

文章编号: 2096-1472(2017)-02-47-02

集成异构数据源的在线考试系统研究

杜月云, 邢文凯

(郑州大学西亚斯国际学院电子信息工程学院, 河南 郑州 451150)

摘要: 本文针对异构环境下在线考试系统的集成提出的一种解决方案, 在不影响原系统的基础上, 构建一个集成平台, 实现异构数据库到XML格式文件的转换, 使得异构环境下的在线考试系统可以共享数据。重点论述集成异构在线考试系统设计与实现: 异构数据源集成的需求、系统功能模块设计、XML数据集成设计。包括包装器、查询处理模块、结构处理模块的设计, 实现由不同数据库到XML格式数据的转换集成, 实现了在线考试信息系统中异构数据的共享。

关键词: 集成; 异构数据源; 在线考试系统

中图分类号: TP311.52 **文献标识码:** A

A Study on the Online Examination System of the Integrated Heterogeneous Data Source

DU Yueyun, XING Wenkai

(Electronic Information Engineering College, Sias International University, Zhengzhou University, Zhengzhou 451150, China)

Abstract: This paper presents a solution to the integration of the online examination system in the heterogeneous environment. Without affecting the original system, an integrated platform is constructed to implement the conversion of heterogeneous database to the XML format file, and data can be shared on the online examination system in the heterogeneous environment. The paper focuses on the design and implementation of the integrated heterogeneous online examination system, such as the requirements of the heterogeneous data source, the design of the system function module, and the design of the XML data integration. Including the design of the wrapper, the query processing module and the structure processing module, the paper implements the conversion integration from different database to XML format data, and the sharing of heterogeneous data in the online examination system.

Keywords: integration; heterogeneous data source; online examination system

1 引言(Introduction)

目前, 我们已经进入了大数据时代, 大数据的典型特征就是海量的信息资源和复杂的异构信息资源。高校作为先进技术的开创者, 一直走在信息化发展的前列。但是伴随着信息化的发展和教育改革的不断深化, 传统的考试形式很难满足高校发展的需求, 于是形形色色的在线考试系统如雨后春笋般涌现。但是由于前期缺乏统一的规划, 导致在线考试系统存在结构化、相互之间不能兼容, 严重影响了在线考试系统的发展。如何集成异构的在线考试系统, 提高在线考试系统的效率成为高校需要迫切解决的关键问题之一。

2 异构数据集成技术^[1,2](Heterogeneous data integration technology)

随着对数据集成技术的日趋成熟, 现在既可以实现数据库中结构化数据的集成, 也能够处理文本文件等非结构化的数据; 不仅能够集成文本类型的常见数据, 也可以集成图片、声音、动画、视频等多媒体形式的数据; 经过数据集成的数据都具有完整性、一致性和集成性三大特征。随着XML等新技术的发展, 异构数据源集成技术的发展也从未停歇。

中间件模式通过中间件进行异构数据源的集成, 它无

需变动原有系统中的信息管理和存储方法。该模式下中间件类似于一个服务接口, 它位于异构数据源和信息管理系统之间。对于异构数据中间件可以协调他们之间的关系, 使异构数据库中的数据转换成标准的数据进行数据交换; 对于各个信息管理系统的用户中间件能够提供一个统一的访问接口供用户使用。其体系结构如图1所示。

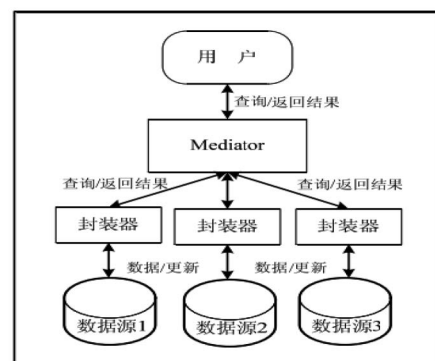


图1 异构数据源集成体系结构

Fig.1 Heterogeneous data source integration architecture

集成(Mediation)模式是中间件体系结构的核心^[3]。采用该模式我们能够将异构数据源的各种数据信息通过专用的封装

器进行相互转换,转换后的数据具有统一的格式,可以实现数据交换和资源共享。用户在该模式下可以实现跨异构数据源的信息查询功能,其实现过程大致如下:首先系统将用户的请求自动转换成标准的格式,再将这些标准的格式找到对应的封装器,利用封装器将其转化为其对应的数据源能够识别的查询语句,数据源经过相应的查询处理后,将查询结果交给包装器进行提取,最后通过集成处理后将最终结果反馈给发出查询请求的用户,普通用户查询时非常方便,几乎感觉不到使用的是异构数据源。

3 系统的设计与实现(System design and implementation)

3.1 异构数据源集成的需求

目前,高校自建了很多形式多样的在线考试系统,在为高校考试信息化改革做出巨大贡献的同时也存在很多问题,如各系统使用的数据库多种多样,版本也各不相同,导致相互之间难以兼容,给高校考务工作带来了很大不便,为了提高考务管理的效率,必须对异构的数据进行集成。主要体现在几个方面:形成考试系统全局视图;统一的查询模式;实现关联查询;保持原有数据库独立和完整;实现跨系统的集成查询。

3.2 系统功能模块设计

本课题针对高校目前存在的体系结构异化的在线考试系统进行集成,首先要对目前正在使用的在线考试系统进行分析,发现很多是学院教师结合自己的技术和在线考试的需求自行开发,前期并没有进行深入的功能模块划分和UML建模,这就加大了数据集成的难度。为了更好的进行数据集成,我们对所有的在线考试系统进行分析,系统主要有学生、教师和系统管理员三类用户,按照功能具体可以分为题库管理、试卷管理、成绩管理、用户管理等11个功能模块。

3.3 XML数据集成设计

首先,为所有异构数据源设计自己特有的包装器,它可以对异构数据进行处理转化为能够被识别的统一格式数据,并通过XML形式进行数据访问。查询处理器主要完成数据库中最常用的查询操作,比如查询学生的成绩。本文设计的查询处理器能够对查询进行分解,将子查询根据实际情况发放到各个异构数据库中。结果处理模块能够把查询处理器得到的查询结果进行XML转换,并将相关内容进行合并处理为一个标准化的XML文档,即是把查询结果以XML形式返回给相应的应用软件。

(1)包装器

包装器(Wrapper)处于逻辑层的最底层,它负责直接抽取数据源中的数据,是XML进行数据集成的必要组件。一般情况下,所有数据源都需有一个与之对应的包装器负责数据的提取工作。不管数据源是那种结构化的关系型数据库,还是非结构化的文本数据都要有一个包装器使用XML将其进行

包装,在包装过程中不能影响原有数据源状态,使得包装器和数据源直接一直保持相互映射的关系。经过包装处理后的XML信息一般不能用于查询条件复杂苛刻的数据,只能进行一系列简单的选择。如果必须进行详细的信息查询,需要XML和Xquery协同构建一个针对数据源能够识别的SQL查询语句,然后使用该语句进行查询,当查询完成后需要将结果转换为标准的XML数据模型供其他模块使用。

包装器还需要解决各种类型数据源与局部视图之间的双向映射。文中将使用XSL对用户视图定义,也就是用XML来解释集成数据,用XML Schema表示集成模式中数据源间的映射。由相应的包装器把各数据源的本地视图就全都溶入到全局视图了。期间,为了方便对XML文档进行定位和解析,包装器将数据源中的数据转换为DOM对象^[4]。

关系数据到XML文档的映射规则:列映射为属性,表映射为元素^[5]。属性标记及元素标记都是自定义的,在Schema文件中都存放相关的描述,用来表示从关系表中所获取的字段信息^[6]。同时经过使用属性标记和元素标记,从而使得XML的文档内容具备了可识别性。

(2)查询分配器设计

集成异构数据源的统一查询系统能够极大方便用户对总体视图进行查询,但要在后台完成对相应物理数据源的查询还需要高效的查询分析器来完成^[7]。查询分析器要能够将查询及时、高效并精准的发送给物理数据源。高校在线考试系统异构数据源的数据集成会产生海量的数据,并且有很多数据直接来源于实际数据库,需要进行数据处理的量还会进一步增长。基于以上情况,我们需要尽早进行权限的识别,以减少取出数据后再判断产生的大量冗余数据,影响系统运行效率。

(3)查询转换器设计

查询分配器分配到数据源的查询采用的是统一的XML格式,而底层的数据源一般无法直接识别需要将其用SQL语句表示出来,这项工作由查询转换器负责完成。查询转换器的工作和包装器工作正好相反,其重点在于如何将XML文档信息转化为SQL语句。本文设计的转换器能够将XML文档中的相关查询请求信息提取出来,并且结合底层数据源生成对应的SQL语句。

(4)结果集成器的设计

查询返回的信息不能被客户端的浏览器直接识别,因此,需要我们对返回的查询结果进行处理,使其转化为浏览器能够处理的标准数据,这一过程的实现需要结果集成器来完成。集成异构数据源的在线考试系统要求能够实现通用查询,并反馈全局数据,这就涉及到针对某一个数据源的局部模式和对整个系统的全局模式进行集成转化,这个过程我们分三个过程来完成,首先进行局部模式的比较,充分认识局部模式的异同。其次进行局部模式的调整,确保系统中模

式的统一。最后进行局部模式的集成,从而实现全局模式^[8]。

(5)结果转换器设计

结果转换器的作用是转换异构数据源提供的子查询结果为XML格式,把结果数据提供给结果集成器。结果数据模式转换为XML Schema是设计结果转换器过程中重点要解决的问题。只有成功的构造了XML Schema,每个子查询的结果以及含义才能无误地提交给结果集成器,从而确保结果数据集成的顺利完成。

结果数据模式转换成为XML Schema的算法是这样定义的:可以认为异构数据源中的查询结果体现为表的形式,结果表的第一行的每一列指定本列查询的结果字段,结果表的第二行以下的数据指定每一个查询字段所取的值。

4 结论(Conclusion)

本文通过研究传统的异构数据源集成技术,以及XML技术,设计了一种基于XML的异构数据库集成的方案。通过中间件技术建立XML数据集成中间件,并通过中间件来完成对异构数据源的查询、访问及结果合成。针对高校在线考试系统中普遍存在异构数据源集成问题进行研究,提出了针对这些异构数据做的集成方案,提出了系统中XML数据集成的设计方法,包括包装器、查询处理模块、结构处理模块的设计,实现由不同数据库到XML格式数据的转换集成,解决在线考试信息系统中异构数据共享的问题。

(上接第56页)

选用适用于餐厨废弃物收运业务的数据挖掘技术并实现相应的数据挖掘软件工具,对系统中存储的海量业务数据进行分析与处理,为餐厨废弃物的产生与收运管理提供辅助决策支持。

4 结论(Conclusion)

本文阐述了基于物联网技术餐厨废弃物回收处理监控系统的研究,研究成果可快速形成产业规模并产生以下社会效益:

(1)改善市容环境

从餐厨垃圾产生源头至收运作业实现全过程多角度监管,有效杜绝随意倾倒、不密闭车厢、抛洒滴漏等影响市容的现象,确保餐厨垃圾不会污染道路环境、民众居住环境,保障市容环境的干净整洁。

(2)遏制“地沟油”的黑色价值链

对餐厨垃圾的收集、运输过程进行监管,实现前端收集量和运输路线的实时监控;与环卫监管平台对接,实时对相关监管数据进行实时采集和上传;与卫生、环保等管理部门平台对接,实现对餐厨垃圾各个环节的监管。杜绝了将餐厨垃圾作为牲畜饲料、地沟油,再进入食物链影响人体健康的危险因素。

参考文献(References)

- [1] Knowledge-Based Wrapper Generation by Using XML. H.Seo,J.Yang,J.Choi.IJCAI-2001 Workshop on Adaptive Text Extraction and Mining(ATEM2001),2001.
- [2] VODAK Kernel Data Model.H.Duchene,M.Kaul,Turau. Proceedings of the 6'0 International Conference on Data Engineering,1999.
- [3] 于帆,王振铎,王振辉.基于XML异构数据库集成中间件的设计与实现[J].计算机应用研究,2007(09):185-188.
- [4] 王星,江志农.基于XML的中间件技术在故障诊断数据集成中的应用[J].机械制造与自动化,2008(05):112-114;122.
- [5] 闵道辉.基于XML的异构数据源集成研究与应用[D].西南石油大学,2006.
- [6] 周运,牟占生,徐久成.基于XML虚拟数据库的异构数据源集成模型研究[J].计算机技术与发展,2008(04):84-91.
- [7] 洪筱菡.基于XML的数据集成中间件研究[D].河海大学,2003.
- [8] 欧玉平.基于XML的机关办公信息系统异构数据库集成设计与实现[D].电子科技大学,2012.

作者简介:

杜月云(1975-),女,学士,副教授.研究领域:计算机软件教学与研究.

邢文凯(1973-),男,硕士,副教授.研究领域:计算机软件教学与研究.

(3)提升管理部门的管理水平

信息化助推餐厨垃圾监管,城市市容和周边环境明显改善,居民正常生活得到保障,管理部门的重复清运成本明显降低,市民对管理部门工作满意度提高。

参考文献(References)

- [1] 廖神海,姜承芝,田浩.基于物联网平台的太阳能热水工程监控系统[J].太阳能学报,2016,37(7):1851-1856.
- [2] 颜波,石平.基于物联网的水产养殖智能化监控系统[J].农业机械学报,2014,7(1):58-65.
- [3] 廖建尚.基于物联网的温室大棚环境监控系统设计方法[J].农业工程学报,2016,32(11):233-243.
- [4] 练胜,王栋,朱祥贤.基于物联网技术的电梯安全监控系统的设计与实现[J].机电信息,2015(03):126-127.
- [5] 徐晓新,等.基于校园网的物联网的建设和管理方法[J].工业仪表与自动化装置,2016(04):109-110.
- [6] 王家利,等.基于物联网的设备监管系统设计[J].山西电子技术,2015(02):53-55.

作者简介:

周子立(1963-),男,本科,高级工程师,副教授.研究领域:计算机技术应用.