

基于多 Agent 的用户上下文自适应站点构架

郑伟平^{1,2}, 齐德昱¹, 向 军¹, 徐克付¹

(1. 华南理工大学计算机科学与工程学院, 广州 510641; 2. 华南师范大学计算机学院, 广州 510631)

摘要: 自适应站点很少考虑对用户环境的自适应。为此, 提出用户上下文自适应站点的概念, 给出基于多 Agent 技术的用户上下文自适应站点构架模型。阐述用户上下文获取、挖掘过程以及站点改造过程中各 Agent 的作用和交互情况。给出根据用户网络带宽和终端屏幕尺寸来调整推荐和显示的策略, 并依据该构架实现了一个网上家具商场的原型站点。

关键词: 自适应站点; 用户上下文; 多 Agent

User-context Adaptive Website Architecture Based on Multi-Agent

ZHENG Wei-ping^{1,2}, QI De-yu¹, XIANG Jun¹, XU Ke-fu¹

(1. School of Computer Science & Engineering, South China University of Technology, Guangzhou 510641;

2. School of Computer, South China Normal University, Guangzhou 510631)

【Abstract】 Current adaptive Websites seldom consider adaptability towards user environment. For this reason, a new concept is presented about user-context adaptive Website and its architecture based on multi-Agent technology. Related Agents and interactions in user-context collecting, mining, and Website transformations are fully discussed. Policies for recommendation and presentation transformations are given according to network bandwidth and terminal screen size from client environment. A prototype Web store for furniture is setup using this architecture.

【Key words】 adaptive Website; user-context; multi-Agent

1 自适应站点的新定义

Web 数据的快速增长引起人们对 Web 访问时间费用的关注。Web 站点纷纷推出定制和个性化服务, 自适应站点 (adaptive Website) 技术应运而生。自适应站点研究用户需求、行为模式等来安排站点的组织与显示, 为不同用户提供不同视图, 使用户尽快访问到所需的信息。目前已有很多自适应站点的相关研究。现有技术大部分通过挖掘 Web 服务器日志 (Weblog) 来发现用户的访问模式, 进而改造站点。

基于 Weblog 的挖掘技术主要有 2 类: 基于事务和基于数据立方体。基于事务的方法^[1]先对 Web 日志文件进行整理, 去除无关信息后进行用户识别和会话识别, 进而识别出频繁序列模式, 最后将分析出来的频繁序列模式应用到站点改造上。文献^[1]提出了最大向前引用 MF 算法, 把 Weblog 转换成最大向前引用集。基于数据立方体的方法^[2]从 Weblog 中建立数据立方体, 利用数据挖掘技术对数据立方体进行挖掘和联机分析处理。

基于 Weblog 的自适应站点仅仅考虑用户“个人因素”, 如访问模式、习惯等, 没有考虑用户的“环境因素”, 如所用终端、网络带宽等。笔者认为真正自适应的站点不单考虑用户个人特性, 还应该考虑环境因素。理由为: 从时间费用来看, 低带宽环境下应该减小数据传输量; 从服务质量来看, 高带宽用户应该得到更高质量的服务; 从组织和显示来看, 在桌面 PC 与移动手机上, 站点的组织和显示应当有所区别。

为此, 本文把用户上下文自适应站点定义为: “根据用户上下文的不同调整站点的组织与显示, 为不同用户上下文提供不同视图的站点 (普适计算中终端信息和带宽等属于计算上下文的范畴)。”用户上下文是指用户本身与用户环境的属性, 包括用户行为模式、偏好、所用终端设备、网络带宽等。

2 Agent 在自适应站点上的应用

Agent 是代表用户完成某种任务, 并具有自治性、交互性、主动性等特性的实体。Agent 技术已在智能机器人等多个领域取得了巨大成功。多 Agent 是有序的 Agent 联合, 各成员之间相互通信、协同、服务和管理的, 共同完成一个复杂任务。由于适合处理数据、知识、控制分散的分布式问题, 多 Agent 正成为一种新的商务智能化技术。

已经有很多文献介绍 Agent 在自适应站点上的应用。文献^[3]提出了一个可重用多 Agent 构架, 由客户、私人助理 Agent、站点 Agent 与雇员组成, 采用通用信息代理 (generic information broker) 的 Agent 模型。并实现了一个保险业务的智能站点。文献^[4]把 Agent 应用到自适应网上商店上, 对商品和顾客进行建模并提出了 SETA 构架, 描述了网上商店的常用功能。SETA 根据顾客对商品属性的重视程度和兴趣为商品打分, 据此动态生成目录页, 提供个性化服务。文献^[5]在网上个性化信息检索中引入 Agent, 提出分层体系结构。

3 用户上下文自适应站点的多 Agent 设计

3.1 设计依据与构架模型

本文采用多 Agent 技术来构建用户上下文自适应站点, 原因如下:

(1) 站点中各构件的角色、功能不同, 所用的技术、知识表示也不同, 多 Agent 系统则提供了统一的协作框架。

基金项目: 粤港关键领域基金资助重点突破项目 (2005A10307007)

作者简介: 郑伟平 (1979 -), 男, 工程师、博士研究生, 主研方向: 软件体系结构, Web 服务; 齐德昱, 教授、博士生导师; 向 军、徐克付, 博士研究生

收稿日期: 2008-03-29 **E-mail:** newmood@tom.com

(2)Agent 能够实现更主动和智能的业务功能。

(3)基于 Agent 的系统提供多种通信机制,很好地支持了站点的并行处理。

(4)利用 Agent 的通信标准,新构件能够无缝地加入系统,站点的扩展能力更好。

(5)支持分布式计算,基于Agent的站点能够方便地把构件分布到不同的机器上^[4]。

作为抽象描述,本文提出用户上下文自适应站点的构架模型,把自适应站点功能分成4个功能模块:访问层,资源层,推理层与改造层,如图1所示。

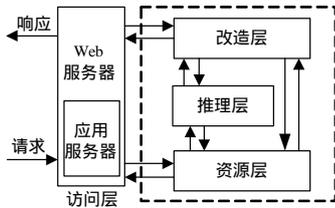


图1 用户上下文自适应站点构架

各功能模块构成如下:

(1)访问层:包括Web服务器和应用服务器,实质上相当于不具备自适应功能的传统站点。

(2)资源层:包括数据管理Agent(下称DMA)、Web信息库、用户上下文库、站点元数据库与推荐模板库。

(3)推理层:是站点自适应智能的实现部分,相当于Agent的推理机。包括用户上下文挖掘Agent(下称UCMA)、推荐策略Agent(下称RPA)、显示策略Agent(下称PPA)和组织策略Agent(下称OPA)。

(4)改造层:实现对站点内容的改造功能,为用户提供个性化推荐,调整页面显示和站点组织等。包括改造控制Agent(下称TCA)、推荐改造Agent(下称RTA)、显示改造Agent(下称PTA)与组织改造Agent(下称OTA)。

资源层是数据存储与管理的模块,挖掘的原始数据和挖掘后的知识都存放在该层。推理层从资源层获取原始数据,数据分析后发现用户上下文,为改造过程提供知识支持。推理层还提供策略服务,为改造层提供决策支持。改造层在资源层和推理层的支持下对输出页面进行适应性调整,调整的数据和动作规划来自资源层,具体格式由用户上下文与推理层的策略决定。在图1中,虚线表示三者共同为站点提供自适应能力。从效果上来看,访问层相当于感知器,推理层相当于推理机,资源层相当于知识库与内部状态,而改造层相当于效应器,整个站点对外表现为一个站点Agent。

3.2 用户上下文获取与挖掘

用户上下文信息从服务器端和客户端两处获取。服务端信息主要来自Weblog文件。在本文中客户端只考虑用户终端和网络带宽信息。为此,在客户端引入了客户环境Agent、带宽监视Agent与终端监视Agent,如图2所示。

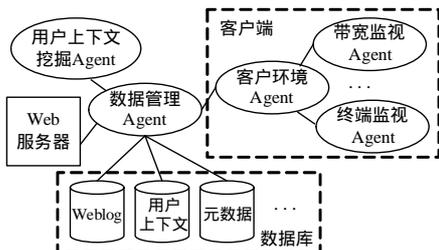


图2 用户上下文获取与挖掘相关的Agent

带宽监视Agent维护客户网络的连接状况、带宽情况;终端监视Agent记录终端类型、处理器使用状况、内存使用状况以及显示器类型。客户环境Agent收集带宽监视Agent与终端监视Agent的数据,融合后传输给服务器端的DMA。客户环境Agent协调监视Agent间的关系,静态数据(如处理机类型)在会话建立时传输,动态数据(如带宽)根据策略决定传输的时机与频率。

服务器上所有与自适应功能相关的数据由DMA统一管理。DMA收到客户环境Agent传来的数据后更新用户上下文数据库。Web服务器的Weblog文件也由DMA负责维护更新。

UCMA读取Weblog数据,应用挖掘算法进行分析,最后把分析结果存到用户上下文数据库中。UCMA主要基于Weblog文件挖掘用户访问模式,在本构架下可以方便地改变挖掘的智能算法。带宽和终端信息对自适应推理的影响主要由推理层的策略Agent体现,图2表示用户上下文信息收集和挖掘过程中各层Agent之间的关联。

3.3 改造站点以适应用户上下文

自适应站点根据用户上下文信息对站点原有内容进行调整的过程称为站点改造(Website transformation)。站点改造视具体应用和站点而定,一般包括页面推荐、链接结构改进、页面聚类、显示突出、排序调整等。本构架中站点改造由TCA、RTA、PTA与OTA配合完成。

如图3所示,TCA是改造过程的协调者,实际的改造工作由RTA、PTA和OTA完成。当需要推荐页面时,TCA与DMA交互,提取与用户上下文相关的推荐信息,与RPA交互获取与当前用户上下文对应的推荐策略,然后将推荐信息和策略一起交给RTA,委托其完成推荐改造。在线推荐时,RTA还必须维护其负责推荐的用户的当前会话活动窗口。同样,TCA分别从DMA和PPA处取得需要突出显示的信息及其显示策略,由PTA完成显示改造。应当注意,OTA一般只给出调整建议,如根据频繁路径增加超链接,或者根据聚类产生新的页面,这些建议常触及站点原有逻辑结构,必须由管理人员决定调整与否。

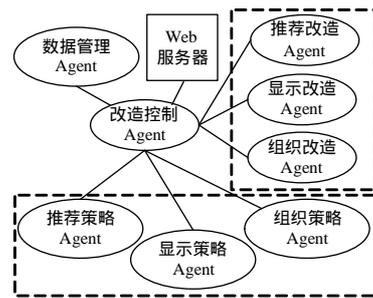


图3 改造过程相关的Agent

TCA的另一职责是根据用户和其会话进行推理,给出改造规划,即决定调用哪些改造Agent。例如,某些页面只突出显示,而有些则除突出显示之外还要推荐页面。TCA还必须协调改造Agent之间的执行次序。最后TCA把改造结果交给Web服务器。

3.4 推荐策略和显示策略

自适应站点关于用户访问模式的相关研究已很多,而组织调整改造主要由访问模式决定,故本文暂不讨论组织策略。

目前用户终端越来越多样化,终端之间的屏幕尺寸和内存大小差别很大。推荐页面时,若终端屏幕很小,就只推荐少数重要条目。内存是要考虑的另一因素。笔者在RPA中引

入了模糊控制模型。模糊控制规则表示推荐策略如下：

If *SrnSz* is large and *TMem* is notSmall then *rmNum* is many
If *SrnSz* is large and *TMem* is small then *rmNum* is medium
If *SrnSz* is medium and *TMem* is notSmall then *rmNum* is medium
If *SrnSz* is medium and *TMem* is small then *rmNum* is few
If *SrnSz* is small then *rmNum* is few

rmNum, *TMem*, *SrnSz* 都是语言变量，对应推荐数量、内存大小和终端屏幕尺寸。*SrnSz* 有 Large, medium, small 3 个语言值，表示大、中与小屏幕。*rmNum* 有 many, medium, few 3 个值，表示推荐很多、一般和很少条。*TMem* 有 small 和 notSmall 2 个值，表示终端内存很小和不小。本文只用模糊规则表示推荐策略，具体语言值的隶属函数视实际而定。

在显示改造策略方面，除了改变内容的字体、样式等以突出关键内容之外，还考虑带宽的影响。基于区分服务的原理，对高带宽用户提供高质量的多媒体数据，而降低低带宽用户的质量，如降低分辨率、提高压缩比等。根据带宽大小和终端处理器的繁忙程度，对图像、声音、视频数据的输出分成高、中、低质量副本以及文本代替(或者不提供)4 种情况。多媒体质量的显示策略也由模糊规则表示。

3.5 多 Agent 的通信和协同

多 Agent 系统中 Agent 的通信和协调很重要。本构架的多 Agent 通信主要分成直接通信和间接通信 2 种方式。如 UCMA 和 DMA 之间就是通过点对点直接通信来交换数据的，而 RTA 要获取推荐策略时要通过 TCA 间接与 RPA 通信。目前主要的 Agent 通信协议有 KQML 和 FIPA ACL 2 种。

Agent 必须具备协作能力。当访问量增多时，要求 DMA 能够复制自身，并在多个 DMA 之间进行负载均衡，以提高响应速度。又如 TCA 必须确定哪些 Agent 参与改造并确保它们之间的同步。

4 系统实现

基于本构架本文实现了一个网上家具商场的简单原型站点。开发部署环境如下：JBuilder2005, WebLogic8.1, JADE3.4。把客户端的 Agent 实现成 Applet。服务器端的所有 Agent 都继承自 jade.core.Agent 类，Agent 的执行任务由行为类表示，行为类从 jade.core.behaviours.Behaviour 派生，通过 addBehaviour 方法添加到 Agent 中。

DMA 利用 JADE 的黄页功能，在 Directory Facilitator 中

发布服务。UCMA 实现了 Apriori 算法以识别用户模式。RTA 基于顾客在购买历史及同类顾客的访问模式推荐产品。页面采用 JSP+XML+XSLT 的技术，由 Servlet 调用有关 Agent 对内容进行改造。

本站点的家具图片都提供高、中、低 3 种分辨率的副本。其命名规则如下：比如 pink-sofa.jpg 是一张高分辨率图片，那么 pink-sofa_m.jpg 和 pink-sofa_l.jpg 则是该图片的中与低分辨率副本。根据策略和命名规则 PTA 替换页面的 URL，实现区别显示。本原型没有实现 OTA 和 OPA。

5 结束语

网络接入无线化、终端多样化是目前计算领域的趋势。结合该形势，本文提出用户上下文自适应站点的概念，并讨论其基于多 Agent 的构架。本构架最大的问题是需要客户端上安装监视插件，这引起的客户隐私安全问题需要进一步研究。虽然开发了一个简单的原型系统，但还缺乏相关评价方法。目前研究重点都在重新组织站点已有信息上，忽略了利用用户的访问模型、兴趣焦点来调整相关商务信息，增加或删除站点内容。这些是今后进一步研究的方向。

参考文献

- [1] Chen M S, Park J S, Yu P S. Data Mining for Path Traversal Patterns in a Web Environment[C]//Proc. of the 16th International Conference on Distributed Computing Systems. Hong Kong, China: [s. n.], 1996: 385-392.
- [2] Zaiane O R, Xin Man, Han Jiawei. Discovering Web Access Pattern and Trends by Applying OLAP and Data Mining Technology on Web Logs[C]//Proc. of International Forum on Research and Technology Advances in Digital Libraries. Santa Barbara, CA, USA: [s. n.], 1998: 19-29.
- [3] Jonker C M, Lam R A, Treur J. A Reusable Multi-Agent Architecture for Active Intelligent Websites[J]. Journal of Applied Intelligence, 2001, 15(1): 7-24.
- [4] Ardissono L, Goy A, Petrone G, et al. Agent Technologies for the Development of Adaptive Web Stores[C]//Lecture Notes in Computer Science. London, UK: Springer-Verlag, 2001: 194-213.
- [5] Pogaenik M. Layered Agent System Architecture for Personalized Retrieval of Information from Internet[C]//Proc. of European Signal Processing Conference. Tampere, Finland: [s. n.], 2000: 421-424.

(上接第 225 页)

参考文献

- [1] Askar H, Li Xiaofeng, Li Zaiming. Performance Analysis of Dim Moving Point Target Detection Algorithm[C]//Proc. of International Conference on Communications, Circuits and Systems and WestSino Expositions. [S. l.]: IEEE Press, 2002: 229-242.
- [2] Askar H. Performance Analysis of Dim Moving Point Target Detection Algorithms[C]//Proc. of International Conference on Communication, Circuits and Systems and WestSino Expositions. [S. l.]: IEEE Press, 2002: 605-609.
- [3] Caefer C E, Silverman J, Mooney J M. Optimization of Point Target Tracking Filters[J]. IEEE Transactions. on Aerospace and Electronic Systems, 2000, 36(1): 15-24.
- [4] 黄 静. 红外目标的检测与跟踪[D]. 大连: 大连理工大学, 2002: 3-6.
- [5] 刘志军, 陈朝阳, 沈绪榜, 等. 基于卡尔曼滤波器的背景抑制及小目标检测[J]. 华中科技大学学报, 2004, 32(12): 7-9.
- [6] 艾斯卡尔, 那斯尔江. 一种基于局部加权非参数回归估计的杂波抑制技术[J]. 激光与红外, 2005, 35(4): 294-296.