

以 CGI 为基础的农业技术分布式查询数据库^①

—以江苏省农业技术分布式查询数据库为例

周慧珍 李 勇 章扬德 石常蕴

(中国科学院南京土壤研究所 南京 210008)

阎国年

(南京师范大学地理信息科学江苏省重点实验室)

摘 要 农业是发展国民经济的基础,农业信息是农业生产力发展的重要核心和战略资源,信息共享技术为充分利用农业信息资源奠定了基础。本研究应用基于 INTERNET 网络、UNIX 为数据通信平台, Foxpro 为数据库平台,使用客户机与服务器 (Client/Server) 为网络通信模型,利用 UNIX 服务器上的 Shell 脚本语言编译器,以简单的 ASCII 码文本文件形式的解释型脚本语言为基础开发了 CGI 技术,采用与 html 表单一道使用的 POST 请求方法,实现了江苏省农业技术信息在因特网上的远程交互式动态查询和信息共享。

关键词 分布式查询;数据库;CGI

农业是国民经济可持续发展的基础,农业信息是促进我国农业进一步发展和进步的主要基础之一,已成为农业生产力发展的重要核心和战略资源。以计算机技术、通信技术、网络技术为代表的现代信息技术,使人类对农业信息资源的利用进入了高效率、专业化、多样化、共享化的现代方式。

江苏省是我国农业基地,江苏的农业在中国农业中占极其重要的地位。如何为江苏农业的进一步开发和持续发展、走农业致富的道路做出贡献,是我们义不容辞的职责。于是建立一个基于 INTERNET 网的农业技术转让数据库被构思和设想。该数据库建立的最终目标要达到便于供省内外、国内外有关部门、科研和技术人员以及致富道路的农民查询,帮助他们找到能解决特定农业开发和农业生产所遇到各种技术问题的单位和专家,或进行农业技术的劳务输出。

1 材料和方法

1.1 数据库结构与信息内容构思

基于上述宗旨,江苏省农业技术服务分布式数据库由 2 个子库组成:(1) 农业服务机构咨询库;(2) 农业服务专家咨询库。2 个子库的数据模型均为关系式数据模型,即每个子数据库相当于一张二维的关系式大表,直观性强。

农业服务机构数据库的每个记录由单位名称、专业范围、服务项目、地址、邮编、电话、传真等字段组成。农业专家数据库的每个记录由姓名、性别、年龄、职称、所在单位、专业范围、特长、通信地址及邮编、家庭电话、单位电话、传真、E-mail 等组成。本数据库力求界面简练、清晰、动态性使用中文查询,所有需要查询的内容都在选项框的下拉式菜单中直接显示,用户可以根据自己所需解决的业务、要求对方的专业及所在行政区域查到所需的单位和专家,这样便于不同知识层次的用户都能理解并进行查询。

1.2 信息源的现势性与信息收集方式

^① 江苏省科技工委与南京师范大学地理信息科学江苏省重点实验室基金资助项目 (BE 97620)。

本数据库的信息源为实时信息,即利用当时收集的信息立即建库,故现势性强。采用 2 种方式收集信息:1.发信方式:按建库需要的信息事先设计好所需的信息内容表,共印发标准化的填写表格 15000 份左右,发往有关部门、单位和专家,请他们按要求填写后寄回作者,共反馈信息 9000 多份;2.实地调查和收集:派专人去有关单位和部门进行调查并填写信息表,一般在本市和临近市郊。通过两种信息收集方式完成的数据记录约 9200 条左右,数据量为 6 兆左右。

1.3 技术物理手段

1.3.1 数据库开发与建立的操作平台 数据库开发与建立的硬件操作平台为奔腾 166 微机,32M 内存,2.1M 硬盘。软件操作平台为 Windows95 下的 Foxpro 2.5B。数据格式为 DBF 格式。建库后放置于 UNIX 服务器上,用 CGI 语言脚本达到查询的目的。

1.3.2 Web 通信模型与运行平台 本研究采用客户机 (CLIENT) 和服务器 (SERVER) 的通信模型。网络通信平台的 WWW 服务器为 UNIX 操作系统的 SUN 20 工作站,其特点是速度快、存量大。将数据库放在服务器一端,用户通过浏览器进行动态式和交互式查询所需要的信息。客户机的硬件可为微机或 UNIX 操作系统的工作站。浏览器软件为 NETSCAPE 3.0 或微软的 IE 产品。不专门设置数据库运行服务器,而是用 UNIX 服务器上的 Shell 脚本语言编译器,以简单的 ASCII 码文本文件形式的解释型脚本语言为基础开发了 CGI 技术,采用与 html 超文本语言编辑的数据库表单一道使用的 POST 请求方法,实现了江苏省农业技术信息在因特网上的远程交互式动态查询和信息共享。

1.3.3 CGI 的运行机制与过程 CGI (Common Gateway Interface) 称为公共网关接口。CGI 程序常称之为脚本 (CGI Script)。实际上它是一种标准,它提供了 WWW 服务器和脚本之间交换数据的一套标准界面,是 WWW 服务器与外部脚本程序的桥梁,用来弥补 WWW 所达不到的功能。通过 CGI,WWW 服务器根据客户机在进行请求时所采用的方法,收集由客户机提供的信息,并将该信息发送给 CGI 脚本,然后 CGI 脚本进行信息处理并将结果返回服务器。服务器再对信息进行分析,最后将结果发送给客户端。图 1 说明了服务器、CGI 和脚本之间的关系。

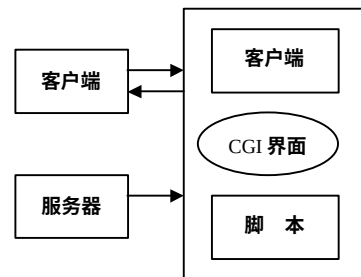


图 1. 服务器、CGI 和脚本的关系

本数据库的通信过程为客户端通过 FORM 表单以 GET 和 POST 命令两种方式向服务器提交查询请求,当查询串长度 < 255 个字符时用 GET 方法。由于本数据库主要采用令访问者使用表单向服务器发送信息,故以 POST 查询方法为主。服务器收到客户端请求后,并将客户端传递过来的参数传递给服务器,服务器再以标准输入 (STDIN) 的方式传递给 CGI 脚本,脚本对要求的事务加以分析处理,最后将处理的结果通过以标准输出 (STDOUT) 的方式传递给服务器,由服务器再传递给客户端。整个过程如图 2。

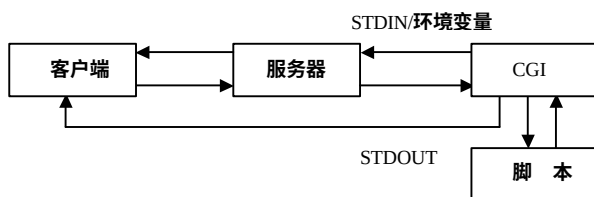


图 2. 服务器、CGI 和脚本的流程图

CGI 脚本存放在 Web 服务器根目录下的一级目录中,目录名为 CGI-BIN,其路径为 /usr/bin/http/webroot/cgi-bin。网址 <http://www.issas.ac.cn/gxsjk.htm>。

1.4 工艺流程

1.4.1 信息收集、分析、分类 按合同内容要求首先设计 2 张关系式信息表：服务机构信息表和专家信息表。反馈的信息表为 9600 多份，经过筛选，录用了 9200 份信息。其中，服务机构 6200 份，服务专家信息 3000 份，专家是指副高级职称以上的人员，只少数几个是经单位推荐的中级人员。信息表到位后，在两大类中将仅收集到的信息以能提供服务的项目为基础进一步细分。服务机构分出 100 个类型；服务专家分出 103 个类型。

1.4.2 数据库建立 在 PC 机平台上用 FOXPRO 数据库软件建立江苏省农业技术咨询服务机构和专家 2 个子数据库作为江苏省农业技术咨询服务分布式查询数据库的组成部分。

1.4.3 CGI 脚本程序与 HTML 网页的编写和制作 基于 INTERNET 网上 http 通信协议，http 的通信包括 4 个步骤：建立通信、发出请求、给出响应、结束通信。用 Shell 语言编写 CGI 脚本程序，并编辑 CGI 与 HTML 动态查询通信界面，该界面的查询方式由 3 部分组成：1. 按行政区域查询；2. 按专业范围查询；3. 按业务特长查询。例如用户要查询南京地区的兽医专家，只要用光标按住下标中的南京市然后再按另一下标中的兽医栏，就能查到南京市所有已被收集的兽医专家名单、通信地址和专长等信息。

一般，客户端向 INTERNET 某地址的 WWW 服务器特定的端口发出请求，这个端口的缺省值是 80h，它专门提供 http 协议服务。当用户呼通了 WWW 服务器的这个端口，客户机和服务器间的 http 连接就建立了。

建立连接后，客户端向服务器发出一个请求。请求程序的内容包括请求方法、请求内容和请求字段 3 方面。本程序基本上采用了与 Html 表单一道使用的 POST 请求方法，便于客户进行动态式查询。请求的内容为放置在 WWW 服务器某目录下的 2 个子数据库，使用统一资源定位器（URL）来识别这 2 个库。请求字段有客户机的地址、客户端使用的浏览器、可接受的数据类型（MIME）等等。

响应程序包含服务器使用的 Html 版本、不同出错的响应状态、数据类型等等。一次通信结束后关闭连接，新的请求则重新开始。

1.4.4 数据检查 数据库建立后，将所有数据打印出来，经过检查和修改，目的是纠正录入过程中的一些误差。

1.4.5 数据上网正式运行 将建好的 2 个子库，通过 Html 编写的网页连接为一个整体，放到 UNIX 服务器上，对 CGI 脚本程序进行调试和纠正，以达到正式上网运行的目的。

2 结论和讨论

1. 与农业发展和开发有关的分布式查询数据库大有可为。中国是世界的农业大国，江苏省是中国的粮食基地，也是现代化农业开发和发展的基地，农业技术发达，农业人才和农业研究大专院校、科研部门和农垦服务的专业单位相对集中。利用 INTERNET 网开发出所有能为中国农业发展服务作出贡献的分布式查询数据库，不仅为中国和江苏的农业发展带来很大的社会和经济效益，也为江苏的农业技术和为农业服务的人才走向世界奠定了基础。实际上江苏还有无数个能为农业服务的数据库等待去开发和利用。

2. 数据输入的规范化和标准化是保证数据库质量的基础。CGI 与 HTML FORM 查询要求在数据库建立的输入过程中具有严格的标准，否则会影响查询界面的质量，甚至出现乱码。

3. UNIX 平台为 WWW 服务器，利用 UNIX 操作系统中的 Shell 编译器进行 CGI（公共
（下转第 294 页）

