化, Count = 0;

Step2: for ART1_Clustering 算法形成的 n 个聚类

Step3: if (Host_ID $\in C_k$)

Step4: { Repeat for $j = 1, 2, \dots, m$

if ($t_{kj} = 1$)

{ // t_{kj} 是 T_k 的第 j 个元素

Prefetched_URLs[Count] = URL_j;

Count++;

Return Prefetched_URLs[]; } //End

3 仿真实验与结果分析

3.1 数据集

为了检验 ART1 聚类算法和传统 K -means 聚类算法的有效性, 文中分别在实际数据集上进行了实验, 做了比较分析, 证明文中提出方法的合理性和适用性。算法采用 C++ 编程实现, 客户端环境为 PC 机, 2 GHz 内存, Windows XP 操作系统; 代理服务器环境为 PC 机, 4 GHz 内存, Windows Server 2003 操作系统。实验运用 Berkley 大学计算机系日志 UCB 和河南省信息网络重点实验室日志 HENU, 选用日志文件的 1/5 作为

训练集,4/5 作为测试集。

3.2 聚类性能

首先测试文中改进的聚类算法在不同输入警戒参数情况下的平均类间距离和平均类内距离。其次,聚类效果评估指标 Davies–Bouldin(DB)用来衡量同类样本的紧密程度和不同类样本的分散程度(如公式(5))^[12]。类内距离小,类间距离大,DB 指标的值就越小,聚类效果越好。

$$V_{DB} = \frac{1}{C} \sum_{k=1}^c \text{Max}_{t,t \neq k} \left\{ \frac{S(C_k) + S(C_t)}{D_{k,t}(C_{k,t})} \right\} \quad (5)$$

设置警戒参数范围在[0.3,0.5],观察其聚类的数量及平均类间/内距离的变化,如表 2 和表 3 所示。

表 2 文中算法的聚类指标

警戒参数	聚类数量	平均类间距离	平均类内距离	平均 DB 值
0.30	20	54.331 2	19.063 8	0.325 0
0.325	22	55.213	19.424 3	0.358 4
0.35	24	56.318 6	19.893 7	0.335 9
0.375	28	54.396 2	17.999 8	0.215 9
0.4	34	57.218 4	19.221 4	0.175 1
0.425	35	53.246 5	19.867 8	0.153 1
0.45	39	54.139 7	20.123 5	0.138 0
0.475	46	52.414 2	20.652 4	0.110 3
0.5	55	53.979 9	19.596 4	0.100 2

表 3 K – means 算法的聚类指标

聚类数量	平均类间距离	平均类内距离	平均 DB 值
20	65.369 0	23.876 6	0.492 8
22	67.011 0	23.309 2	0.470 4
24	67.950 5	20.872 4	0.488 3
28	68.899 1	21.599 8	0.410 1
34	69.986 6	19.065 7	0.388 7
35	67.986 5	18.841 4	0.371 1
39	68.723 4	17.148 2	0.369 9
46	67.235 4	15.782 9	0.333 4
55	65.199 7	12.515 7	0.341 0

观察表 2,随着警戒参数的递增,文中算法聚类的数量逐渐增多,平均类间距和类内距分别在 50 和 20 左右浮动。对比其他算法,平均类间距离逐步稳定说明算法在类间距上表现并没有太大差别。对比表 3,在平均类内距离方面,K – means 算法表现在 12.515 7 至 23.876 6 之间,文中算法在 17.999 8 至 20.652 4 之间。对于 DB 指数,文中算法比 K – means 算法低 0.2 个单位。与 K – means 算法相比,随着聚类数量逐步增多,文中算法的聚类差异更趋于稳定。

3.3 预取效果

通过分析预测准确率(Accuracy)^[13]、延迟比(Ra-

tio_Latency)和流量开销比(Ratio_Traffic)衡量预取模型的有效性^[14],预测准确率越高、延迟比和流量开销比越小,说明系统性能越好。

实验中设置算法警戒参数为 0.38,统计 4 个聚类的预取情况指标,如表 4 所示。

表 4 预取情况统计

聚类	用户团体	用户	请求页面数量	预取数量	命中数量	准确率/%
C ₁	U ₁ ,	U ₁	178		43	95.56
	U ₂ ,	U ₂	202		42	93.33
	U ₅ ,	U ₅	145		41	91.11
	U ₉ ,	U ₉	185	46	42	93.33
	U ₁₁ ,	U ₁₂	209		43	95.56
	U ₁₄ ,	U ₁₄	209		43	95.56
	U ₄₁	U ₄₁	149		42	93.33
C ₂	U ₃ ,	U ₃	220		64	94.12
	U ₄ ,	U ₄	149		60	88.24
	U ₈ ,	U ₈	161	68	61	89.71
	U ₁₂ ,	U ₁₂	222		63	92.65
	U ₁₇	U ₁₇	169		62	91.18
C ₃	U ₆ ,	U ₆	205		33	91.67
	U ₁₀ ,	U ₁₀	–		–	–
	U ₁₅ ,	U ₁₅	162	31	32	88.89
	U ₁₆	U ₁₆	94		31	86.11
C ₄	U ₅ ,	U ₅	–		–	–
	U ₇ ,	U ₇	139	26	22	91.67
	U ₁₃	U ₁₃	178		21	95.56

如表 4 所示,基于 UCPM 的预测准确率达到 86% 以上。当对一个用户群聚类后,也存在用户没有访问已经预取过的对象的现象。例如 C₃ 聚类中的 U₅ 和 C₄ 聚类中的 U₁₀。尽管如此,文中预取机制的平均预测准确率仍能达到较高的预测准确率 91.73%。

平均延迟比/流量开销比如图 2 所示。观察可知,带宽越高,平均延迟比/流量开销比越低,表现出较好的性能。随着流量开销增大,延迟呈递增趋势,尤其是在带宽较小时,递增格外明显。

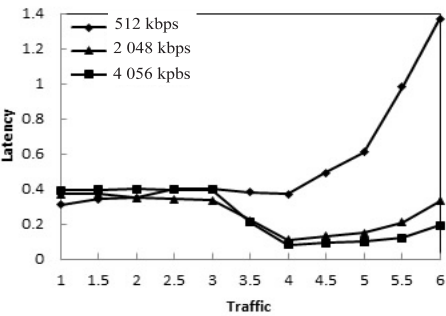


图 2 平均延迟比/流量开销比趋势图

4 结束语

文中提出改进的 ART1 算法对用户群进行聚类并

应用到 Web 页面预取中。与其他方法相比,该方法的优点在于,首先从挖掘用户群的访问兴趣入手,为用户群预取而不是单一用户;其次,根据主机发出请求序列的随意性,改进了 ART1 聚类算法,与传统 K -means 算法相比,提高了聚类的准确度和适用度;最后,应用改进的 UCPM 模型在真实日志上模拟实验,预取准确率最高可达到 95.56%。研究表明,采用基于用户访问特性的 ART1 聚类算法能够对用户有效聚类,并且对聚类后的用户群预取时能够在保证较高准确率的前提下降低访问延迟,充分提高资源利用率。

参考文献:

- [1] Zong Ziliang, Fares R, Romoser B, et al. FastStor: improving the performance of a large scale hybrid storage system via caching and prefetching[J]. Cluster Computing, 2014, 17(2): 593–604.
 - [2] 尹文科,朱明,陈天昊. 基于 Wiki 链接结构图聚类的领域词典构建方法[J]. 小型微型计算机系统, 2014, 35(6): 1286–1292.
 - [3] 张万山,肖瑶,梁俊杰,等. 基于主题的 Web 文本聚类方法[J]. 计算机应用, 2014, 34(11): 3144–3146.
 - [4] 陈克寒,韩盼盼,吴健. 基于用户聚类的异构社交网络推荐算法[J]. 计算机学报, 2013, 36(2): 349–359.
 - [5] Dr A K, Jayasudha S S. An efficient cluster based web object filters from web pre-fetching and web caching on web user navigation[J]. International Journal of Computer Science Issues, 2012, 9(3): 483–489.
 - [6] Liu Qinghui, Solis – Oba R. Web prefetching with machine learning algorithms[C]//Proc of international conference on internet computing. [s. l.]: [s. n.], 2008: 142–148.
 - [7] Wan Miao, Jönsson A, Wang Cong, et al. Web user clustering and Web prefetching using random indexing with weight functions[J]. Knowledge and Information Systems, 2012, 33(1): 89–115.
 - [8] de la Ossa B A, Sahuquillo J, Pont A, et al. Key factors in web latency savings in an experimental prefetching system[J]. Journal of Intelligent Information Systems, 2012, 39(1): 187–207.
 - [9] 黄清兰. 基于聚类划分的 Web 日志关联规则增量式挖掘方法研究[D]. 南昌:南昌大学, 2013.
 - [10] 邓智龙,张海粟,黄立威. 一种基于社区结构的用户兴趣关联规则发现方法[J]. 计算机应用研究, 2012, 29(5): 1799–1801.
 - [11] 袁书寒,陈维斌,傅顺开. 位置服务社交网络用户行为相似性分析[J]. 计算机应用, 2012, 32(2): 322–325.
 - [12] 郑富兰,吴瑞. 基于用户特性的 Web 会话模式聚类算法[J]. 计算机应用与软件, 2014, 31(2): 283–286.
 - [13] 姚瑶,王战红,石磊. 基于页面聚类的 Web 概念化模型研究[J]. 科学技术与工程, 2014, 14(25): 272–276.
 - [14] Ban Zhijie, Wang Sansan. A framework of online proxy-based web prefetching[J]. Web Information Systems and Mining Lecture Notes in Computer Science, 2012, 7529: 610–620.
- +++++
- (上接第 105 页)
- 化方法及其应用[J]. 计算力学学报, 2010(3): 451–456.
 - [3] 薛丹,杨宸,周健. 一种基于区间值的模糊访问控制策略研究[J]. 计算机技术与发展, 2012, 22(1): 246–249.
 - [4] Jiang Chao, Han Xu, Liu Guirong, et al. A nonlinear interval number programming method for uncertain optimization problems[J]. European Journal of Operational Research, 2008, 188(1): 1–13.
 - [5] Jiang Chao, Han Xu, Li Ding. A new interval comparison relation and application in interval number programming for uncertain problems[J]. Computers Materials and Continua, 2012, 27(3): 275–303.
 - [6] 蒋峥,刘斌. 自适应主从式并行遗传算法在区间非线性规划问题求解中的应用[J]. 信息与控制, 2006, 35(3): 314–318.
 - [7] Kamakar S, Bhunia A K. An alternative optimization technique for interval objective constrained optimization problems via multiobjective programming[J]. Journal of the Egyptian Mathematical Society, 2013, 22(2): 292–303.
 - [8] 张著洪,陶娟. 求解非线性区间数规划的微免疫优化算法研究[J]. 计算机研究与发展, 2014, 51(12): 2633–2643.
 - [9] Moore R E, Cloud M J, Kearfott R B. Introduction to interval analysis[M]. Philadelphia: Society for Industrial and Applied Mathematics, 2009.
 - [10] Ishibuchi H, Tanaka H. Multiobjective programming in optimization of the interval objective function[J]. European Journal of Operational Research, 1990, 48(2): 219–225.
 - [11] 陈娟,徐立鸿. 动态小生境遗传算法在多模函数优化中的应用[J]. 同济大学学报:自然科学版, 2006, 34(5): 684–688.
 - [12] 彭利兵,黄辉先,阮挺,等. 多峰函数优化的自适应小生境克隆选择算法[J]. 计算机工程与应用, 2011, 47(9): 48–50.
 - [13] Kwon Y D, Kwon S B, Jin S B, et al. Convergence enhanced genetic algorithm with successive zooming method for solving continuous optimization problems[J]. Computers & Structures, 2003, 81(17): 1715–1725.
 - [14] 靳宗信,郑良仁,樊红娟. 一种改进的用于多峰值函数优化的自适应克隆选择算法[J]. 现代计算机, 2010(3): 14–17.
 - [15] Jiang C, Han X, Liu G P. A sequential nonlinear interval number programming method for uncertain structures[J]. Computer Methods in Applied Mechanics and Engineering, 2008, 197(49): 4250–4265.

基于ART1用户聚类的Web预取模型研究

作者：[姚瑶](#)，[张慧](#)，[YAO Yao](#)，[ZHANG Hui](#)
作者单位：[中州大学 信息工程学院, 河南 郑州, 450044](#)
刊名：[计算机技术与发展](#)[ISTIC](#)
英文刊名：[Computer Technology and Development](#)
年，卷(期)：2015 (9)

引用本文格式：[姚瑶](#). [张慧](#). [YAO Yao](#). [ZHANG Hui](#) [基于ART1用户聚类的Web预取模型研究](#)[期刊论文]-[计算机技术与发](#)
[展](#) 2015 (9)