

应用 XML 技术在互联网上发布数据库

王 静, 傅秀芬

(广东工业大学 计算机学院, 广东 广州 510090)

摘要: 综合介绍了 XML 技术, 重点讲述了利用 XML 技术在 Internet 上发布数据库的方法. 说明了该方法与传统的用 ASP 技术发布数据库的区别, 给出了一个在互联网上发布数据库的例子. 最后总结了用 XML 技术进行 Web 开发的优点.

关键词: XML; 互连网; 数据库; ASP

中图分类号: TP311.52

文献标识码: A

文章编号: 1007-7162(2001)04-0035-05

基于 Web 的数据库开发是当今 Internet 技术应用的一个热点. 在 Web 的早期阶段, 较为通用的数据库发布技术有 ISAPI、NSAPI、IDC、JDBC 和 ASP 等, 它们通过 Web 服务器调用程序模块或在 HTML 中内嵌程序的办法实现发布数据库. 以 ASP 为代表, 这个阶段信息页面从本质上讲是静态的对象, 一旦信息页面显示出来, 它就固定了, 直到 Web 服务器上重新下载该文档, 造成对服务器的访问次数过于频繁. 现在简单的脚本语言如 JavaScript、Java 以及 ActiveX 控制改变了这种静态发布的现象, 为 Web 页面加入了动态内容; HTML 已经可以作为创建用户界面、处理数据及与服务器产品程序进行对话等的动态环境. 这种灵活性以及可编程性提供了在熟悉的 HTML 环境中以及现在的 Web 浏览器中利用 XML 文档的丰富结构的可能性. 而用插件技术来实现 Web 上的应用有一个缺陷就是下载部分太多, 那么在插件部分内部的内容搜索引擎是无法访问的, 因此不利于 Web 上数据的搜索. 而 XML 技术使我们能够从数据的角度来看待 Internet, 所以更利于数据库的发布, 它掀起了一场 Web 页面语言的变革.

1 XML 技术综述

1.1 XML 技术^[1]

XML(Extensible Markup Language, 可扩展标示语言)是从 SGML 所衍生出来的简化格式, 它使制作页面有了一个更具弹性、更容易添加新功能而又不失统一标准的语言格式, 并使 SGML 的优秀品质能方便而直接地被用在 Web 开发上. XML 主要有三个部分: 文档类型定义 DTD(Document Type Definition)用于规定文档的逻辑结构, 定义页面的元素、元素的属性以及元素和属性之间的关系; 可扩展样式语言 XSL(Extensible Style Language)用于规定 XML 文档样式; 可扩展链接语言 XLL(Extensible Link Language)支持目前 Web 上已有的简单链接, 并将进一步扩展链接, 包括结束死链接及间接链接.

1.2 XSL 技术^[1]

一个 XSL 样式表包括一套规则, 它们告诉 XML 处理器(如 msxml)如何将 XML 文档转化为其他的

收稿日期: 2001-04-15

作者简介: 王静(1978-), 女, 99 级研究生, 主要研究方向为网络与多媒体软件.

表示法. 总体说来, XSL 可以创建独立于任何输出表示法的样式表, 如 HTML, Tex, RTF 等等. XSL 样式表的基础是在 XML 文档中遇到特定元素时激活的规则. 在一条规则中, 你可以指定你想要 HTML 输出什么, 它的范围可以从什么都没有(即删除)到运行 JavaScript 应用程序所产生的结果.

以下是 XSL 中规则包含的几种模式:

a. 选择模式

< xsl:for-each> 、< xsl:value-of> 和< xsl:apply-templates>

选择模式语句将数据从 XML 中提取出来, 是一种简单获得数据的方法, 这几个标记都有一个 select 属性, 选取 XML 中特定的结点数据.

b. 测试模式

< xsl:if> 、< xsl:choose> 、< xsl:when> 和< xsl:otherwise>

如果我们对 XML 数据不需要全部输出, 而只需要其中的满足某条件的部分数据, 那么就要使用测试模式.

c. 匹配模式

< xsl:template>

XSL 模板将 XSL 的设计细化成一个个模板(块), 最后再将这些模板(块)组合成一个完整的 XSL; 从 XML 选取满足条件的节点, 针对这些特定的节点形成一个特定输出形式的模板.

2 利用 XML 技术在 Web 上发布数据库

2.1 XML 技术在 Web 上发布数据库的原理^[2]

由于 XML 中的数据是以数据库的结构来存放的, 因此我们可以很方便地将数据库中的数据转变为 XML 文件, 只需进行类似于文件之间读写操作即可. 把转变后形成的 XML 文件放在客户端, 从而可以关闭与数据库的连接, 与 XML 文件直接交互就可以了. 这时我们用 XSL 样式表来对 XML 文件进行重构以及计算操作, 生成 HTML 文件.

现在浏览器发展很迅速, 像 Internet Explorer 4.0 之类的浏览器可以在客户端处理 XML, 所以就提供了这种选择的可能性——在服务器端将数据转变为 XML, 让浏览器直接处理它而不是在服务器端便全部转变为 HTML.

应用 XML 在 Web 上发布数据库的原理如图 1 所示.

2.2 与 ASP 技术在 Web 上发布数据库的区别^[3]

ASP 是通过组件和对象实现客户和服务之间的信息交互, 它的语法规则是以 HTML 作为骨架再嵌入脚本语言 ASP 指令. ASP 脚本首先在服务器上打开数据库, 由 ADO 组件对象通过 ODBC 实现到数据库的连接, 脚本文件按要求生成最终的 HTML 文件, 送到浏览器端输出显示.

应用 ASP 在 Web 上发布数据库的原理如图 2 所示.

2.3 浏览器的选择

用 XML 在 Web 上发布数据库不需要专门的服务, 它只需要由服务器将数据库转变为相应的 XML 文件, 而这种用于转变的脚本语言可以用多种语言. 转变后所有的工作便可以交给浏览器来完成. 这种方法对浏览器要有一定的要求和

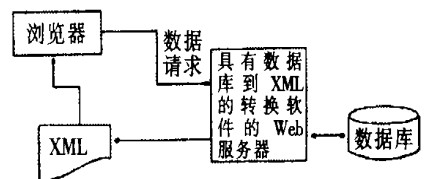


图 1 应用 XML 在 Internet 中发布数据库的原理

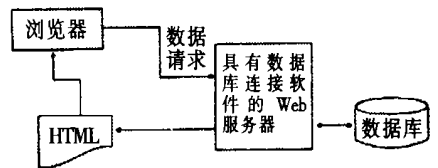


图 2 应用 ASP 在 Internet 中发布数据库的原理

规定,因为并非所有支持 HTML 的浏览器都会支持 XML. 目前支持 XML 的浏览器包括 IE4.0 及其后来的版本.

3 在互连网上发布数据库的一个实例

3.1 利用 ASP 技术

下面举一个查询电脑价格的应用. 本应用的开发环境是 IIS5.0+ Win2000. 首先设置数据源 ODBC DSN 和配置服务器, 根据任务要求, 建立一个 ACCESS 数据库 com.mdb, 内有一个数据表 pc(名称、容量、价格), 为该数据源设置一个系统 DSN, 命名为 computer, 密码和用户名为空. 另外, 在 IIS 宿主目录下专门为此应用配置一个虚拟路径/jwang. 这个应用需要两个页面, 一个是提供查询功能的 search.htm 页面; 另一个是服务器端返回查询结果的 result.asp 页面, 保存在虚拟路径/jwang 的目录下. 源程序在/论文目录.

传统的用 ASP 的源程序部分源代码即对最贵电脑查询部分代码如下:

```
<script language="VBScript" RUNAT="Server">
Dim Conn
Dim rs
Dim SQLstr
Dim DearestPrice
DearestPrice= 0
Set Conn= Server.CreateObject("ADODB.Connection")
Set rs= Server.CreateObject("ADODB.RecordSet")
Conn.open "computer"
SQLstr= "select* from pc"
rs.open SQLstr, Conn
... ..
Do While not rs.EOF
    ... ..
    if DearestPrice< rs("price") then
        DearestPrice= rs("price")
    end if
    rs.MoveNext
Loop
response.write("< B> 最贵的电脑的价格为:< /B> ")
response.write(DearestPrice)
rs.close
Conn.Close
< /script>
```

3.2 利用 XML 技术

本应用的开发环境是 IE5.0+ Win2000.

1) 把数据表 pc 导出为文件 table1.txt 数据库的导出表示法种类是一种行格式的、以某种分

隔符分开的简单表结构,这样可以避免在 XML 转变中引入任何特定于数据库的代码.保留表头字段,各字段间数据以逗号为分隔符,字符字段的前导符设为空,即 null.

2) 运行 C++ 程序 form.cpp 将 tablel.txt 文件转变为 result.xml 文件. Form.cpp 程序的功能是从 tablel.txt 中读取数据库中的数据,然后将它们以 xml 文件的格式写入到一个新文件 result.xml 中去.当然本程序也可以用其他多种语言开发.

3) 用 XSL 文件 pc.xsl 将 pc.xml 文件转变为 pc.htm 文件.生成 XML 文件之后,采用 XSL 样式表语言实现从 XML 向 HTML 的转变,这是一种实现 HTML 与 XML 无缝结合的强大技术.尽管 XSL 标准提案还处于初级状态,但是某些公司(如 Microsoft 和 Arbotext)已经推出大量的 proof of concept 工具.这里使用 Microsoft 公司的 XML 工具(msxsl)将 PC 目录的 XML 文档转变为不同形式的 HTML.

在命令行中运行语句

```
msxsl pc.xml pc.xsl -o pc.htm
```

mscsl.exe 命令行处理程序是一个大小不超过 16K 的小程序,它调用微软 xml 解析器 3.0 (msxm13.dll) 来执行从 xml 文件到 htm 文件的转变,因此运行它之前应先确保安装了 msxm13.dll 文件.

xsl 中两个元素 <xsl:eval> 与 <xsl:script> 可以实现对 xml 文档内容加入一些统计信息.<xsl:eval> 含义这计算脚本表达式,输出一个文本字符串.<xsl:script> 含义为声明全局变量或定义函数.具体代码如下来输出在 pc 表的最昂贵的 pc 的价格.

pc.xsl 文件代码如下:

```
<?xml version="1.0" encoding="GB2312?">
<HTML xmlns:xsl="http://www.w3.org/TR/WD-xsl">
<xsl:for-each select="Records">
<HEAD>
<TITLE> 电脑报价</TITLE>
</HEAD>
<BODY>
<TABLE border="1" cellspacing="0">
<CAPTION style="font-size:150%;font-weight:bold">
电脑报价
</CAPTION>
<TH> 名称</TH> <TH> 容量</TH> <TH> 价格</TH>
<xsl:for-each select="Record">
<P/>
<TR>
<TD> <xsl:value-of select="Name"/> </TD>
<TD> <xsl:value-of select="Capacity"/> </TD>
<TD> <xsl:value-of select="Price"/> </TD>
</TR>
</xsl:for-each>
```

```
< /TABLE>
最贵的电脑价格为: < xsl: eval> max( this, "Price" )< /xsl: eval>
< xsl: script>
function max( node, Price) {
temp= 0;
mark= '/Records/ Record/' + Price;
v= node. selectNodes( mark);
for( t= v. nextNode(); t; t= v. nextNode() ) {
if ( Number( t. text) > temp)
temp= number( t. text);
}
return temp;
}
< /xsl: script>
< /BODY>
< /xsl: for-each>
< /HTML>
```

DOM(Document Object Model, 文档对象模型) 是程序访问和维护 XML 文档的 API. DOM 是使用接口和对象来表示文档的, 面向对象是现在编程的主流方式, 在表示为对象的前提下, 脚本语言就可以很方便地访问文档了. 只要用 DOM 提供的 API, 程序就可以创建基本的文档. selectNodes() ——是XMLDOM 对象的一个方法, 返回文档中所有满足条件的结点的集合, 条件与< xsl: for-each> 和 select 属性的取值采用同样的写法, 可以加筛选、下标等限制. 这种方法主要用来实现对 XML 文档的查询.

3.3 两种技术比较

从运行结果(图 3, 图 4) 来看两种技术的实现是相似的, 然而本文提出了一种基于 XML 的动态网页技术, 浏览器可以直接完成查询等功能而不需经过服务器, 新的中间层服务器将所有内容转化为特定的 XML 形式, 这种技术有着更为强大的数据处理功能, 减小了服务器的负荷, 同时又克服了插件技术的下载部分太多的缺陷, 并且可以很容易地转化为包括 HTML 在内的多种形式, 因此相对来说具有在 web 上更大的发展潜力. 可以预见, 随着浏览器技术的进一步发展, XML 在 Web 上将会得到更广泛的应用和发展.



图 3 asp 运行结果



图 4 xml 运行结果

(下转第 49 页)

种实体之间的相互作用的关系, 它使程序开发效率提高, 维护费用降低. 桥梁程序设计应该使用面向对象方法, C++ 是进行面向对象的桥梁程序设计的有力工具.

参考文献:

- [1] 万明坤主编. 桥梁漫笔[M]. 北京: 中国铁道出版社, 1985.
- [2] 范立础, 林元培编. 桥梁工程[M]. 北京: 人民交通出版社, 1986.
- [3] 石洞编著. 桥梁结构电算[M]. 上海: 同济大学出版社, 1989.

The Application of Computer in Bridge Design

LIN Ji-chun, WANG Chang-wu

(College of Traffic and Communication, Huazhong Univ. of Sci. and Tech, Wuhan, 430074, China)

Abstract: In this paper, the development of bridge engineering is introduced, and the advantage of bridge design aided by computer is illustrated. The shortcomings of traditional methods are analysed. Furthermore, the Object Software design method is a tendency in bridge design.

Key words: bridge engineering; computer-aided calculation; bridge design

(上接第 39 页)

参考文献:

- [1] [美] Simon St. Laurent 著. 康晓林等译. XML 基础教程[M] 北京: 电子工业出版社, 2000. 3-37.
- [2] Sean McGrath 著, 潇湘工作室译. XML 应用实例[M]. 北京: 人民邮电出版社, 2000. 122-132
- [3] [美] Scot Johnson 著, 新智工作室译. Active Server Pages 详解[M]. 北京: 电子工业出版社, 1999. 367-475.
- [4] 孙一中. XML 理论和应用基础[M]. 北京: 北京邮电大学出版社, 2000. 141-148.

Publishing Database on Internet with XML

WANG Jing, FU Xiu-fen

(Faculty of Computer, GDUT, Guangzhou 510090, China)

Abstract: The paper generally introduces XML technology, with focuses on the method of using XML technology to publish database on Internet and explains the differences between this method and the traditional method of using ASP technology. An example of publishing database on Internet is shown. In the end a conclusion of the advantages of using XML to develop Web application is made.

Key words: XML; Internet; Database; ASP