

文章编号:1001-9081(2006)02-0267-03

## 基于混合学习策略的多 Agent 信息过滤系统

费洪晓, 巩艳玲, 谢文彪, 倪敏

(中南大学 信息科学与工程学院, 湖南 长沙 410075)

(gylcsu@yahoo.com.cn)

**摘 要:** 针对用户个性化服务的要求, 给出了一种基于混合学习策略和 BP 神经网络的多 Agent 信息过滤系统实现方案。系统采用蒙特卡罗强化学习算法进行多 Agent 协作学习, 同时运用三层 BP 神经网络计算用户的满意度, 根据算出的满意度对用户兴趣模型进行更新。本系统中用户无须反复提供显示反馈, 由 Agent 跟踪并记录用户的浏览行为而得到用户的隐式反馈信息, 从而减轻了用户的负担。

**关键词:** Agent; 个性化信息过滤; 强化学习; BP 神经网络; 蒙特卡罗算法

**中图分类号:** TP393 **文献标识码:** A

## Multi-Agent system for information filtering based on hybrid learning approach

FEI Hong-xiao, GONG Yan-ling, XIE Wen-biao, NI Min

(College of Information Science and Engineering, Central South University, Changsha Hunan 410075, China)

**Abstract:** With the requirement of user's specific information service, a framework of multi-Agent system for information filtering, based on hybrid learning approach and BP neural network, was proposed in this paper. This system aimed at helping the user obtain the relevant information precisely. The Modified Monte-Carlo method was used to collaborate multi-Agent system, and the intelligent Agent was exploited to watch the user's behaviors. Besides, the BP neural network was applied to learn the user's interests in this system. A conclusion is drawn that this system is able to obtain the user's interests precisely without the user's repeatedly feedback, so as to reduce the labor of the user.

**Key words:** Agent; personalized information filtering; reinforcement learning; back propagation neural network; Monte-Carlo method

## 0 引言

随着因特网的迅猛发展, 如何从数以亿计的网页中迅速地搜索到满足用户需求的信息, 已经成为人们迫切需要解决的问题之一。传统的搜索引擎虽然在一定程度上解决了资源定位问题, 但仍不能满足人们对信息质量日益增长的需求。很多主流搜索引擎提供的服务都没有考虑每个个体用户的需求, 没有考虑用户的差异。对于任何用户只要输入相同的查询关键词, 返回的检索结果就完全相同。实际上, 不同的用户由于背景知识、兴趣爱好等方面的不同, 需要的信息往往不同。因此很多研究者试图把 Agent 技术运用到信息检索和信息过滤中<sup>[1,2]</sup>, 以设计出个性化的信息过滤系统。由于单个 Agent 能力有限, 很难完成大规模的复杂任务, 多个 Agent 组合通过协作将会提高系统的智能。然而, 在设计多 Agent 系统时, 要准确定义系统的行为以适应各种需求是比较困难的。这就要求多 Agent 系统具有学习能力。标准的监督学习、非监督学习和强化学习已成为多 Agent 进行学习的有效途径<sup>[4]</sup>, 在具体的应用中几种学习策略往往需要相互结合, 取长补短。

在设计多 Agent 信息过滤系统时面临以下两个问题: 怎样在不提供显式反馈的情况下获取用户的兴趣; 怎样保证系统所获取的用户兴趣是准确的用户兴趣。仅仅依靠有监督的学习无法解决第一个问题, 而单依靠强化学习又无法解决第二个问题。

考虑到信息过滤中需要解决的这两个问题, 本文设计了基于强化学习和 BP 神经网络的多 Agent 个性化信息过滤系统, 系统在用户无需提供显示反馈的基础上, 能够准确获取用户兴趣, 从而帮助用户准确获取有用信息。系统运行于客户端, 是一个具有学习机制, 能够监视用户浏览行为和浏览动作, 并能实时学习到用户兴趣, 以满足各个用户不同信息服务需求的一个个性化信息过滤系统。

## 1 多 Agent 信息过滤系统设计

### 1.1 信息过滤系统结构

图 1 是多 Agent 信息过滤系统的结构, 系统中包含 4 个 Agent: 用户 Agent、搜索 Agent、兴趣学习 Agent、信息过滤 Agent。用户通过用户 Agent 与系统进行交互, 系统将信息过滤结果在用户 Agent 中呈现给用户, 同时由用户 Agent 跟踪用

收稿日期: 2005-09-03; 修订日期: 2005-11-14

**作者简介:** 费洪晓(1967-), 男, 浙江嵊县人, 副教授, 主要研究方向: 网络管理与网络安全; 巩艳玲(1981-), 女, 山西河曲人, 硕士研究生, 主要研究方向: 中文信息处理、多 Agent 系统; 谢文彪(1980-), 男, 湖南宁乡人, 硕士研究生, 主要研究方向: 网络协议和网络安全; 倪敏(1981-), 女, 湖南常德人, 硕士研究生, 主要研究方向: 智能教学系统。

户的浏览行为。搜索 Agent 从用户兴趣模型中提取出查询关键词送往搜索引擎进行搜索。信息过滤 Agent 将搜索返回结果进行过滤。兴趣学习 Agent 根据用户的浏览行为更新用户的兴趣模型。

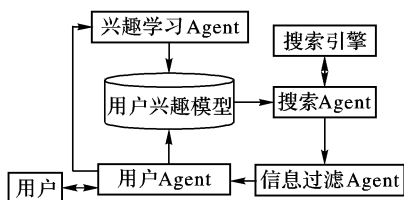


图1 多 Agent 信息过滤系统结构

各 Agent 间具体协作过程如下:

- 1) 用户 Agent 从用户那里得到用户感兴趣的针对某一主题的查询关键字,并将该查询关键字保存到用户兴趣模型中,作为初始的兴趣模型;
- 2) 搜索 Agent 从用户兴趣模型中提取查询关键字,将得到的查询关键字送往搜索引擎;
- 3) 搜索 Agent 将搜索引擎返回的结果送到信息过滤 Agent;
- 4) 信息过滤 Agent 将 3) 得到的搜索结果进行过滤;
- 5) 信息过滤 Agent 将聚类结果传送到用户 Agent;
- 6) 用户 Agent 将 5) 得到的聚类结果呈现给用户,观察并记录下用户对这些文档的反应,也即监视用户的行为,并将用户的行为信息送往兴趣学习 Agent;
- 7) 兴趣学习 Agent 根据 6) 得到的用户的行为信息计算出用户对各个 Agent 行为(也即多 Agent 系统过滤结果)的满意度 utility,并根据满意度更新用户兴趣模型;
- 8) 返回 2)。

## 1.2 信息过滤系统组成部分及功能描述

系统由以下四大模块组成。各模块之间彼此独立而又相互协作,实现了整个系统的功能。

### 1) 用户 Agent

用户 Agent 是一个接口 Agent,负责与用户交互,主要由输入、结果显示和用户行为记录单元三部分组成,用户通过登陆界面进入系统的查询界面,在查询界面中用户可以选择自己感兴趣的类别,输入查询关键字进行查询。用户 Agent 接收到用户的查询请求之后,首先检查用户的兴趣模型,如果为以往搜索过的兴趣,则提取该用户兴趣,否则把用户的查询请求添加到用户的兴趣模型中。

### 2) 搜索 Agent

搜索 Agent 负责信息的检索,并将检索的结果传送给信息过滤 Agent 进行处理,由于 WWW 环境是动态变化并且不断更新的(如更新现存文档、移除现存文档、添加新文档),对于 Agent 来说,访问同一状态的次数超过一次的概率是非常小的,几乎接近于零。Agent 每次检索都处于不同的状态下,因此本文中 Agent 的所有动作(action)都是针对特定状态 $s$ 来说的,每一幕(episode)从“开始”到“结束”这个时间段称为一个状态(state)。系统中的搜索 Agent 可以利用 Google 的 Web service 来实现。Google 提供了基于 SOAP 的 Web service,因此用户可以向 Google 服务器提交的查询请求,而后 Google 服

务器将处理这些请求,并返回 SOAP 格式的查询结果。

### 3) 信息过滤 Agent

信息过滤 Agent 将搜索结果进行分词、特征提取、聚类等一系列处理,最终将搜索结果聚类成簇(cluster),每一个簇是用户所查询主题的子主题。由于用户所查询主题可能不止一个,因此一个文档可能属于不同的簇。聚类结果包括有多少个簇,每个簇的公共属性,哪些文档属于哪些簇。由于中文信息处理的特殊性,使得中文文本聚类的算法远远不同于西文,论文采用模糊聚类算法进行过滤。

### 4) 兴趣学习 Agent

兴趣学习 Agent 负责学习用户的兴趣,根据用户的行为信息利用三层 BP 神经网络计算出用户对各个 Agent 行为的满意度,效果是一种把状态映射到实数的函数,Agent 采取不同的动作会产生不同的效果,然后依据相应的策略更新用户模型。BP 神经网络的输入为:A,B,C,D,E,F,输出为用户的满意度。经过训练的神经网络,能够有效地对所输入的用户行为进行判断,给出用户的满意度。

## 2 信息过滤系统中 Agent 学习算法实现

### 2.1 基于强化学习的用户兴趣学习算法

用户兴趣的学习是根据用户对浏览信息的选择,采取某种学习方法来逐步明确用户兴趣的一个过程。实际上它是一个机器学习的过程,可以采用多种机器学习的方法来实现。作为一种无教师学习方法,强化学习具有其他学习方法所无法比拟的优点:它考虑的是在没有外界指导的情况下,Agent 通过与环境之间的交互进行学习,通过动作引起奖惩回报,从而获得最优策略的学习过程。强化学习主要有四个组成要素:

策略:决策函数,规定了在每个可能的状态,Agent 应该采取的动作集合;

奖励函数(Reward Function):就是环境交互的过程中获取的奖励信号,奖励函数反映了 Agent 所面临的任务的性质,同时也可以作为修改策略的基础;

状态(或动作)值函数(Value Function):与奖励函数是两种不同的概念,奖励函数是 Agent 在执行具体任务时的目标(即获取最大的奖励信号),值函数则从长远的观点来判断 Agent 当前所采取动作的好坏;

环境模型:对外界环境状态的模拟,Agent 在给定的状态下执行某个动作,环境模型将预测下一个环境状态和奖励信号,为简化学习问题,有的强化学习算法没有采用环境模型。

强化学习主要算法有:动态规划算法、蒙特卡罗算法、时序差分算法、Q 学习算法。论文采用改进的蒙特卡罗强化学习算法。

在某一状态 $s$ 下,Agent 采取不同的动作会产生不同的效果,学习效果通过用户的满意度来体现。本文中所指的动作是呈现代表用户兴趣的相关文档;策略指的是从用户的兴趣模型中提取出查询关键字,使该关键字将能产生最大的效用,多 Agent 强化学习的目标就是寻找这样的最优策略 $\pi$ 。本文中的强化学习方法不同于标准蒙特卡罗学习方法,标准蒙特卡罗方法中值函数取平均回报,而本文中值函数取实时回报,

由于在同一状态下 Agent 重复执行动作  $a$  的概率很小,基本接近于零,因此实时回报也就近似于平均回报。

如图2所示用户兴趣学习流程,首先初始化动作值函数  $Q(s, a)$ , 并初始化用户兴趣模型,然后提取查询关键词进行搜索,一旦查询关键词选定就标志着一幕 (episode) 开始, Agent 依据策略  $\pi$  选择动作  $a$  (将过滤后得到的代表用户兴趣的相关文档呈现给用户), 然后依据所得回报修改  $Q(s, a)$  的值。当用户对各个簇的兴趣度都相同时,即达到学习要求,学习结束;否则继续学习直到满足学习要求时为止。

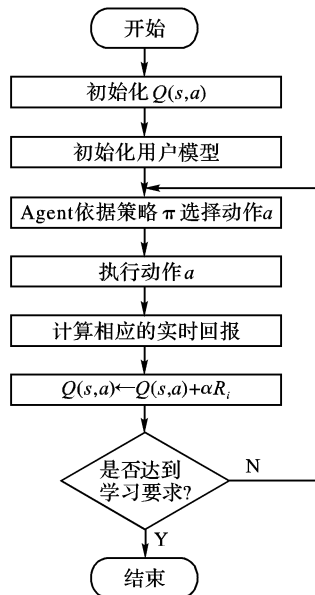


图2 用户兴趣学习流程

## 2.2 基于 BP 神经网络的用户兴趣更新算法

用户兴趣模型的更新算法的基本思想是通过对用户的浏览行为进行跟踪记录,用三层 BP 神经网络算法计算出用户的满意度,根据该满意度使用户的兴趣模型得到更新和修正。用户在浏览当前搜索过滤的结果时,会对结果进行某些操作,如:打开浏览、将页面保存、添加收藏等,用户接口 Agent 会进行分析和判断该信息是否符合自己的要求和需要。用户 Agent 对每篇文档都从以下六方面来记录用户的浏览行为: A) 用户是否 bookmark 该页面; B) 用户是否访问该网页的超链接; C) 用户是否拖动滚动条; D) 用户是否打开该文件; E) 用户是否保存该页; F) 用户在该文件停留时间。

神经网络的输入端: A, B, C, D, E, F; 输出端: 用户对 Agent 行为的满意度。图3给出了神经网络的模型。

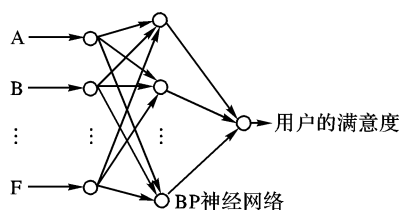


图3 BP神经网络模型

当用 BP 神经网络计算出用户对 Agent 行为的满意度之后,就可以根据该数值对用户的兴趣模型进行更新和修正。详细的用户兴趣更新算法如下程序段所示。用户对经系统过滤后呈现到界面的某文档  $j$  的满意度  $r_j$  通过 BP 神经网络计算

得出,将每个 cluster (簇) 中所有文档的  $r_j$  相加,即得到用户对该 cluster 的满意度  $R_i$ 。对于用户兴趣模型中的每一个项  $k$ , 如果  $k$  是某个 cluster  $i$  的公共项,那么  $k$  的权值将被更新。更新分两种情况:

1) 如果  $R_i$  等于零,那就表示用户对该 cluster 中所有的文档都不感兴趣,此时将该 cluster 的公共项  $k$  的权值  $W_k$  置为 0;

2) 如果  $R_i$  大于零,那么  $k$  的权值将通过  $\alpha R_i$  来更新,令  $W_k = W_k + \alpha R_i$  (其中  $\alpha$  为学习率)。

当用户兴趣模型中项  $k$  的权值逐渐减为 0 时,该项将被删除。反之,对于 cluster 中权值较高而用户模型中没有的项,系统将其添加到用户兴趣模型中。

用户兴趣更新算法如下:

```
Get user's behavior information of each document from User Agent
for each cluster  $i$ ,  $R_i = 0$ 
  for each document  $j$  inside cluster  $i$ 
    calculate user's utility  $r_j$  using BP Neural Network
     $R_i += r_j$ 
for each cluster  $i$ 
  for each term  $k$  in user profile
    If term  $k$  is a common term of cluster  $i$ 
      If  $R_i = 0$ , then set  $W_k = 0$ 
      else  $W_k = W_k + \alpha R_i$ 
For top scored cluster, extract common terms and insert into user
profile if the common term is not in current user profile
```

## 3 结语

本文设计了一种基于 Agent 的个性化信息过滤系统结构,适用于一般万维网上的智能信息搜索系统。与其他过滤系统相比<sup>[3]</sup>,本系统的优势在于:用户不需要给出显式的反馈,而是通过 Agent 跟踪并记录用户的浏览行为而得到用户的隐式反馈信息。然而,该结构下系统的实现面临着一些难点问题,突出表现为用户兴趣更新、Agent 间的通信、概念的标准化定义和多系统的集成等问题,这些研究的进展将为多 Agent 系统的进一步实用化奠定基础。

## 参考文献:

- [1] MLADENIC D. Personal WebWatcher: Implementation and Design [R]. Technical Report, IJS-DP-7472, Department of Intelligent Systems, J. Stefan Institute, Slovenia, 1996.
- [2] 冯翔,刘斌,卢增翔,等. Open Bookmark——基于 Agent 的信息过滤系统[J]. 清华大学学报(自然科学版), 2001, 41(3): 85 - 88.
- [3] SEO Y-W, ZHANG B-T. Personalized Web-document Filtering Using Reinforcement Learning [J]. Applied Artificial Intelligence, 2001, 15 (7): 665 - 685.
- [4] MITCHELL TM. 机器学习[M]. 曾华军, 张银奎, 译. 北京: 机械工业出版社, 2003.
- [5] 庄镇泉, 王煦法, 王东生. 神经网络与神经计算机[M]. 北京: 科学出版社, 2002.
- [6] 潘金贵, 胡学联, 李俊, 等. 一个个性化信息搜集 Agent 的设计与实现[J]. 软件学报, 2001, 12(7): 1074 - 1079.