

JENA 在 CIM 模型中的应用研究

樊荣, 杨峰, 左吉昌

(华中科技大学水电与数字化工程中心, 湖北 武汉 430074)

摘要: 通用信息模型 CIM 提供了用于不同 EMS 系统间交互数据的一个通用模型, 在基于 CIM 的 EMS 系统间交互信息是以 CIM/XML 为载体, 而 CIM/XML 数据本质上是基于 CIM 语义的 RDF 数据。在开发基于 CIM 模型的 EMS 系统中, 通过引入基于 RDF 的 JENA 技术来处理 CIM/XML, 着重讨论了 JENA 对 RDF 数据的表达、解析、查询技术在 CIM 中的应用。JENA 不仅能自动地生成 CIM/XML 数据, 而且可以高效地对 CIM/XML 进行查询从而可以进一步实现 CIS 接口。JENA 技术为 IEC61970 标准的进一步的实用提供了一条新的有效途径。

关键词: 电力系统; IEC61970; CIM; JENA; RDF

Application and research of the JENA in CIM

FAN Rong, YANG Feng, ZUO Ji-chang

(Digital Engineering Research Center, Huazhong University of Science & Technology, Wuhan 430074, China)

Abstract: CIM (Common Information Model) provides a common model for exchanging in different EMS, CIM/XML as a carrier of EMS systems using CIM is based on RDF technology essentially. In the development of EMS, JENA technology that focuses on RDF has been introduced to process CIM/XML. JENA can automatically produce CIM/XML data, and highly effective query helping to implementation of CIS. JENA technology provides a new and effective way to practical usage for IEC 61970 standard.

Key words: power system; IEC61970; CIM; JENA; RDF

中图分类号: TM76; TM73

文献标识码: A

文章编号: 1003-4897(2007)04-0044-05

0 引言

通用信息模型 CIM(Common information model) 是一个抽象模型, 它表示包含在 EMS 信息模型中的电力企业的所有主要对象。通过提供一种用对象类和属性及他们之间的关系来表示电力系统资源的标准方法^[1]。具体地说, CIM 给电力系统中所有的实体对象提出了一个标准化的定义, 把所有这些实体对象抽象成“类”来进行描述, 进而又把这些类分别归入了不同的“包”。CIM 对主要的电网设备和相关对象进行了建模, 包含这些对象的公共类和属性, 以及它们之间的关系。可见, CIM 提供了一种通用的电力系统模型来对电力系统资源进行统一描述, 进而在接口层上实现数据交换和数据共享。

由于目前运行的各 EMS 系统的数据模型各不相同, EMS 系统无法读取、操作其它 EMS 系统的数据, 从而使得各个 EMS 系统成为单个的信息孤岛。CIM 模型的提出正是提供了这样一个通用的电力系统信息模型, 因此不同的系统要数据共享、相互操作,

就要以这种通用的模型来相互交换数据。导出就是把某一 EMS 系统的数据导出形成 CIM/XML 格式的数据, 传统的导入导出方法通常使用基于 XML 的 DOM 技术来解析 CIM/XML 文件, 这往往需要手工编写 XML 的 DTD 文件作为解析规则, 过程繁杂且容易出错, 缺乏灵活性, 并难以实现复杂的查询。CIM/XML 实际上就是使用 RDF 和 RDF Schema 来组织 CIM 模型的 XML 文档结构。本文通过引入基于 RDF 的 JENA 技术, 研究并实现了传统 EMS 关系数据库 CIM 导出与 CIM 数据的查询, 对实现过程中的一些关键技术问题进行了着重的阐述, 并讨论了处理 CIM 数据过程中, 基于 RDF 的 JENA 技术相对于传统基于 XML 技术的优点, 为进一步通过 JENA 来实现在 CIM 中的应用提供了参考。

1 CIM、RDF 与 JENA

1.1 资源描述框架 RDF (Resource Description Framework)

资源描述框架 RDF 提出了一个简单的模型用来

表示任意类型的数据。RDF是一种基于XML的应用,其基本思想是:通过简单二元关系表达的对资源陈述,其中每个陈述都是由主体(subject),谓词(predicate),客体(object)组成的^[2]。在RDF中,如下的英文陈述:

http://www.example.org/index.html has a
creation-date whose value is August 16, 1999

主体是URL http://www.example.org/index.html
谓词是词“creation-date”

客体是短语“August 16, 1999”

可以通过XML进行表示如下:

```
<?xml version="1.0"?>
<rdf:RDF
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:exterms="http://www.example.org/terms/">
<rdf:Description
rdf:about="http://www.example.org/index.html">
<exterms:creation-date>August 16, 1999</exterms:creation-date>
</rdf:Description>
</rdf:RDF>
```

RDF模型可以方便地描述对象(或者资源)以及它们之间关系。RDF的数据模型实质上是一种二元关系的表达,由于任何复杂的关系都可以分解为多个简单的二元关系,因此RDF的数据模型可以作为其他任何复杂关系模型的基础模型。

1.2 CIM的提出与RDF的关系

为了使电力系统模型间能够进行信息交换,电业部门需要就电力系统的实体以及实体间关系的定义达成一致。电力能源研究协会(EPRI)开发了一个通用信息模型CIM来完成这一任务。在CIM模型中,明确了电力系统的资源、属性以及关系等的通用语义。CIM模型也通过了国际电工委员会的国际化认证过程,并被加入IEC61970标准的30x部分,成为IEC61970标准对电力系统建模的标准模型。为了能够进一步增强电子化地交换CIM模型的能力,电力业已经开发了一种IEC61970,规定了用XML表达CIM模型的方法即CIM/XML。CIM/XML是一种RDF的应用,使用RDF和RDF Schema来组织CIM模型的XML文档结构。

CIM模型被明确表示为一系列使用统一建模语言(http://www.uml.org/)UML(Unified Modeling Language)描述的类图。CIM模型的基类是PowerSystemResource类,同时CIM模型还包括了Substation、Switch、Breaker等一些特殊化的子类。CIM/XML将CIM表示为RDF模式的一个词汇表,并

将RDF/XML作为交换特定系统模型间的语言。本质上说,CIM/XML数据是基于CIM语义的RDF数据。

1.3 JENA

既然CIM/XML本质上是基于RDF技术的,就可以直接使用基于RDF工具来对CIM/XML进行解析。Jena就是一个由HP实验室开发的表示和处理半结构化数据(主要是基于RDF的管理、查询等)的Java开源工具。目前是Jena2版本,Jena1做到了支持RDF的数据表达、解析、查询(采用RDQL)和简单的推理,Jena2的目标在支持本体(包括DAML、RDFS和OWL的应用程序接口API)。Java程序员可以方便地使用JENA来创建和填充RDF模型,并且能使用JENA的RDQL查询语言以程序方式查询这些模型。

2 使用JENA从EMS数据库导出CIM/XML文件

为了实现应用系统之间的数据交换,CIM采用RDF来实现电力系统中各个对象的描述,这就需要有一个描述的模式Schema,也就是所说的CIM RDF Schema,同时这个模式要与CIM的UML定义相匹配,这就需要有一个从UML模型生成一个Schema的规则。这个规则在IEC 61970的第501部分定义。CIM RDF Schema文件可以通过Xpetal工具从Ritonal Rose的mdl模型文件来生成。

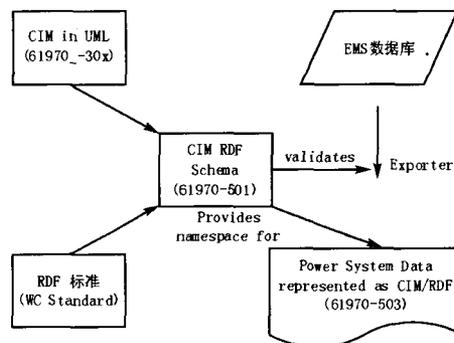


图1 EMS数据库导出符合CIM标准的CIM/XML文档的过程

Fig.1 XML-based IEDC-related data exchange

有了给定的CIM RDF Schema,一个电力系统EMS就能转换成一个导出的RDF文档,这个文档就被称为CIM/XML文档。在CIM/XML文档中所使用的所有标记由CIM RDF Schema定义。这个文档能够被解析,其中的信息能够被导入到别的系统中。图1描述了如何利用从EMS数据库导出符合CIM标准的CIM/XML文档的过程。注意是通过UML描述的CIM模型(IEC61970-30x部分描述)与RDF标准(W3C

给出的标准) 来得到 CIM RDF SCHEMA。由 CIM RDF SCHEMA 来规范 EMS 导出的 RDF 文件的格式,使其能以 CIM 标准来正确表示 EMS 数据库的数据。

通过 JENA 生成 RDF 文件就是创建与填充 RDF 模型,而使用 JENA 来实现 EMS 数据库导出 CIM/XML 文件就是创建和填充基于 CIM 的 RDF 模型。这个过程主要分为两个过程:①首先必须建立的是基于 CIM 模型的 JENA 词汇来保证基于该词汇生成的 RDF 文件是符合 CIM 标准的。②然后读取 EMS 数据库,并将数据通过 CIM JENA 词汇来创建与填充基于 CIM 的模型从而生成相应的 CIM/XML 文件。

2.1 从 CIM RDF SCHEMA 生成 CIM JENA 词汇

CIM 对电力系统中主要的电网设备和相关实体进行了建模,将其分为不同的对象与属性。CIM 标准就包含这些对象的公共类和属性,以及它们之间的关系。CIM 标准是通过 UML 语言来描述的一个 mdl 文件。只有符合 CIM 标准定义的 RDF 文件才可以称为 CIM/XML 文件。CIM 标准描述的开关 Switch 包括如下的属性 normalOpen、switchOnCount、switchOnDate 等。这些属性一般都可以从 EMS 数据库中得到,并且体现在导出的 CIM/XML 文件中。如果有些 CIM 标准中定义的属性 EMS 数据库中不存在则导出的 RDF 文件可以没有该属性,但需要注意的是导出的 RDF 文件的属性名称必须严格与 CIM 标准保持一致,以此来确保 CIM 的兼容性。

由于导出的 CIM/XML 文件必须满足 CIM 标准, JENA 需要得到描述 CIM 标准的词汇。通过该词汇导出的 RDF 必然满足 CIM 标准。JENA 提供了 Schemagen 工具从 RDF SCHEMA 来生成 JENA 词汇。而 CIM 的 RDF SCHEMA 又可以通过 Xpetal 工具从 mdl 模型文件来得到。图 2 描述了从 CIM 标准生成 JAVA CIM 词汇的过程。

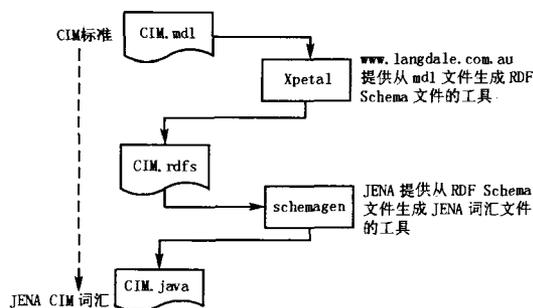


图 2 从 CIM 标准生成 JEANCIM 词汇

Fig.2 Conversion of JAVA schema from CIM mdl file

下面是 JENA CIM 词汇关于 Switch 属性的部分。

JENA CIM 词汇完整地表现了 CIM 标准在 CIM.mdl 的对象的各个属性,以及对象间的相互关系。

```

Public static final Property Switch_normalOpen = m_model.createProperty
("http://iec.ch/TC57/2001/CIM-schema-cim10#Switch.normalOpen");
public static final Property Switch_switchOnCount =m_model. Create
Property
("http://iec.ch/TC57/2001/CIM-schema-cim10#Switch.switchOnCount");
public static final Property Switch_switchOnDate =m_model.createProperty
("http://iec.ch/TC57/2001/CIM-schema-cim10#Switch.switchOnDate");
  
```

2.2 通过 JENA 创建和填充 RDF 模型

通过 JENA CIM 词汇就可以创建和填充 RDF 模型来生成符合 CIM 标准的 RDF 文件。Jena 的 ModelFactory 类是创建不同类型模型的首选方式,通过调用 ModelFactory 的方法 ModelFactory.createDefaultModel() 来得到一个空的模型,这种方法返回 Model 实例,使用它可以调用 model.createResource 创建表示 CIM 中每个对象的 Resource。创建了资源后,就可以将这些成员的属性添加到模型中。将属性添加到模型中的最简单方法是通过调用 Resource.AddProperty()。此方法以 Resource 作为主题在模型中创建语句。该方法使用两个参数,表示语句谓词的 Property 和语句的对象。

```

// create an empty model
Model model = ModelFactory.createDefaultModel();
sql = "select * from Station";
rs = stmt.executeQuery(sql); //执行查询
IDType = "station";
while (rs.next()) {
    ID = rs.getString("STANO");
    model.createResource(IDPrefix + IDType + ID, CIM.Substation)
.addProperty(CIM.Naming_aliasName, rs.getString("STATIONNAME")
)
.addProperty(CIM.Naming_description, rs.getString("STATIONDESC
")); }
  
```

以上代码以导出 EMS 数据库中站点信息为例展示了完整的创建与填充 RDF 模型的过程。主要分为 4 步:①查询 EMS 数据库得到相关信息。② ModelFactory.createDefaultModel() 创建一个空的模型。③通过 model.createResource() 创建资源,该过程往往用于将 EMS 数据库的表表示为 CIM 中的对象。④通过 Resource.addProperty() 创建属性,该过程往往用于将 EMS 数据库的表中的字段表

示为 CIM 中的资源的属性。需要注意的是, 创建资源与创建属性步骤需要应用 JENA CIM 词汇, 故在编写这部分程序时需要包括 JENA CIM 词汇文件。

表 1 是 EMS 的 ORACLE 数据库记录系统中变电站的名称与 ID 信息的表的前三条记录。

表 1 ORACLE 中电站信息表

Tab.1 Table of station information in ORACLE

STA_ID	STATIO_NNAME	STATION_DESC
1	QT	青台
2	NY	南阳
3	TH	唐河

以下是上表的第一条记录进行通过 JENA 转换后的 CIM/XML 文档。转换后的 CIM/XML 文档同样描述了一个名叫“青台”的站点, 它的英文名字是“QT”。该站点的 ID 是“station1”。文档中的所有标记具有“cim”的命名空间前缀, 表示这些标记在 CIM RDF Schema 中定义。

```
<rdf:Description rdf:about=
  "http://www.hust.edu.cn/derc/dts#station1">
  <rdf:type rdf:resource=
    "http://iec.ch/TC57/2001/CIM-schema-cim10#Substation"/>
  <cim:Naming.aliasName>QT</cim:Naming.aliasName>
  <cim:Naming.description>青台</cim:Naming.description>
</rdf:Description>
```

实际系统中, 对于不同的 EMS 数据库, 由于表达数据的表与字段是不同的, model.CreateResource() 创建资源可能需要多张表来表示 CIM 中某个对象, 也可能一张表可以表示多个 CIM 对象。同样, Resource.addProperty() 可能需要访问多张表的字段。

3 使用 JENA 查询 CIM/XML 文件

在建立了 CIM/XML 文件后, 新的基于 IEC61970 的 EMS 就可以直接使用 CIM/XML 文件作为其数据源。IEC61970 在 CIM 的基础上提出组件接口规范 CIS (Component Interface Specification), CIS 能够以一种标准方式对基于 CIM 的信息进行访问^[3,4], 在访问过程中如何高效地对 CIM/XML 数据进行查询成了关键的问题。JENA 的出现为解决查询问题提供了一条简单而高效的途径。

由于 CIM/XML 数据本质上就是基于 CIM 语义的 RDF 文件。而 JENA 可以通过 RDQL 来实现对 RDF 数据的查询。RDQL 是 RDF 的查询语言。虽然 RDQL 还不

是 W3C 的正式标准, 但已由 RDF 框架广泛执行。RDQL 允许简明地表达复杂的查询。RDQL 的语法表面上类似 SQL 的语法, 它的一些概念对已经使用过关系数据库查询的人来说将比较熟悉。Jena 的 com.hp.hpl.jena.rdql 包包含在 Java 代码中使用 RDQL 所需的所有类和接口。要创建 RDQL 查询, 将 RDQL 放入 String 中, 并将其传送给 Query 的构造函数。通常直接设置模型用作查询的源, 除非在 RDQL 中使用 FROM 子句指定了其他的源。一旦创建了 Query, 可以从它创建 QueryEngine, 然后执行查询。以下代码显示如何从基于 CIM 模型的 RDF 文件实例中查找所有 Breaker 的 Switch.normalOpen 的值。

```
String queryString =
  "PREFIX cim: <http://iec.ch/TC57/2001/CIM-schema-cim10#> " +
  "PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> " +
  "SELECT ?x ?y1 ?y2 ?y3 " +
  "WHERE {" +
  " ?x rdf:type cim:Breaker ." +
  " ?x cim:Naming.name ?y1 ." +
  " ?x cim:Switch.normalOpen ?y2 ." +
  "?x cim:Equipment.MemberOf_EquipmentContainer ?y3 ." +
  " }";
Query query = QueryFactory.create(queryString);
QueryExecution exc = QueryExecutionFactory.create(query, model);
ResultSet results = exc.execSelect();
// Output query results
ResultSetFormatter.out(osw, results);
```

传统的 CIS 接口的实现是基于数据库的查询语言 SAL^[5], RDQL 查询 RDF 数据的方式与 SQL 类似; SELECT 子句: 识别程序返回的变量。如果程序不需要所有的变量, 则指定所需要的结果, 从而减少内存开销; WHERE 子句识别限制性条件; PREFIX 子句: 缩短 URIs 长度的一种方法, 通过较短的字符串来代替较长的 URIs。变量则由“?”符号导入。变量可以代替 RDF 中的主体, 谓词, 客体各个部分, RDQL 会据此与 RDF 数据进行匹配来返回查询的结果。

4 基于 JENA 技术处理 CIM 数据优点

传统处理 CIM 数据的方法是基于 XML 的 DOM 技术来生成与查询 CIM/XML 数据。

如图 3 所示, 在基于 XML 的 DOM 技术处理 CIM 数据过程中实际上是将 RDF 描述的 CIM 数据当作普通的 XML 数据来进行处理^[6], 无法充分利用 RDF 特性。

首先,在生成与查询 CIM/XML 过程中需要手工编写 XML 的 DTD 文件作为解析规则,过程繁杂且容易出错。其次,IEC61970 是未最终定稿的标准,每次 IEC61970 标准在变动后,都需要再次手工生成 DTD 文件,这使得基于 DOM 技术的 CIM 系统缺乏必要的灵活性。最后,DOM 技术本身的查询能力有限,无法直接实现复杂的查询。

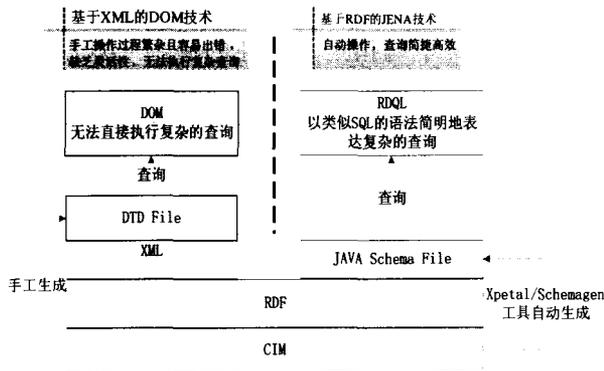


图3 基于RDF的JENA技术与基于XML的

DOM技术处理CIM数据

Fig.3 Operation of CIM with JENA based on RDF and DOM based on XML

在基于RDF的JENA技术中,JENA技术直接基于RDF技术,可以充分利用RDF的特性来操作CIM数据,而且JENA可以通过相关工具来自动生成Java Schema文件,以此建立CIM的RDF模型,并基于此模型,就可使用JENA的RDQL技术来执行查询。通过引入JENA,不仅使得CIM Schema的生成自动化而且使得对CIM数据的查询变得简洁与高效。

5 结论

通过在CIM应用中引入JENA技术,JENA不仅能够自动地生成CIM/XML数据,而且可以高效地对

CIM/XML进行查询从而可以进一步实现CIS接口。JENA作为RDF数据表达、解析、查询的强大工具。使得新的基于CIM的EMS应用开发集中在EMS本身的应用设计上,对于数据的建立、访问、查询都可以直接依赖于JENA。总之,JENA为IEC61970标准的进一步的实用提供了一条新的有效途径。

参考文献

- [1] IEC 61970,Energy Management System Application Program Interface, Part 301: Common Information Model (CIM) Base[S].
- [2] <http://www.w3.org/TR/rdf-primer,2004>.
- [3] IIEC 61970,Energy Management System Application Program Interface,Part 401: Component Interface Specification (CIS) Framework[S].
- [4] IEC 61970, Energy Management System Application Program Interface, Part 402: Base Services[S].
- [5] 王志南,吴文传,张伯明,等.基于IEC61970的CIS服务与SVG的研究和实践[J].电力系统自动化,2005,29(22):65-68,85.
WANG Zhi-nan,WU Wen-chuan,ZHANG Bo-ming,et al.Study and Implementation of CIS and SVG Based on IEC 61970[J].Automation of Electric Power Systems,2005,29(22):65-68,85.
- [6] 仇宏祥,王康元.能源管理系统数据CIM导出的实现[J].继电器,2006,34(1):71-75.
QIU Hong-xiang,WANG Kang-yuan. Export Technique Based on Common Information Model(CIM)in EMS[J].Relay,2006,34(1):71-75.

收稿日期:2006-09-15; 修回日期:2006-11-27

作者简介:

樊荣(1979-),男,硕士研究生,主要从事EMS与异构数据交互的研究;E-mail:coosty@163.com

杨峰(1973-),男,博士,讲师,主要从事电力系统仿真的研究工作;

左吉昌(1981-),男,硕士研究生,主要从事电力系统调度与优化的研究。

许继高压直流输电控制保护系统获国家重点新产品证书

日前,许继集团公司研制的DPS-2000高压直流输电控制保护系统荣获2006年度国家重点新产品证书。证书由科技部、商务部、国家质检总局和环保总局共同颁发。

DPS-2000型高压直流输电控制和保护系统,主要在运行控制系统、运行人员培训系统、直流控制保护功能和配置以及硬件平台等部分对引进技术进行了升级改进,研发成果在性能试验和安装调试中得到了客户以及外商的高度评价和认可,使国内直流输电技术和控制保护设备的国产化水平实现了质的飞跃。许继研制的DPS-2000型高压直流输电控制和保护系统,在技术性能方面已经达到了目前国际先进水平,不仅能够满足±500 kV超高压直流输电工程的技术要求,并且经与有关专家认真研讨,可以适用于±800 kV特高压直流输电工程,满足特高压直流对控制保护系统在运算速度、功能和可靠性等多方面的要求,为促进国产化技术和设备的成熟做出了贡献。