

# 基于 Agent 的农业个性化搜索系统研究与应用

刘月仙<sup>1</sup>, 罗长寿<sup>1</sup>, 蒋国武<sup>2</sup>, 程新荣<sup>2</sup>

(<sup>1</sup>北京市农林科学院农业科技信息研究所, 北京 100097;

<sup>2</sup>中国农业大学信息与电气工程学院, 北京 100083)

**摘要:** Web 信息资源的增长出现了“信息过载”和“资源迷向”现象。基于 Agent 的特性, 研究设计了针对农业用户的个性化信息搜索系统, 并将其应用到农业信息咨询服务业中, 对于为用户提供个性化与智能化的信息服务具有重要的参考价值。

**关键词:** Agent; 个性化; 搜索系统

中图分类号: S126

文献标识码: A

论文编号: 2009-1391

## Research on the Application of Agent in Agricultural Personalized Information Search System

Liu Yuexian<sup>1</sup>, Luo Changshou<sup>1</sup>, Jiang Guowu<sup>2</sup>, Cheng XinRong<sup>2</sup>

(<sup>1</sup>Institute of Information on Science and Technology of Agriculture, Beijing Academy of Agriculture and Forestry Sciences, Beijing 100097;

<sup>2</sup>China Department of Computer Science, China Agriculture University, Beijing 100083)

**Abstract:** The growth of Web information resources leads the phenomenon of "information overload" and "resources disorientation". In this paper, based on the characteristics of Agent, we design and develop an information search system Meeting agricultural users' individual requirements. This system has applied in agricultural information consulting services, and has very important reference value for providing individual and intelligent information services to the users.

**Key words:** agent, personal, search system

## 0 引言

随着 Internet 网络信息量几何式的增长, 出现了信息利用的“信息过载”和“资源迷向”现象。如何在海量的农业网络数据中搜索、挖掘、发现有用的信息, 满足农业用户的个性化的农业信息服务需求, 对于推动农业生产与发展、缩小农村“数字鸿沟”等具有重要意义与作用。该研究利用 Agent 技术, 根据农业用户的兴趣和偏好, 设计了一个农业个性化搜索系统, 系统通过 Agent 自主收集用户感兴趣的信息并将查询到的信息加以过滤, 去掉用户不需要的信息, 为用户提供智能化、主动性的信息服务, 对于满足农民对农业信息的需求具有重要作用。

## 1 基于 Agent 的农业个性化搜索系统

### 1.1 Agent 的定义及特性

广义来讲 Agent 代表一切具有智能的实体, 它置身于某种环境中, 通过感应器感知环境, 通过效应器作用于环境。狭义来讲它指的是信息领域中的智能软件实体, 这种软件实体能自治的运行以代表其设计者或使用者实现一系列目标。作为 Agent 的软件或硬件系统应具有以下的特征<sup>[1-2]</sup>:

**自主性:** 也叫自治性, 指的是智能体能在没有人类或其他智能体直接干涉和指导的情况下持续运行, 并能控制其内部状态和动作。这是 Agent 区别于对象 (object) 的一个重要特征。

**基金项目:** 北京市科技新星项目“基于 Agent 技术的农业个性化咨询系统的研究”(2006B40)。

**第一作者简介:** 刘月仙, 女, 1976 年出生, 北京人, 博士, 助研, 研究方向农业信息技术。

**通讯作者:** 罗长寿, 男, 1974 年出生, 山西, 副研, 博士, 研究方向农业信息智能处理技术, 发表论文 30 余篇。通信地址: 100097 北京市农林科学院农业科技信息研究所, E-mail: luocs@agri.ac.cn。

**收稿日期:** 2009-07-08, **修回日期:** 2009-07-20。

**社会性:** Agent 拥有其他 Agent 的信息和知识, 并能通过某种 Agent 通讯语言与其他 Agent 进行交互、协同和合作。

**反应性:** Agent 能够理解周围的环境, 并对环境的变化做出实时的响应。

**主动性:** 也叫能动性, Agent 不仅简单地对其环境做出反应, 也能通过接受某些启动信息, 表现出有目标的行为。

一个 Agent 在具备以上四条最基本特性后可以再根据其应用情况拥有其他特性, 包括移动性、通信能力、理性、持续性、自启动等特性。

### 1.2 基于 Agent 的个性化搜索系统的优势

Agent 是具有感知能力、自我学习能力、问题求解能力和与外界进行通讯能力的一个自主性抽象实体, 因此将 Agent 应用于个性化搜索引擎系统中将极大提高系统的性能, 基于 Agent 的个性化搜索引擎与通用搜索引擎相比, 具有如下优势<sup>[3-5]</sup>。

(1)较强的交互能力。智能 Agent 具有社会性, 基于 Agent 的个性化搜索系统通过跟踪用户的反馈意见, 了解用户对查询结果的评价, 及时更新用户兴趣。

(2)分析、应变能力。智能 Agent 具有反应性, 基于 Agent 的个性化搜索系统在分析处理用户需求的过程中, 可以根据环境和用户意图的变化, 及时改进工作方式, 返回用户最需要的结果。

(3)提供个性化的服务。智能 Agent 具有主动性, 基于 Agent 的个性化搜索系统中设有专门的用户信息库, 记录了各种用户信息, 用户登录时, 引擎可以根据查询历史记录和习惯偏好提供合适的信息。

(4)学习、改进的能力。智能 Agent 具有自主性和学习性, 能持续的学习用户兴趣, 随着查询次数增多, 基于 Agent 的个性化搜索系统能逐渐深入、全面地了解用户的兴趣, 进而大大提高查询准确率。

(5)推理、思考能力。智能 Agent 具有推理性, 基于 Agent 的个性化搜索系统能对输入的查询语句进行推理, 而不是机械的切分语句, 从而返回最符合问题的答案。

### 1.3 基于 Agent 的农业个性化搜索系统结构

系统由多个 Agent 组成, 自顶向下分为三层: 人机交互层、信息处理层、信息搜集层, 主要由爬取 Agent、索引 Agent、用户兴趣 Agent、反馈 Agent、检索 Agent、用户兴趣模型库、用户历史信息库、注册信息库、问题模板库、网页库、农业知识库组成。系统体系结构如图 1 所示。系统首先由爬取 Agent 从农业网站爬取网页并将其存在网页库中, 索引 Agent 则利用问题模板库、农业知识库和用户兴趣模型库的内容对网页建立索引, 用户通过用户接口 Agent 向检索 Agent 提出检索请求, 检索 Agent 查询索引库, 对查询所得的结果加以过滤并排序后返回给用户。系统各 Agent 功能如下:

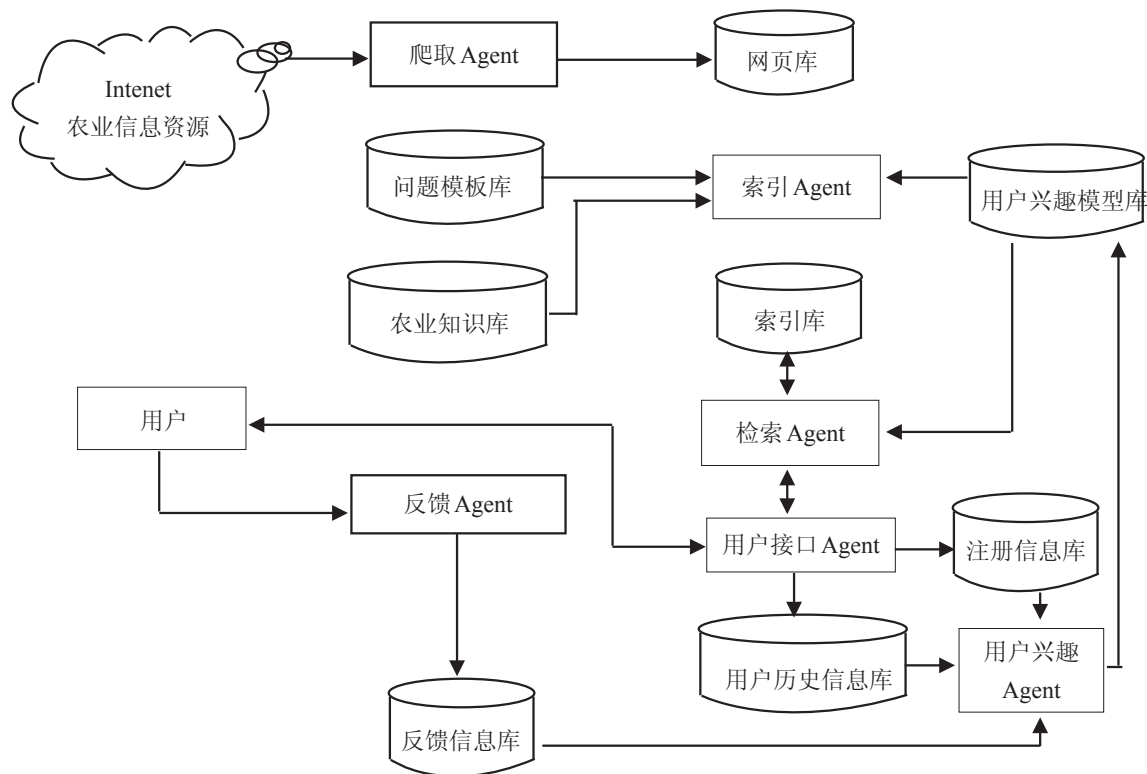


图1 基于 Agent 的农业个性化搜索系统结构

(1)爬取 Agent。用于爬取所有重要农业网站的网页并将其保存在网页库中,它能及时发现网页的更新情况,及时调整网页库,从而使网页库满足全、新的特点。

(2)索引 Agent。用于对网页库的网页建索引,它综合利用问题模板库、农业知识库和用户兴趣模型库的内容对网页进行分析、建立倒排文件,并将其存在索引库中。这种索引充分考虑了用户的兴趣,为实现个性化搜索做好准备。

(3)用户接口 Agent。该 Agent 是用户与系统交互的接口,能接受关键词、问题等查询方式,通过用户主动提供的方式获取注册信息并保存在注册信息库中,综合采用服务器端挖掘和系统被动学习的策略来捕捉用户浏览的历史记录,并将其保存在用户历史信息库中。

(4)用户兴趣 Agent。该 Agent 对注册信息库、用户历史信息库、反馈信息库中内容采取提炼、去重等步骤后,使用空间向量模型将其保存在用户兴趣模型库中,并定期使用信息增补技术对用户兴趣模型库更新。

(5)反馈 Agent。该 Agent 能将用户对查询结果的评价情况保存在反馈信息库中,保存的反馈信息通过用户兴趣 Agent 反映到用户兴趣模型库中,进而影响后面的检索效果。

(6)检索 Agent。该 Agent 根据用户接口 Agent 提供的关键词或问题,从索引库中查找匹配的网页链接,并对查询返回的集合采取过滤、消重和排序等操作,是系统很重要的部分。

(7)数据库。问题模板库保存了一些农业领域比较常见的问题样例;农业知识库保存了农业方面的专业术语、常用词语;用户兴趣模型库保存了用户的兴趣、爱好方面的信息。

#### 1.4 基于 Agent 的农业个性化搜索系统工作流程

农业个性化搜索系统的工作流程见图 2 所示,具体如下:

(1)爬取 Agent 在 Internet 上抓取农业信息资源,并将网页内容存入网页库中。

(2)索引 Agent 对网页库的内容做相应处理后,利用问题模板库、农业信息库、依照用户兴趣模型库对处理结果建立索引,并存入索引库中。

(3)用户通过查询输入框,发出检索请求,用户接口 Agent 将请求传递给检索 Agent。该接口以浏览器的方式实现。

(4)检索 Agent 在从索引库中筛选出跟检索请求相关的初始结果,接着依照用户兴趣模型对搜索结果

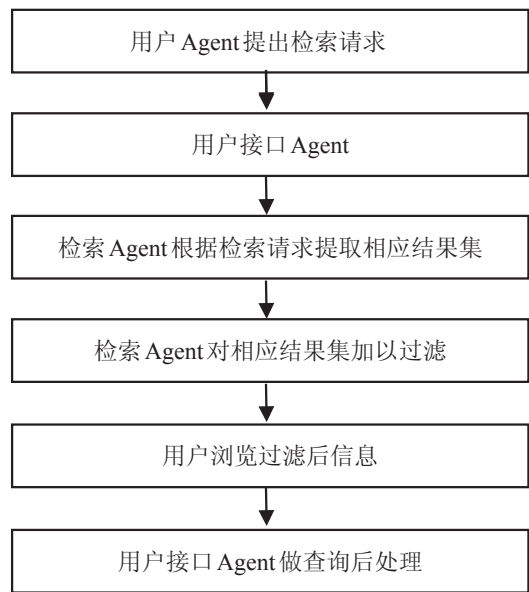


图 2 基于 Agent 的农业个性化搜索系统工作流程

进行过滤。

(5)检索 Agent 处理完毕后,将最终结果通过用户接口 Agent 返回给用户,同时用户接口 Agent 将本次查询情况存入用户历史信息库。

(6)用户在本次查询结束后将反馈信息提交反馈信息库,用户兴趣 Agent 将用户历史信息库和反馈信息库的新信息做相应处理后存入用户兴趣模型库。

## 2 系统应用

应有上述研究结果作者开发了针对农业实用技术的个性化咨询服务系统,见图 3。系统采用 Java 语言开发,版本 jdk5.0 及以上;开发工具为 MyEclipse,版本 6.0 及以上;开发框架采用 struts,所用的开发包为 Struts 1.x;服务器采用 Apache Tomcat,版本 5.0;数据库应用开源数据库 Mysql,版本 5.0;数据库 Mysql 图形操作工具采用 Navicat,版本 8.0。

### 2.1 系统模块

系统主要实现了五个模块,如下:

菜单模块:该模块有四个小按钮:“首页”、“我的兴趣”、“我要提问”、“使用指南”。点击“首页”将回到系统首页,点击“我的兴趣”将跳转到某页面,该页面会显示“我”当前的兴趣。点击“我要提问”会跳转到“我要提问”的页面,点击“使用指南”会跳转到某页面,该页面会显示系统使用帮助方面的信息。

用户登录:注册用户输入用户名和密码点击登录就进入该系统个性化首页,非注册用户点击注册将打开“用户注册表单”,填完相关信息后提交将看到欢迎界面。

知识查询: 知识查询下面分四个类: “病虫害实用技术查询”、“养殖实用技术查询”、“种植实用技术查询”、“其他实用技术查询”。点击某一个类, 会打开该类下面的问题列表。

专家咨询: 专家咨询模块分成四个小部分: “我要提问”、“回答未解决问题”、“浏览已解决问题”、“浏览全部问题”。“我要提问”是针对登录用户提出问题所用; “回答未解决问题”是登录该系统的专家回答问题

所用; 其他两个部分是针对所有用户的。

其他相关: 该模块分两个部分: “留言反馈”和“关于我们”。前者是为用户对该系统提出使用反馈和系统优化建议的; 后者是介绍系统开发和维护的相关信息的。

## 2.2 系统特色

农业实用技术个性化咨询服务系统充分结合了 Agent 技术与个性化搜索系统的关键技术, 主要具有以下特点:



图3 农业实用技术个性化咨询服务系统

(1) 很强的个性化。系统针对的是农民用户, 在获取兴趣爱好时通过挖掘用户注册信息、检索历史记录和反馈信息等方式, 综合采用服务器端挖掘、用户主动提供、系统被动学习等多种方法, 以充分地掌握用户信息, 因而具有很强的个性化搜索能力。

(2) 很好的扩展性。系统的各个 Agent 各就其职, 利于发挥自己所长, 也利于系统的扩展, 完善。比如: 爬取 Agent 的职责就是尽可能地爬取所有最新的、重要的农业网页, 系统实现者构建出的 Agent 主动性、持续性、自治性越强, 系统的性能就越好。再如, 用户兴趣更新 Agent 的职责是提取和更新用户兴趣信息, 因此, 系统实现者只要使构建出的 Agent 的推理能力、学习能力越强, 用户兴趣模型库的信息就越能体现用户最新兴趣, 就越利于系统发挥其在个性化搜索方面的优越性。

(3) 很好的自我优化性。和其他的系统有所不同, 该系统的优化不用开发人员修改其代码, 系统中的 Agent 有很好的推理和学习能力, 它们之间也有很强的交流能力, 随着用户使用次数逐渐增多, 每个 Agent 能力逐渐变强, 系统也越来越全面地掌握了用户兴趣, 系统性能将逐渐提升。

## 3 结语

Agent 具有的感知能力、自我学习能力、问题求解能力和与外界进行通讯能力, 因此在个性化信息服务领域中具有广泛的应用前景。作者对基于 Agent 的农业信息个性化搜索系统进行了研究, 并将其应用农业个性化咨询服务中, 系统可为农业用户提供个性化、准确的信息服务。为了为用户提供更人性化的服务功能, 系统模块中各 Agent 的功能有待于进一步深化完善。

## 参考文献

- [1] 宋国柱, 陈俊杰. 基于 Agent 的个性化搜索引擎关键技术研究[J]. 山西农业大学学报: 自然科学版, 2007, 27(2): 333-336.
- [2] 李燕. 多 Agent 技术在信息检索中的应用[J]. 广西师范大学学报: 自然科学版, 2007, 25(2): 111-114.
- [3] 张敏. 基于 Agent 的个性化信息检索系统设计[J]. 信息技术, 2006 (6): 143-145.
- [4] 何静媛, 张程. 基于 Agent 的搜索引擎智能个性化设计[J]. 重庆大学学报: 自然科学版, 2005, 28(12): 48-50.
- [5] 胡吉明. 个性化搜索引擎中的用户兴趣提取技术[J]. 图书馆学刊, 2006 (4): 118-119.