

## Heterogeneous Data Exchange Technology Based on XML

Jia-hua Zhao

Department of Computer Science and Technology,  
Cheng-Dong College of Northeast Agricultural University  
Harbin, 150025, China  
jiahua927@163.com

**Abstract**—Database technology is of great variety, structure diversity, solving data integration and exchanges sharing of the heterogeneous database system have become the important problem in all fields. Because of XML is the very important data exchange technology to realize data exchange and transmission ability, this paper mainly explores the data exchange problem how to use XML in heterogeneous database system, , and .Net platform have good support to XML and heterogeneous database system, so the combination of XML and.net, can realize data exchange of heterogeneous database system.

**Keywords**-component heterogeneous database; data exchange technique;XML

### I. INTRODUCTION

Data exchange between heterogeneous databases is mainly to eliminate the conflict between heterogeneous data. In the different application platform and operating system, to exchange both data can be achieved between each of the transparent access, in order to achieve data sharing purposes. Heterogeneous data exchange is the critical pathways of implementation data sharing of systems, business collaboration and solving "information isolated island" phenomenon. At present, the common exchange techniques of heterogeneous data can be divided into three types: many to one conversion method, data warehouse method, bidirectional middleware method. But the traditional data exchange of heterogeneous database is tedious, time-consuming, and involves a lot of manual coding; poor data resource reuse and transplanted; lack of flexibility[1-2].

In addition, there are a lot of researches at home and abroad. The United States University of California Dr.Lee proposed NET algorithm and COT algorithm. The two algorithms is the goal for the DTD mode conversion and completed in two steps, but lack of maintain the integrity of data constraint. Home also studied using DTD document to directly realize the method of conversing the XML document to relational database. This study use a data structure to replace the object tree, but it is the precision lack of complex XML documents information conversion. The middleware and database companies have also introduced the products of realizing data exchange between XML and database[3].

### II. DATA EXCHANGE OF HETEROGENEOUS DATABASE BASED ON XML TECHNOLOGY

XML has the formidable data performance ability, strong ability of self description and expandability, while the realization of the separation of data and the form of expression. These characteristics decided that XML can be used as media of data exchange, thus realizing the information exchange of heterogeneous system. The study of data exchange of heterogeneous database based on XML by nature is to actually study the mutual conversion between the XML model and relational database schema. This paper through establishing a mapping relationship between the XML document model and database model realize data conversion between them[4-5].

#### A. Exchange rules analysis module design

Exchange rules: first to open source data table through the given database connection information, to select the fields that need to be converted in the source data table, and then to select the target database corresponding field, final object to save the source data table and the target field mapping information in the HashMap. Figure 1 shows the data exchange rules analysis module[6].

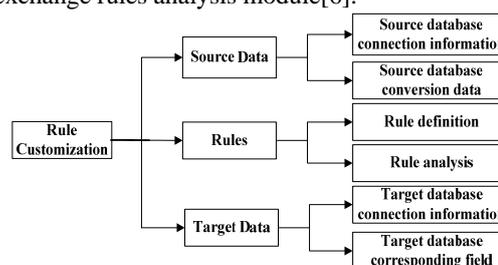


Figure 1. The Data Exchange Rules Analysis Module

#### B. Mapping file generation

To realize the heterogeneous data exchange will have to put the database structure mapping become the structure of an XML document, or will XML document structure mapping for the structure of the database, can transfer data between XML documents and the database. We must follow the principle of "Complete Map" in order to achieve from relational database to XML document mapping. Complete mapping means that the relationship between the relation tables fully reflected in the XML documents, when the

database or more tables to XML "complete map", we should mainly consider from three aspects. First turning out each data in the table, then rolling out communication between each table, and finally rolling out relation model. Figure 2 shows the data conversion and write program[7].

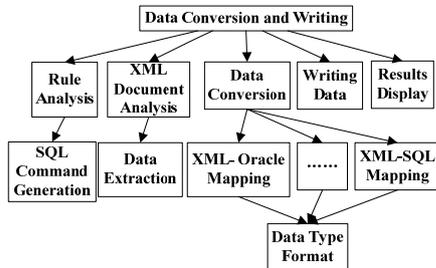


Figure 2. Data Transfer and Write Program Chart

1)Generating XML Schema format documents

This module mainly complete custom data extraction rules: the source database which information data needed convert into a target database information, the rule definition will save document in extraction exchange rules, for the user to select calls, and use XML file save the corresponding various transformation rule[8].

Application must first establish connection between the background service of application layer and database, will derive specified data table in the database or the data of view, then query the database, extract the corresponding data from the database, and background service is responsible for the data organization into a tree node and stored in the document with XML format. Design that mapping from the concrete the relational database table and its relationship to the XML Schema document is as follows: 1) Creating empty XML document; 2) Generating a root node; 3) Definition contains record number N; 4) Reading records from the data table, and then adding control description information in the corresponding node[9].

In the transition from relational database to transform XML documents, we need preserved some of the source database structure information with XML document format. Therefore, it is very important and key problems to establish a correct mapping relations data and pattern mapping relationship between database and XML document. Specific conversion can shows as table 1.

TABLE 1 XML SCHEMA AND DATABASE COMMAND CONVERSION TABLE

Database Explain	XML Schema
Description Section:It includes the identification, namespace, etc	<xs:schema...>
Element Part: Four table field in the database were mapped to elements and child elements of the XML document	<xs:element...>
Specifies the order of appearance of the sub-elements	<xs:sequence...>
Element uniqueness (can't have repeated value), "msdata:Primarykey =" indicates that the elements as the primary key (not empty), these elements	<xs:unique...>

correspond to the field on behalf of the primary key in the database table and non-empty.	
check Constraint	<xs:restriction...>...</xs:restriction>
Field Type	base="..."
Field Length	<xs:maxlength value="10">
IN Value	<xs:enumeration value="...">/>
Element binding provisions, which corresponds to foreign key constraint relations in the database table	<xs:keyref...>
Primary Key	<key...></key>
Column Names	<field xpath="...">

2)Extract data from an XML document

The data in the database to generate XML documents, when using the data, we need to parse an XML document and extract data from documents with XML + DOM parsing method, using XML Object Model (XML DOM) object which provided by Microsoft. We can read XML document by creating "Xml Node Reader" object. "Xml Node Reader" object is mainly used to read XML node data, and the attribute of "NodeType" can be used to judge read the node type, attribute of "Value" means the value of the node. Partially achieved code that read XML file and display the ListView is as follows:

```

while (XMLRreader.Read())
{
    // Judge the node type of current reading made
    switch (XMLRreader.NodeType)
    {
        case XmlNodeType.Element :
            name = XMLRreader.Name ;
            break ;
        case XmlNodeType.Text :
            if (name.Equals ("Name"))
                xmlItem = Listview1.Items.Add (XMLRreader.Value) ;
            else
                xmlItem.SubItems.Add(XMLRreader.Value) ;
            break ;
    }
}
    
```

III. HETEROGENEOUS DATA EXCHANGE .NET IMPLEMENTATION BASED ON XML

.NET is Microsoft's platform designed to solve enterprise-class applications, it is a framework set, which contains a variety of techniques, and which can be used to achieve the technology of XML, Web Services, SOA and agility.

.Net Framework is widely support for the use of the XML document, XML in the .NET Framework plays an important role, which makes cross-platform development applications easier. .Net support for XML mainly includes: XML Namespaces, XML Designer, XML Web Servers, as well as XML DOC. In the Microsoft .NET, database access technology is ADO.NET. In ADO.NET can be easily create and use XML data by XMLReader, XMLWriter,

XMLNavigator, XmlDocument, and support the W3C XSLT, DTD, XDR standard.

Ticket information management system software is designed based on the above mapping conversion mode. Ticketing system requires forwarding and sharing of data between multiple management servers, and management servers' data source may not be the same, so in the realization of the system you can use XML as the medium for data exchange. There are a large number of heterogeneous data in the ticket information management system, for example, heterogeneous database exchange of two aspects of the ticketing system and financial settlement system. Ticketing system produce corresponding ticketing information, this time ticketing system have the ticket information data mapped for XML format and sent to the ticket processing system platform, then converse ticket information format, finally mapping database which is used in financial settlement site, and processing this information in the financial settlement. In general, we achieved information sharing between two heterogeneous databases through XML middleware, and use .NET platform development.

In this instance ticketing information system uses SQL Server databases, but financial settlement management system using Oracle database, and using ADO.NET technology access the database. Figure 3 shows a heterogeneous database mapping module.

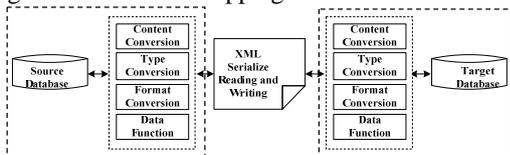


Figure 3. Heterogeneous Database Mapping Module

After user have selected requiring the conversion of data through the system interface, system generated the SQL command which is needed by data extraction through data analysis. According to the source database related information and analysis of the SQL command to connect to database source, and the implementation of the corresponding SQL command in order to complete the data extraction.

To read the TicketBookInfo table to generate XML tree node as example, the specific implementation code is as follows:

```

XmlDocument xdoc = new XmlDocument();
XmlElement ticketInfo = xdoc.CreateElement(TableName)
// Determination of the content of XML documents
XmlDeclaration xd = xdoc.CreateXmlDeclaration("1.0", "GB2312", "no");
xdoc.AppendChild(xd);
xdoc.AppendChild(ticketInfo);
while (dr.Read())
{
    for (i = 0; i < dr.FieldCount; i++)
    {
        XmlElement node = xdoc.CreateElement(dr.GetName(i).ToString());
        ..... // Obtain column attribute
        XmlElement childNode = xdoc.CreateElement(dr[i].ToString());
        ..... // Obtain fields constraint
        node.AppendChild(childNode);
        ticketInfo.AppendChild(node);
    }
}

```

```

}
}
xdoc.Save("C:\\Booking\\TicketBooking_Data.xml");
}
}

```

The following example, based on the order information in the ticketing system database table, XML data file generated according to the mapping rules:

```

<?xml version="1.0" encoding="ISO_8859-1" ?>
-<TicketSystemDB>
-<OrderInfoTable>
    <tsid>K3286</tsid>
    <uid>100001</uid>
    <upir>120.00</upir>
    <bnum>3</bnum>
    <tpir>360.00</bnum>
-</ OrderInfoTable>
-<OrderInfoTable>
    <tsid>T20</tsid>
    <uid>100002</uid>
    <upir>80.00</upir>
    <bnum>5</bnum>
    <tpir>400.00</bnum>
-</ OrderInfoTable>

```

In the XML data documents through the analysis and validation, we need to consider the XML document is mapped to the Oracle database. This mapping is not directly realized, so we can't put the XML data document directly map to the library, but must first transform ADO.NET provided by .NET platform. First XML document data is loaded into the DataSet, and then mapped to the database. ADO.NET technology can remotely transmit through XML, so the data can be exchanged between applications and systems which don't support ADO.NET.

ADO.NET provides powerful function to read and write XML document. Use ADO.NET relational data class is included in the namespace of System.Xml. If the XML document is loaded into the Oracle database, the first to transform XML documents into ADO.NET DataSet. The main code is as follows:

```

DataSet dataSet_xml=new DataSet("TICKETXML");
dataSet_xml.ReadXml("C:\\Booking\\TicketBooking_Data.xml",Xml
ReadMode.ReadSchema);

```

The ReadXml function is mainly used to deal with XML data, we select ReadSchema mode, namely to XML Schema as description language of XML data document model. It includes two parameters, which is used to read the XML data documents and settings read mode.

Although we have converted the XML data document into data in the dataSet\_xml, but we need a data adapter to directly update the database, otherwise it is impossible to directly update. Therefore, we need create a new DataSet\_Server, and make the data set and database connection through the data adapter DataAdapter. This dataset directly update the database, .NET is provided Merger function which is used to merge the two data sets with the same data model, Because the data model of dataSet\_xml and DataSet\_Server is the same, that they are XML schema format, we can keep them "merge", and finally through the DataAdapter to update database.

#### IV. CONCLUSION

So far, the heterogeneous database based on XML between information exchanges has all come true. Seen from the implementation process of the entire heterogeneous database exchange, XML as a data exchange middleware can well achieve the conversion of data between the various databases, so as to achieve the purpose of data sharing and information exchange.

#### REFERENCES

- [1] Zi-yao Wang, Jun-jie Nan and Zi-hui Duan "SOA Core technology and applications ",Bei-jing, Publishing House of Electronics Industry,2008,PP.34-39
- [2] Hao Shaohua, Han Xie. Heterogeneous Relational Database Integration Model Based on XML Technology.Computer Engineering and Design, 2010,31 (24), pp. 5285-5288
- [3] Andreas M. Weiner, Theo Härder. An integrative approach to query optimization in native XML database management systems.IDEAS '10 Proceedings of the Fourteenth International Database Engineering & Applications Symposium.2010,3, pp. 64-74
- [4] Ji Zhaohui, Hu Kongfa. Based on ADO.NET and XML to Realize Data Exchange.Journal of Micro Computer Information. 2009,25(6-3), pp. 265-266
- [5] Hu Nengfa, Tang Weiping. Universal Heterogeneous Data Exchange Model Based on XML .Computer Engineering and Design. 2010,31 (8). PP: 1743-1745
- [6] Elli Tyrkkö, Anna Pelander, Ilkka Ojanperä. Differentiation of structural isomers in a target drug database by LC/Q-TOFMS using fragmentation prediction.Drug Testing and Analysis. Drug Testing and Analysis. 2010, (6):259–270
- [7] Geng Biao, Song Yuqing, Liang Chengquan and Chen Jianmei. XML Document to The Relational Database Mapping Method. Computer Application Research. 2010,27(3). PP: 951-955
- [8] Zhang Wei.Brief analysis of format standard of isomerism data converting.Electronics,Communications and Control (ICECC), International Conference on.2011,9(11):1925 – 1927
- [9] Shi Yeling, Li Jianhui. Relational Database Schema to XML Schema Generic Mapping Model. Computer Engineering. 2009,35(7). pp: 35-51