

基于.NET的XBRL数据转换引擎研究与实现

葛永琪^{1,2}, 吕科²

(1. 宁夏大学 数学与计算机学院, 宁夏 银川 750021;

2. 中国科学院 研究生院 计算与通信工程学院, 北京 100049)

摘 要:针对目前基于XBRL统一数据集成化处理平台中从各种不同数据库读取及转换数据等方面的困难及不足,分析了XBRL技术的特点和数据转换需求,然后探讨了数据转换引擎的工作原理及其实现过程。该研究有利于数据快速整合,为统一数据集成化处理平台数据转换提供了依据。

关键词:XBRL; XML; .NET; 数据转换

中图分类号: TP311

文献标识码: A

文章编号: 1673-629X(2009)02-0148-03

Research and Realization of XBRL Data Transformation Engine Based on .NET

GE Yong-qi^{1,2}, LÜ Ke²

(1. Mathematics and Computer College, Ningxia University, Yinchuan 750021, China;

2. Coll. of Computer and Communication Eng., Graduate Sch., Chinese Academy of Sciences, Beijing 100049, China)

Abstract: Aim at the difficulty and shortage of XBRL reading and converting data from various different databases, it firstly analyzed the XBRL technical characteristics and the need of data convert in this paper, and then brought forward the work principle of the data transformation engine. This research is advantage to integrate data quickly, provides a basis for data transformation of unified data processing platform.

Key words: XBRL; XML; .NET; data transformation

0 引言

基于XBRL统一数据集成处理平台主要应用于业务报告的自动处理,为业务报告信息的准备、生成、分析和传输提供了便利,为企业节约了成本,提高了效率并且提升了在报告生成过程中的准确性和可靠性。其所处理的信息内容包括:金融、税务、证券、财务等相关行业,这就要求该平台必须从各种不同数据库读取数据,转换为XML,进而进行XBRL数据处理。

如果没有一个统一的数据转换引擎,数据的交换与共享将非常烦琐。因此,针对XBRL统一数据集成处理平台的这项需求,研究并实现了XBRL数据转换引擎。

1 XBRL技术与.NET技术

1.1 XBRL技术

XBRL是eXtensible Business Reporting Language的缩写,中文译作“通用业务报告语言”。它是XML(可扩展标记语言, eXtensible Markup Language)在业务报告信息交换方面的一种应用,是目前应用于非结构化信息处理,尤其是财务信息处理的最新技术^[1]。XBRL是一套公开免费的标准,这使得XBRL更容易成为业界的标准,现在已经逐渐成为业务数据报告领域的统一标准^[2]。

XBRL主要应用于业务报告的自动处理,为业务报告信息的准备、生成、分析和传输、比较提供了便利,XBRL为企业节约了成本,提高了效率并且提升了在报告生成交换过程中的正确性和可靠性,XBRL技术为整个信息供应链提供了一种理想的载体方案^[3]。XBRL可以处理不同语言和会计准则下的数据,它提供了足够的灵活性来满足不同的需求。数据可以通过影射工具转换成XBRL数据,也可以通过XBRL软件生成XBRL数据。

收稿日期: 2008-05-16

基金项目: 北京市自然科学基金(4073042); 中国科学院研究生院科研启动基金(2006010)

作者简介: 葛永琪(1980-),男,宁夏银川人,硕士研究生,主要研究方向为计算机软件与理论; 吕科,副教授,主要研究方向为XBRL技术与应用和数字图像处理。

1.2 .NET 技术

.NET 是一个可以用于创建 XML WEB 服务的平台,该平台将信息、设备和人以一种统一的、个性化的方式联系起来。借助于 .NET 平台,可以创建和使用基于 XML 的应用程序和 WEB 站点以及服务,它们之间可以按设计,在任何平台或智能设备上共享和组合信息,以向单位或个人提供定制好的解决方案。

.NET 框架下提供了一组高效的数据访问组件,总称为 ADO.NET,提供了对关系数据库和各种数据源的高效访问^[4]。.NET 内置了 XML 支持,可以处理非 Windows 平台导入或导出的数据。ADO.NET 附带四个数据库客户命名空间,一个用于 SQL Server,另一个用于 Oracle,第三个用于 ODBC 数据源,第四个用于通过 OLEDB 实现的数据库。如果数据库不是 SQL Server 或 Oracle,就应该使用 OLEDB,除非还能使用 ODBC^[5]。

2 引擎结构及数据流程

2.1 引擎的体系结构

该引擎的总体目标是利用 ADO.NET 技术读取来自不同的数据源的各种数据类型,并使读取的数据与 XML 节点生成对应关系,利用 .NET 技术对 XML 的良好支持,最终将读取的数据转换成满足初期 XBRL 需求的 XML 文件。其通过图形化的操作界面,使得用户直观,简单地就可以进行数据转换,提高了系统平台的简洁性和易用性。从而使得该平台更加易于操作,更加人性化。

该引擎主要针对六类数据库进行操作,包括 Access、SQL Server、Oracle、MySQL、DB2、Sybase,这六类数据库对 SQL 都具有良好的支持。要对数据库进行访问就必须对数据库建立连接,为此,针对这六类数据库,首先建立统一的数据库连接接口,根据用户设置的远程服务器 IP 地址及选择的数据库类型进行数据库连接验证(Access 属于桌面数据库,其远程连接需要进行额外的权限及共享设置),连接成功后,根据用户设

置的要转换的表,自动读取表中全部字段,并自动生成字段与 XML 节点关系映射设置会话。用户根据需要选择要转换的数据进行相关的设置。最终将数据转换成 XBRL 初期所需要的 XML 文件。

其总体体系结构如图 1 所示。

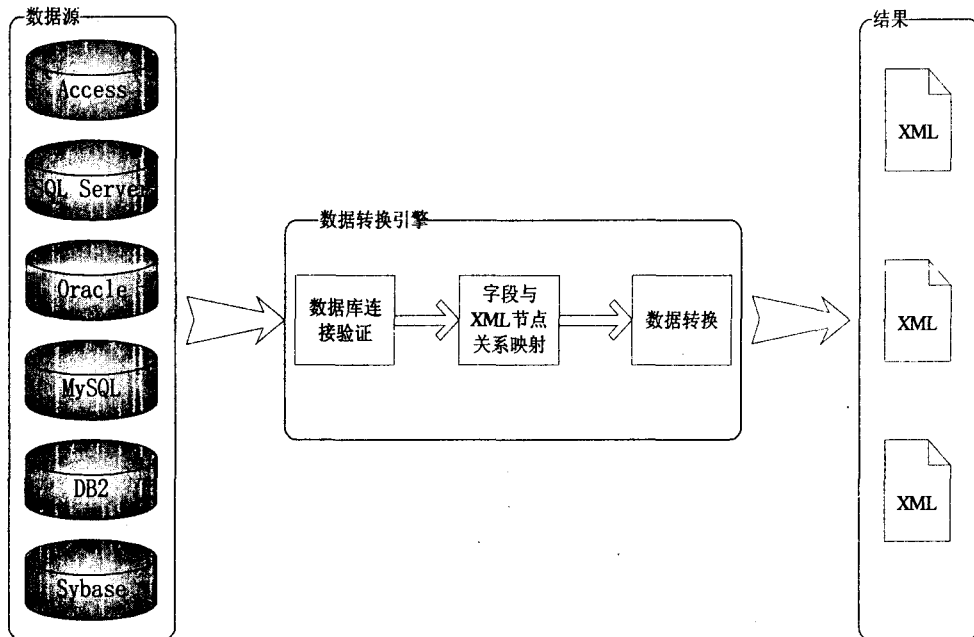


图1 引擎体系结构图

2.2 数据流程

经过对该引擎进行全面的研究总结后,首先按照不同数据源相应的数据库连接驱动,制定统一的数据库连接接口。该接口具有传入参数DT,该参数表示所进行连接数据库的类型。接口根据传入参数,选择相应的数据库驱动对远程数据库服务器进行连接,连接成功则进行下一步的操作,若失败则返回,重新设置后再尝试进行连接。连接成功后读取表中的各个字段,生成字段与XML节点关系映射对照表。如果不对字段与XML节点关系进行映射对照,直接生成的XML节点将与数据库字段保持一致,这将造成从各个数据库取出的数据不统一,不易于理解,也同时不满足XBRL初期对XML文件格式的要求。SQL标准提供了别名机制,这样可以很简单地进行字段映射,映射完成后,利用.NET内置的对XML的支持,将取出的数据存入数据集,然后再转换为XML文件。

其数据流程如图2所示。

3 数据转换引擎的实现

3.1 统一数据库接口实现

此接口的实现是以常见的六大关系型数据库作为用例数据库,包括 Access、SQL Server、Oracle、MySQL、DB2、Sybase,利用 .NET 框架下 ADO.NET 数据库访

问组件对以上数据库良好高效的支持,将连接各种不同数据库的复杂的逻辑封装起来,减少了并发冲突,并可以保证数据库以最高效率运行,将连接数据库逻辑从用户界面和代码分类出来,提高了易用性。

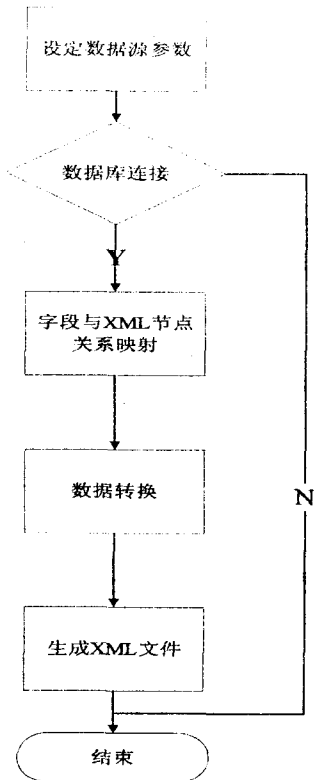


图2 数据流程图

首先通过 `private void setSqlconn(string DT)` 获取用户通过 WEB 界面设置的数据库信息,参数 DT 表示数据库的类型,以此来判断来引入相应的命名空间并设置相应的数据库连接驱动。

然后通过 `private bool testSqlconn(string DT)` 来判断数据库连接是否成功。

其设置界面如图3所示。

图3 数据库信息设置图

3.2 字段与XML节点映射关系实现

由于对生成的XML文件的节点名称有特殊要

求,XML文件的节点名称需要指定,所以该部分动态取出需要数据转换的表的字段,应用字段所对应的XML节点名称用户可以自由指定。

如果不设置XML节点名称,直接将查询结果转换为XML,那么XML节点对应的即是字段的名称,这不符合要求。SQL标准有别名机制,所以支持SQL标准的数据库都可以使用别名机制,这就为我们设置XML节点提供了便利。

其节点对照设置如图4所示。

表名:	Sample Company	查询字段	转换数据
		S_Cash	Cash Flow from(Used In) Operations
		S_Free	Free Cash Flow
		S_Num	Number of Employees at End of Period
		S_Profit	Profit(Loss)
		S_Profit_O	Profit(Loss) from Operations
		S_Sale	Sales
		S_Year	Year

图4 节点对照设置图

3.3 数据转换实现

ADO.NET使用XML进行远程传输,所以数据可以在不支持ADO.NET的应用程序和系统之间交换。因为XML在ADO.NET中非常重要,所以ADO.NET提供了一些强大的功能来读写XML文档。

本研究使用ADO.NET和XML把从不同数据库取出的数据存入DataSet,然后使用WriteXml()方法将结果转换成XML。首先创建SqlDataAdapter对象da和DataSet对象ds后,然后调用DataSet类的WriteXml()方法,它将生成一个XML文档。WriteXml()有两个

重载方法,一个重载方法的参数是带有文件路径和名称的字符串,另一个重载方法还有一个参数,即模式。其值可以是 IgnoreSchema, WriteSchema, DiffGram。

下列代码采用WriteSchema模式,让Schema以内联模式写入XML文件的开头。

SQL Server实例核心数据转换代码:

```
DataSet dsxml = new DataSet
("XMLDATA"); //建立 dataset
```

(下转第186页)

求,同时也保证了操作界面的简易。系统对空间位置关系大的目标处理能力强,可以框选或任意多边形选的方式选定目标,进而测算、分析目标的详细情况,最后以报表形式显示。

(3)图层的分类管理。

利用 ArcGIS Engine 能显示多个图层组成的地图的功能,将录入的空间数据归类后,以图层方式分层存放。可以在目标图层上进行增加、删除地块等编辑,或对加载在此图层上的数据库进行操作,而不对其他图层产生影响。图层之间可以进行排序、叠加。

(4)方便的数据更新和维护。

Visual c# 2005 和 SQL server2005 开发的系统能较好地提供数据维护、数据检查、数据备份、数据整理功能。数据维护界面友好,能直接对电子地图数据和数据库信息进行维护编辑,保障数据的准确可靠。相关的数据文件之间具有智能化的录入功能。

(5)良好的权限访问。

设置管理员、一般用户、参观者三级操作权限。不同用户只能在相应的权限下对系统进行访问,得到相应密级的系统资料,从而避免系统数据被盗窃和擅自修改。

5 结束语

由于 ArcGIS Engine 采用了嵌入式的开发技术,使得建立的林业资源信息管理系统能更好地兼容目前已

有的管理系统。利用 ArcGIS Engine 建立的林业资源信息管理系统不仅可以帮助管理部门全面系统地掌握林业资源各项最新情况,而且还能提供林业可持续发展做出决策科学依据。随着应用需求的提高,系统的功能将进一步地增强。最后衷心感谢方陆明教授和徐爱俊副教授提供的资料和帮助!

参考文献:

- [1] 邬伦,刘瑜,张晶.地理信息系统原理、方法和应用[M].北京:科学出版社,2001.
- [2] 肖海,武伟,刘洪斌.基于 ArcGIS Engine 的农业资源信息管理系统研究[J].计算机与现代化,2006(1):76-78.
- [3] Watson K, Nagel C. Beginning Visual C# 2005. C# 入门经典[M].第3版.齐立波译.北京:清华大学出版社,2006.
- [4] 魏安世,李伟,陈鑫,等.基于 ArcGIS Engine 的森林资源管理信息系统设计与开发[J].广东林业科技,2006,22(2):31-36.
- [5] 李佳田,刘洪斌,武伟,等.基于 SDE 的数据库应用研究[J].微机发展(现名:计算机技术与发展),2003,13(4):8-11.
- [6] 桂润堂,钟霞,薛重生.基于 ArcSDE 空间数据库引擎技术的应用研究[J].微机发展(现名:计算机技术与发展),2003,13(6):52-53.
- [7] 方陆明,唐丽华,徐爱俊.县级林业资源管理信息系统的结构研究与应用县级林业资源管理信息系统的结构研究与应用[J].浙江林学院学报,2005,22(3):249-254.

(上接第 150 页)

```
SqlConnection connxml = new SqlConnection(sqlconn);
SqlDataAdapter da = new SqlDataAdapter(tfsql, connxml);
da.Fill(dsxml, "Data");
dsxml.WriteXml(Server.MapPath("xml/") + tname + ".xml", XmlWriteMode.WriteSchema);
//保存 xml 文件
转换后的 XML 文件部分数据:
<Sample Company>
  <Cash Flow from (Used In) Operations>620,000</Cash
from (Used In) Operations>
  <Free Cash Flow>400,000</Free Cash Flow>
  <Number of Employees at End of Period>1,231</Number
of Employees at End of Period>
  <Profit(Loss)>554,000</Profit (Loss)>
  <Profit (Loss) from Operations>650,000</Profit (Loss)
from Operations>
  <Sales>1,300,000</Sales>
  <Year>2004</Year>
</Sample Company>
```

4 结束语

基于 .NET 的 XBRL 数据转换引擎为 XBRL 统一数据集成化处理平台访问不同数据源提供了便利,可以高效、迅速地完成了对远程数据库的访问及数据转换,降低了应用程序管理的复杂度,并相应地提高了利用率。本研究对于不同数据源快速访问及数据的快速转换具有一定的意义。

参考文献:

- [1] Ramin K. XBRL as a new language for business and intangibles reporting[J]. PRISMWP4,2002(9):2-4.
- [2] 吕科,刘晓峰.业务报告领域中的一场革命[N].金融时报,2006-05-24(9).
- [3] 吕科,刘晓峰. XBRL 技术原理与应用[M].北京:电子工业出版社,2007.
- [4] 谈政,蔡明.基于 XML 与 .NET Remoting 的分布异构数据转换[J].微计算机信息,2006,22(7):230-231.
- [5] Nagel C, Evjen B. C# 高级编程[M].第4版.北京:清华大学出版社,2006.