

电子新闻管理系统的研究

顾韵华 龚 俭

(南京气象学院计算机系 南京 210096) (东南大学计算机系)

摘 要 本文简要介绍和分析电子新闻系统的工作原理和电子新闻服务器的系统结构,阐述了电子新闻管理系统在电子新闻服务器中的地位和作用,提出电子新闻管理系统的目标及管理活动,给出一个电子新闻管理系统的详细设计方案。

关键词 电子新闻 电子新闻管理系统 NNTP 新闻管理员

APPROACHING ELECTRONIC NEWS MANAGEMENT SYSTEMS

Gu Yunhua

(Department of Computers, Nanjing Institute of Meteorology, Nanjing 210096)

Gong Jian

(Department of Computers, Southeast University, Nanjing 210096)

Abstract This paper introduces the principle of the electronic news system and the architecture of an electronic news server. It also interprets the position and function of an electronic news management system in an electronic news server. The object and management activity of an electronic management system are put forward, and a detailed scheme of the management system is described.

Keywords Electronic news Electronic news management system NNTP News administrator

1 引 言

近年来,随着 Internet 的逐步普及,它的应用正被人们所认识和接受。电子新闻是其中使用较为广泛的一种,它是采用 NNTP(网络新闻传输协议),利用计算机网络,在“新闻组”下公布消息或组织讨论的一种应用。Internet 上有许多电子新闻服务器,它们构成电子新闻发布网络,其中传播的是人们感兴趣的各个领域的信息。在电子新闻系统中,每条信息称作一则条款(article,或称文章)。

电子新闻服务器所处理的条款众多,用其服务的诸用户在阅读要求、兴趣等方面又各不相同,从而使新闻服务器的管理十分重要且繁琐。但目前 Internet 上流行的电子新闻服务软

件大多在系统管理功能上尚欠完善,本文将通过对电子新闻系统组成及工作原理的简要分析,论述电子新闻管理系统的目标和管理活动,提出一个电子新闻管理系统设计方案,并对其功能、结构及实现进行了讨论。

2 系统目标与管理活动

电子新闻系统使用 NNTP 协议,NNTP 是基于可靠流传输、采用 client/server 模式,对新闻条款进行发布、检索、邮寄、索引及失效等控制的协议。新闻条款从一个新闻服务器发布到另一个或多个新闻服务器,如此扩散开去,便形成了新闻传播网。新闻服务器是电子新闻系统的关键,每个新闻服务器提供多个新闻组,用户可按自己的兴趣选择新闻组。

2.1 系统目标

电子新闻服务器的核心是新闻接收转发系统和新闻阅读处理系统。前者负责接收、存储本网点用户邮寄条款和其馈送网点(是另一个或几个电子新闻服务器)发布来的条款,并可转发条款副本到被本网点馈送的其它电子新闻服务器去;后者对客户端来的阅读请求进行响应、查询新闻数据库、将条款送至用户标准输出,供其阅读。新闻管理系统是为了保证新闻服务系统有效地运行而设立的,它建立在两个核心系统基础之上,又为其提供服务,是电子新闻服务器的重要部分。电子新闻服务器系统模型如图 1 所示。电子新闻管理系统应实现的主要目标是:对新闻数据库、用户及服务器运行进行监控和信息收集,为新闻管理员了解服务器状况、限制用户访问、调整新闻组等管理问题提供决策依据。

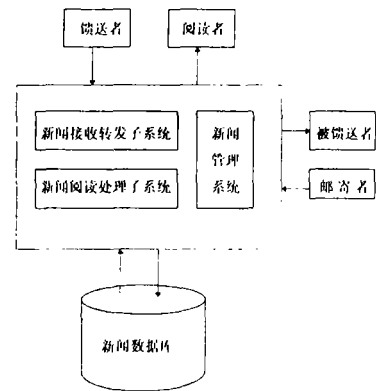


图 1 电子新闻服务器系统模型

2.2 管理活动

电子新闻管理系统的管理活动主要有:(1)新闻数据库中条款质量控制;(2)新闻数据库中重要条款归档;(3)用户管理;(4)新闻服务系统运行数据收集;(5)新闻服务系统运行的在线监测;(6)新闻服务质量分析和系统调整。

3 系统功能与逻辑结构

电子新闻管理系统主要与新闻服务器主机相关,较少涉及网络交互。这里提出的电子新闻管理系统是一个“外挂式”系统,即将主要管理功能与核心系统分离,这样可以使核心系统致力于处理新闻接收、转发与阅读,同时也考虑到 Internet 电子新闻服务器核心系统的成熟性。

3.1 系统功能

电子新闻管理系统的功能可随所用新闻服务软件及设计者的需要而定,一般可有以下几点:

(1)定期获取新闻馈送者馈送条款、阅读者访问新闻组及阅读条款、邮寄者邮寄条款、本服务器馈送给其它服务器新闻条款以及对本服务器非法访问等重要运行数据;对所获得的数据进行统计、分析,得出相关结论,以图形、表格等形式输出;

- (2)新闻条款审查、过滤;
- (3)服务器使用者的在线监测;
- (4)条款归档保存。

3.2 系统逻辑结构

该电子新闻管理系统分为四个子系统:新闻订阅子系统、条款审查子系统、条款归档子系统和分析子系统,它与核心系统及管理员之间的关系如图2所示。

(1)新闻自动订阅子系统:实现阅读自动订阅及对邮寄者邮寄新闻组意向的自动处理。

(2)新闻条款审查子系统:对邮寄条款进行过滤,阻止质量低劣或不良的条款加入新闻数据库。可通过新闻仲裁员人工检查与自动过滤两条途径:对较重要或较敏感的新闻组可以人工审查,其它新闻组可自动审查。

(3)条款归档子系统:将有价值的条款保存指定时间,即保存“精华本”。

(4)分析子系统:对系统运行数据在线获取以及事后获取、统计、分析,使新闻管理员及时了解系统的当前运行状况和近期运行、访问情况,后者作为新闻管理员调整系统运行参数的依据。

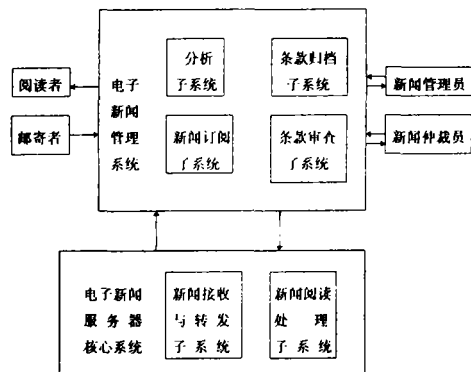


图2 电子新闻管理系统逻辑结构

4 系统模块结构

本系统由图3所示的模块构成。

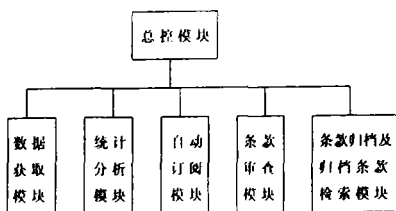


图3 电子新闻管理系统模块结构

4.1 数据获取模块

包括在线监测和事后数据获取。在线监测可随时报告连接访问情况;事后数据获取是将新闻系统的运行数据定期(如一天一次)收集整理,供统计分析模块使用。

4.2 统计分析模块

该模块与在线数据获取、事后数据获取模块一起完成分析子系统功能。统计分析模块按给出的任意时段对馈送新闻、阅读新闻、转发新闻、邮寄新闻及非法访问情况进行统计,并根据统计分析的结果得出相关结论(如馈送高峰新闻组、新闻组访问分布等)。

4.3 自动订阅模块

它完成新闻订阅子系统功能。分别设计新闻组订阅表和邮寄预定表,由用户填写。根据阅读者或邮寄者反馈回的表格,新闻管理员酌情予以批准,以批准后的结果由本模块自动修改

用户访问控制表,从而实现新闻的自动订阅。

4.4 条款自动审查模块

它实现条款审查子系统中自动审查功能。可根据多种自动审查策略来完成对条款的检查。审查策略可视新闻服务器的具体情况而定,如关键词过滤策略等。

4.5 条款归档与归档条款检索模块

“条款归档”部分通过定期对新闻数据库扫描,将选中的条款装入“精华”新闻库中。“归档条款检索”部分提供方法让新闻管理员及某些用户完成查询等工作。

5 系统实现要点讨论

Internet 上常见的新闻服务软件的核心系统功能大多提供一定量的管理信息,如系统运行日志。这是电子新闻管理系统实现的基础,在管理系统实现中,我们认为以下几点是关键。

5.1 电子新闻管理系统与服务软件核心模块的联接

找出新闻管理系统与服务软件核心模块产生数据联系之处,在管理系统中作适当处理,而尽量不改动原服务软件。在本方案中,管理系统与原系统数据发生联系有四处:

(1)数据收集处,这是单向联系,即数据从原系统流向管理系统,这在管理系统中不必作特殊处理;

(2)条款归档处,具有与(1)相同性质;

(3)条款审查处,数据双向流动,当管理系统决定条款进入新闻数据库时要修改新闻数据库的状态数据,该数据是管理系统与原系统共同操作的数据,在管理系统实现中要保证该数据的正确性;

(4)新闻订阅处理中,数据双向流动,与(3)类似处理。

5.2 数据收集频度及数据预处理

现有新闻服务系统的核心系统在运行时都记录访问数据于其日志文件中,但若按日志原样保存运行数据而不加处理,则会有两个问题:一是保存的数据量大且多数属无用数据;二是日志文件的更新周期一般为一星期或稍长,这样历史较长的运行数据将被丢弃。故在数据获取时,应将日志中关键数据提取出来,保存到一个永久文件中。数据收集频度取值一般与日志更新频度一致。

5.3 交互界面

电子新闻管理系统的多数模块都涉及交互界面设计,尤其新闻订阅模块中涉及到与用户的远程交互,其界面更应直观、易于操作。界面设计方法有多种,目前 WWW 服务正被越来越多的 Internet 用户使用,利用其方便友好的交互功能设计对我们的交互界面来说是一种很好的选择。采用这种方法的要求是在客户端必须运行浏览器软件(如 Netscape Navigator)。

5.4 模式匹配

模式匹配在电子新闻管理系统中也是一项重要的技术,在统计分析、条款审查、条款归档等模块中该技术是关键之一。在实现模式匹配时可利用已有的模式匹配工具,如 Unix 系统的 awk 实用程序。

5.5 条款归档策略和条款审查策略

这是很本地化的问题。归档策略应避免在硬盘上长期保存过多的条款,可按主题(新闻

组)、日期索引等方式保存条款,并应定期删除已无价值的条款。审查策略应多从条款意义上考虑,故重要新闻组中的条款审查要由人工进行,自动审查仅能作机械性的匹配检查。

6 结 论

本文通过对电子新闻管理系统在电子新闻服务器中的地位的分析,说明了电子新闻管理系统是电子新闻服务器中不可缺少的有机组成部分,其性能直接影响到电子新闻服务器的服务质量。文中提出的外挂式电子新闻管理系统设计方案,在功能上是较完善的,其与核心系统的衔接设计是有效合理的。该方案的一部分已在 CERNET 华东(北)地区网络中心的 Sun 工作站上实现。通过数据获取和分析,可得出指定的任意时段的服务器访问情况。整个电子新闻管理系统的实现将使电子新闻服务系统成为一个更可用、有效的系统。文中的设计思想可推广到其它网络新闻服务(如 BBS)及其它网络应用中。

参 考 文 献

- [1] K. Brain, L. Phil, Network News Transfer Protocol, A Proposed Standard for the Stream - Based Transmission. RFC: 977, 1986.
- [2] H. Hareley, S. Rick, Internet 参考大全, 第一版, 学苑出版社, 1994.
- [3] 孙玉方等, UNIX 系统工具与应用, 第一版, 清华大学出版社, 广西科学技术出版社, 1992.

(上接第 20 页)

面,使得对实际生产过程的实时控制成为可能。证明本套软件确实具有组态软件中图形组态应该具备的通用性、实时性、可靠性、灵活性、易于维护性和人机界面友好性等特点。而且,在实现动态图素的绘制时运用了面向对象的编程思想,这大大提高了该软件的扩充性,有助于以后结合面向对象技术进一步增强该图形组态的功能。

参 考 文 献

- [1] 王常力、廖道文,集散控制系统的设计和应用,清华大学出版社,1993年6月,pp. 159 ~ 187.
- [2] 李丹、许少云,“用于工业过程控制的一种新工具——组态软件”,《计算技术与自动化》,pp. 5 ~ 6, 1995, 9.
- [3] 李枫、王军鹰,“国内工控组态软件发展动向”,《电子技术应用》,pp. 53 ~ 55, 1996, 2.
- [4] 徐平,集散型控制系统监控组态软件的设计与实现,清华大学毕业论文,1996年7月,pp. 12 ~ 13.
- [5] 林启隆, VISUAL BASIC FOR WINDOWS 程序设计,清华大学出版社,1994年,pp. 32 ~ 155.
- [6] [美] Tim Farrell, Runnoe Connally 著(李莉等译), WINDOWS 3.1 编程指南,清华大学出版社,pp. 33 ~ 56.