

基于 CIM/E 的电网模型解析^①

崔 丽^{1,2}, 王占国², 吴 奕²

¹(中国科学院大学, 北京 100049)

²(中国科学院 沈阳计算技术研究所, 沈阳 110168)

摘 要: 电力系统中模型数据交换的实现具有重要的实际意义. 为了更好的解决这一问题, 提高系统的互操作性, 在研究由国际电工技术委员会 (IEC) 制定的 IEC 61970 系列标准中的公共信息模型 (CIM)、分析 XML 文件的解析方法和电网物理模型的相关设备类属性描述的基础上, 提出了一种使用 Java 语言对 CIM/E 文件进行解析的方案. 即将 CIM/E 电网模型数据解析成内存中的结构化数据, 并且将数据存入数据库, 以防下次使用时再次解析, 提高重用性, 这里使用的是 MySQL 数据库, 并且最后用导出的电网模型 CIM/E 文件进行了验证. 提出的解析方案对实际电力系统中模型数据的解析具有一定的指导和参考意义.

关键词: 电力标准; 公共信息模型; E 语言; CIM/E 文件解析

引用格式: 崔丽, 王占国, 吴奕. 基于 CIM/E 的电网模型解析. 计算机系统应用, 2018, 27(1): 168-173. <http://www.c-s-a.org.cn/1003-3254/6151.html>

Analysis of Power Grid Model Based on CIM/E

CUI Li^{1,2}, WANG Zhan-Guo², WU Yi²

¹(University of Chinese Academy of Sciences, Beijing 100049, China)

²(Shenyang Institute of Computing Technology, Chinese Academy of Sciences, Shenyang 110168, China)

Abstract: The realization of model data exchange in power system is of great practical significance. In order to solve this problem and improve the interoperability of the system, a solution of analyzing CIM/E file using Java language is proposed, which is based on the study of the Common Information Model (CIM) in the IEC 61970 series standard developed by the International Electrotechnical Commission (IEC), the analysis method of XML file and the description of related equipment class attribute of grid physical model. The solution means to analyze the CIM/E grid model data into the structured data in memory, then store the data into database to prevent analyzing it again the next time to improve the reusability. This database we use is MySQL, and finally the exported grid model CIM/E file validates this solution. The analytical scheme has much significance of guidance and reference to the analysis of the model data in the actual power system.

Key words: power standard; common information model; E language; CIM/E file analysis

随着现代经济的高速发展, 电力需求持续增加, 电网的规模在不断扩大, 电网形态也在逐渐变化^[1]. 目前电力系统已经不可否认的发展成世界上最复杂的人造工业网络之一, 大电网安全问题已经成为人们关注的热点问题. 保障电网经济、稳定、安全的运行是电力

行业工作者一直以来追求的宗旨. 周孝信院士于 2011 年在国家电网报上提出了三代电网理论^[2], 将电网和电力系统的发展分为三个阶段, 即第一代小型电网, 第二代大型互联电网, 以及第三代智能电网. 关于电网的发展水平世界各国是有差异的, 但总的来说, 是处于第二

① 收稿时间: 2017-04-11; 修改时间: 2017-04-26; 采用时间: 2017-05-04; csa 在线出版时间: 2017-12-22

代大型互联电网过渡到第三代智能电网的阶段. 电网的稳定、安全、高校运行离不开电网模型数据和运行数据的支撑. 应用系统之间的信息共享和应用集成逐渐成为迫切需要解决的问题.

电力系统的不断发展和自动化技术的提高, 使电网调度自动化系统得到广泛的应用. 出现了变电站自动化系统、电网能量管理系统 (EMS) 等各种自动化系统, 涉及电力行业的各个环节. 这些系统通常由不同的厂家提供, 可能采用不同的开发平台、数据库技术和应用接口, 应用间的信息共享成本很高, 不能很好的实现系统之间的互联互通. 对此, 美国电力科学研究院 (EPRI) 启动了 CCAPI 项目^[3], 主要目标就是推动由不同厂商开发的 EMS 应用之间, 独立开发的完整 EMS 系统之间, 以及 EMS 系统与有关电力系统运行的其他系统之间的集成, 并为数据交换提供方便. 目前电力系统中对于 CIM/XML 的研究和解析技术已比较成熟, 但对于 CIM/E, 在研究的过程中发现很少有针对 CIM/E 解析的研究介绍, 缺乏对 CIM/E 的解析过程的详细说明, 所以本文在研究 CIM/E 结构的基础上, 提出了一种 CIM/E 文件的解析方案并详细介绍了具体的解析过程, 弥补该部分的空白, 为实际的应用提供一定的参考和指导价值.

1 公共信息模型 (CIM)

1.1 IEC 61970 简介

美国电力科学研究院 (EPRI) 于 1993 年启动了“控制中心应用接口”(Control Center Application Program Interface, CCAPI) 研究项目 (RP-3654-1), EPRI CCAPI 项目的主要目标是: 减少向 EMS 中增加新应用所需要的费用和时间; 保护对 EMS 中正在有效工作的现有应用的投资. 随着各种组件软件技术的出现, 1999 年开始确立了基于组件软件的相对稳定的 CCAPI 解决方案.

基于 EPRI CCAPI 项目的工作, 国际电工委员会 57 技术委员会与 EPRI 紧密合作, 启动 IEC EMS-AIP 项目^[4], 制定了 IEC 61970 系列标准, 该系列标准定义了能量管理系统的应用程序接口 (EMS-API), 包括导则和总体要求、术语表、公共信息模型 (CIM)^[5]、组件接口规范 (CIS) 和组件接口规范映射 5 大部分内容. 该协议提出为现有的应用系统提供一个基于公共信息模型、公共体系结构和组件技术的系统集成框架. 各

个 EMS 应用内部可以有各自的信息描述方式, 但是只要在应用程序接口语义级上基于公共的信息模型, 不同厂商开发的应用程序或不同系统的应用间就可以以同样的方式访问公共数据, 实现应用间的相互操作和插件兼容. 通过定义应用编程接口 (API), 使这些应用或系统可以访问公共数据或进行信息交换, 而不依赖于信息的内部表述形式. 公共信息模型 (CIM) 规定了此信息交换内容的语义. 组件接口规范 (CIS) 规定的消息交换的内容.

1.2 公共信息模型 (CIM)

IEC 61970 系列标准主要包括公共信息模型 (CIM) 和组件接口规范 (CIS) 两方面内容, 其中 CIM 是整个协议的核心. 公共信息模型 (CIM) 是一个抽象模型, 它描述电力企业的所有主要对象, 用对象类和属性及它们之间的关系来表示电力系统资源. 通过定义一种基于 CIM 的公共语言 (即语法和语义), 使得应用或系统能够不依赖于信息的内部表示而访问公共数据和交换信息来实现系统间的集成. CIM 中描述的对象类本质上是抽象的, 可以用于各种应用. CIM 的使用远远超出了它在 EMS 中应用的范围. 应当把本标准理解为一种能够在任何一个领域实行集成的工具, 只要该领域需要一种公共电力系统模型, 使得应用和系统之间能够实现互操作和插入兼容性, 而与任何具体实现无关.

由于完整的 CIM 的规模较大, 所以将包含在 CIM 中的对象类分成了几个逻辑包, 每个逻辑包代表整个电力系统模型的某个部分. CIM 规范使用统一建模语言 (UML) 表达方法, 它将 CIM 定义成一组包. CIM 中的每一个包包含一个或多个类图, 用图形方式展示该包中的所有类及它们的关系. 然后根据类的属性及与其它类的关系, 用文字形式定义各个类. IEC 61970-301 内容包括核心包、域包、发电包、发电动态包、负荷模型包、量测包、停运包、生产包、保护包、拓扑包和电线包. 每一个 CIM 包的类图展示了该包中的所有类及它们的关系. 在与其它包中的类存在关系时, 这些类也展示出来, 而且标以表明其所属的包的符号. CIM 301 部分的包图如图 1 所示.

2 电网模型交换标准

2.1 CIM/XML

通过 CIM 提出的电力系统信息交换标准可以实现不同系统间的互联互通. 但 CIM 模型是个抽象模型,

它没有定义模型数据库的规范和数据交换格式,在具体的工程应用中,对 CIM 模型的实现方式需要有明确、可行的规定. XML 语言解决了这个问题, CIM 以 XML 语言为载体,使用 CIM RDF 模式作为元模型框架,并采用一个 RDF 语法的子集作为描述规则,描述电网静态数据模型,构建电力系统信息模型的 XML 文档^[6].

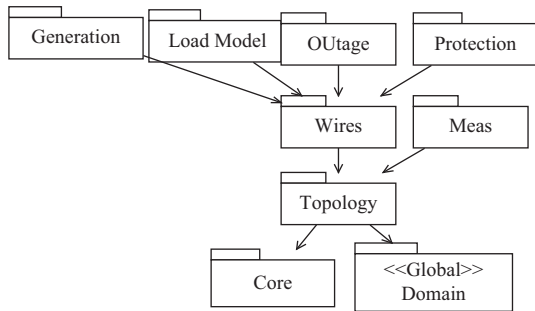


图1 IEC61970-301包结构

根据 IEC 61970 501 标准所描述的 CIM RDF 模式,一个 EMS 电力系统模型能被转换导出为一个对应的 CIM XML 文档, CIM RDF 模式提供了 CIM XML 文档所使用的资源描述格式,最终的 CIM XML 模型文件能够被解析,将其中的信息导入到另一个外部系统.目前基于 CIM/XML 的数据导入导出是不同系统和软件间进行数据交换和信息共享的主要方式.

2.2 CIM/E

为了提高电力系统模型数据的描述效率和大量数据的在线交换的效率,出现了电力系统数据模型描述语言 E^[7], E 来源于 3 个英文单词的首字母: easy(简单)、efficiency(高效)、electric-power(电力). E 语言有 XML 的基本特点和优点,使用它的特殊符号和描述语法,能实现电力系统中各种数据模型的高效描述,效率比 XML 高很多,而且更符合人类的表述习惯^[8].

E 语言使用几个英文半角符号位于文件每行开头的第 1 个与或第 2 个字符组合,表示不同的特定含义,比如: <, >, @, #, / 等符号, E 语言有三种数据描述方式: 衡表式, 单列式和多列式. E 语言出了基本功能, 还有扩展功能, 可以用于表示类属性的类型、单位、限值等. CIM/E 文件的文件名格式为, 区域_日期_时间. CIME, 比如, 驰宏站_20170330_141315.CIME. CIM/E 文件按照 E 格式规范对电网物理模型的相关设备类属性进行组织. 为了使电网调度自动化系统的运行要求得到满足, 应对电网物理连接模型进行描述, 包含区域、负

荷、间隔、厂站、变压器、基准电压、母线段等各类对象, 对各类对象包含的属性项有相关的要求.

3 CIM/E 文件解析

3.1 XML 文档解析分析

XML 文档的解析^[9]主要有两种方式: DOM 和 SAX. DOM 是由 W3C 制定的一套编写 XML 分析器的标准接口规范, SAX 是 XML_DEV 邮件列表中的成员根据应用的需求自发的定义的对 XML 文档进行操作的一套接口规范. 这两种接口规范各有长短, 互有侧重, 都有着广泛的应用. 应用程序并不是直接对 XML 文档进行操作的, 先由 XML 分析器对 XML 文档进行分析, 而后, 应用程序通过 XML 分析器所提供的 DOM 接口或 SAX 接口对分析结果进行操作, 从而间接的实现了 XML 文档的访问, DOM 和 SAX 在应用程序开发过程中的作用, 如图 2 所示.

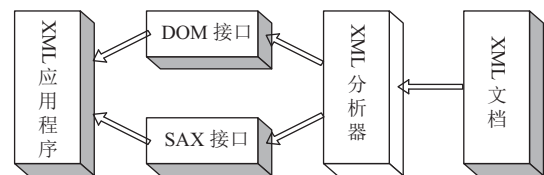


图2 DOM 和 SAX 在应用程序中的作用

文档对象模型 (DOM) 的全称为 Document Object Model. DOM 就是一个标准的对象化的接口规范, 与具体语言和平台无关. 在程序中, 基于 DOM 的 XML 分析器首先将一个 XML 文档转换成一个 DOM 树 (对象模型的集合) 放在内存中, 然后应用程序以通过该对象模型的操作而实现对 XML 文档中数据的操作. 利用 DOM 接口的解析机制被称作随机访问机制, 因为通过 DOM 接口, 应用程序可以在任何时候访问 XML 文档中的任何数据, 这种随机访问方式对于应用程序来说是非常灵活的.

XML 实际上是一种分层的结构, DOM 使用树模型来描述 XML 文档信息非常有效, DOM 接口提供通过分层对象模型的方式来访问 XML 文档中的信息, 根据 XML 文档的结构, 分层对象模型形成一颗节点树的形式. 解析形成的 DOM 树是存储在内存中的, 当 XML 文档比较复杂庞大时, DOM 解析器对内存的要求就比较高. 图 3 描述了 DOM 解析器解析 XML 文档形成内存中的节点树的过程. DOM 首先将文档加载进内存里, 然后再内存中构造出一个与文档对应的一棵树, 树

上的每个节点都是一个对象 (Node 类型), 放在内存里, 访问是随机的。

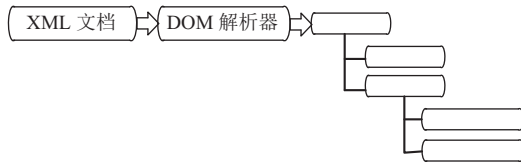


图3 DOM 解析器功能

XML 简单应用程序接口 (SAX), 全称为 Simple APIs for XML. 支持 DOM 的解析器是 W3C 的官方标准, 而支持 SAX 的解析器是事实上的工业标准. SAX 不同于 DOM, SAX 不必将整个 XML 文档事先加载到内存当中, 它快速读写 XML 数据, 是一种基于事件驱动的循序的访问模式, 不能重新访问已经分析过的内容. 对 XML 文档进行分析的时候, 会触发一系列的事件, 并激发相应的事先定义好的事件处理回调函数, 应用程序通过事件函数从而实现对 XML 文档的处理和访问. SAX 只做了一些简单的工作, 应用程序要自己去做大部分的工作, SAX 适用于只需要访问不要更改的 XML 文档数据的应用程序. 文档的读入过程就对应着 SAX 解析器的解析过程, 图 4 描述了 SAX 解析器解析 XML 文档的过程.

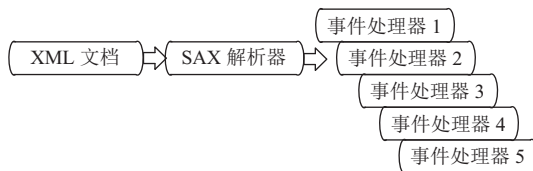


图4 SAX 解析器功能

3.2 CIM/E 文件解析方案

XML 文档解析器实际上就是一段程序代码, 同样 CIM/E 文档的解析器在本文中也是一段 Java 程序代码, 它读入一个 CIM/E 文件并分析其结构, 在分析文档的同时, 采用基于事件的方法在内存中构造出一个对应的自定的文档类型对象, 该文档类型实现了 W3C 标准中的 Document 接口. 然后程序可以对该文档对象进行操作, 将文档中的元素的不同信息分别提取出来存储到集合 (ArrayList 类型) 中元素 (TableEClass 类型) 的不同数据结构中.

其中 TableEClass 类型是构造出的一个数据结构, 用来存放从 CIM/E 文件解析出的元素数据内容, 该类包含 tableName、columnNames、columnTypes、

columnUnits、columnLimitValues、rowData 等属性, 分别存储 CIM/E 文件中对应的类名、属性名, 与横表式 CIM/E 文件中对应的属性类型、属性单位、属性限值和数据行的内容. 上述集合对象封装了 CIM/E 文档中的数据信息, 有了该集合对象之后, 就可以在该对象的基础上进行数据的入库等操作, 为后续的应用做好准备, List 集合中的每个元素 (TableE 类型) 对应于数据库中的一张表, 并包含了表中的数据.

本解析方案根据 CIM/E 文件结构特点, 在 W3C 提供的 DOM 接口和 SAX 接口的基础上, 针对 CIM/E 文件解析定义了一系列的节点类型等数据结构, 文件中一个类对应为一个元素, 将类起始符合类结束符中间的内容设置为该元素的内容, 在封装数据的时候在具体分析取出. 该解析方案主要分为基于事件的文档对象的构建与操作文档进行数据的封装两个功能, 涉及到的主要类和方法如图 5 所示.

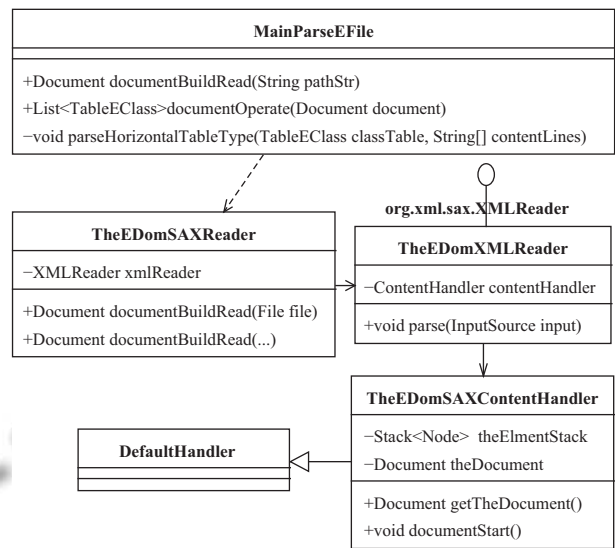


图5 解析方案中的部分主要类和方法

3.2.1 基于事件的文档对象的构建过程

首先定义 MainParseEFile 类 Document documentBuildRead (String pathStr) 方法, 通过 CIM/E 文件名来构建一个与之对应的自定义文档对象 Document. 该方法的实现依赖于 TheEDomSAXReader 类型对象的 documentBuildRead() 方法, 该类有个 XMLReader 类型的引用, 该引用指向的对象的 void parse(InputSource input) 方法完成底层的文件的具体分析操作.

TheEDomXMLReader 类实现了 XMLReader 接口, 而 XMLReader 接口的作用就是用于使用回调读取

XML 文档, 所以定义了 TheEDomSAXContentHandler 类并在类中提供了各种回调函数. parse() 方法根据 CIM/E 文件中的标识符号分析到文件的不同位置时会调用处理器类中定义的各种回调方法. 这些回调方法在处理过程中借助于栈类型成员变量的进栈、出栈、取栈顶元素操作来实现自定义文档对象的构建.

最后在 TheEDomSAXReader 类型对象的 documentBuildRead() 方法中通过处理器类型引用的 getTheDocument() 方法返回自定义的文档对象, 如图 6 所示.

```
public Document documentBuildRead(InputSource in) throws Exception {
    try {
        XMLReader reader = new TheEDomXMLReader(); //创建TheEDomXMLReader类的对象
        TheEDomSAXContentHandler handler = new TheEDomSAXContentHandler(); //创建处理器对象
        reader.setContentHandler(handler); //将处理器设置为reader对象的引用
        reader.parse(in); //完成具体的文件分析工作
        return handler.getDocument(); //返回构造的文档对象
    }
}
```

图 6 documentBuildRead() 方法的片段代码

3.2.2 操作文档进行数据封装

MainParseEFile 类中的 List<TableEClass> documentOperate(Document document) 方法实现了将由 documentBuildRead() 方法得到的文档对象中的信息存储到集合变量 list 中的操作. 该方法首先取出文档对象下的所有子元素, 遍历每个元素, 生成一个 TableEClass 对象, 将元素名对应的类名设置为对象的 tableName 属性值, 并将该对象添加到 ArrayList<TableEClass> 类型变量中, 同时, 根据元素内容的的首行初始符号判断元素的类型 (即元素所属 CIM/E 文件的结构类型, 是单列式、多列式、还是横标式), 针对不同的文件结构使用不同的解析方法将元素内容填充到 TableEClass 对象中的不同属性上.

比如横表式文件的解析, 使用 void parseHorizontalTableType(TableEClass classTable, String[] contentLines) 方法. 遍历元素内容的各行, 将属性引导符 @ 引导的元素内容第一行的各个属性, 设置到 TableEClass 对象的 columnNames 属性上, 针对其余的内容行, 若是以类型引导符 “%” 开头, 则将该行的各个类型保存到 TableEClass 对象的 columnTypes 属性上; 若是以量纲引导符 “\$” 开头, 则将该行的各个单位保存到 TableEClass 对象的 columnUnits 属性上; 若是以限值引导符 “:” 开头, 则将该行的各个限值保存到 TableEClass 对象的 columnLimitValues 属性上; 若是以数据引导符 “#” 开头, 则将该行的各个数据追加到

TableEClass 对象的 rowData 集合属性中.

documentOperate() 遍历完文档中的每个子元素后, 在方法的最后返回已填充数据的 ArrayList<TableEClass> 类型的对象, 之后应用程序即可在该对象基础上进行数据入库等操作.

4 实例解析测试

下面以一个变电站的电网模型文件为例 (CIM/E 文件) 进行解析方案的实验验证, 如图 7 所示是导出的电网模型文件 (CIM/E 文件) 的模拟片段.

```
<!--Entity=某区域 type=电网模型 time='2017-03-30 14:13:15'!>
<ControlArea::某区域 date='2017-03-30 14:13:15'> -----表名
@Num nRID name Parent p q region_id -----字段
//序号 标识 区域名 父区域标识 总有功 总无功 所属系统区域
# 1 115713890602242 某区 NULL NULL NULL 1
# 2 115713890757186 谋区 115713890602243 NULL NULL 10
# 3 115713890929346 某某电网6 NULL NULL NULL 20
.....
</ControlArea::某区域>
<BaseVoltage::某区域 date='2017-03-30 14:13:15'> -----表名
@Num nRID name noakV region_id -----字段
//序号 标识 基准电压名 基准电压 所属系统区域
# 1 113871265677274 20kV 20.00 1
# 2 113871265677275 500kV 525.00 1
# 3 113871265677276 35kV 35.00 1
.....
</BaseVoltage::某区域>
<Substation::某区域 date='2017-03-30 14:13:15'> -----表名
.....
```

图 7 CIM/E 文件片段

XML 文档对象模型有根节点 (Document) 和根元素节点 root, 但电网模型文件 CIM/E 文件对应的文档模型是没有根元素节点的, Document 节点下就是各个并列的子元素. 若将 CIM/E 文件导入关系型数据库中, 文件中的类名如 ControlArea、BaseVoltage、Substation 等应该作为表存在, 类中的属性作为表的字段、数据作为表中记录的相应字段值.

图 8 和图 9 是该 CIM/E 文件应用该解析方案过程中在内存中形成的自定义文档对象和对应的存储数据的 ArrayList 集合对象.

得到 ArrayList 形式的信息后, 就可以对这个信息进行操作和处理了, 比如将其存储到关系型数据库^[10]. 使用的是 Java 进行数据库操作, Java 提供的单纯的使用 sql 语句方式进行编程的机制, 即 JDBC (Java Database Connectivity) 技术. 使用纯的 Java 代码来去实现对应的数据库的编程, 写相应的 SQL 语句, 调用相应的语句把 SQL 代码进行执行, 把语句发送到数据库端, 让数据库端去执行 SQL 代码. 比如执行一条插入语句, 数据库中就具备了插入进去的内容. 实际的大型应用

中后端一般会使用框架, 比如 Hibernate、MyBatis 等。在程序中取得数据库的连接, 调用建表的方法, 并填充字段的值, 形成数据表结构和数据。如图 10 所示为形成的 ControlArea 表结构数据。

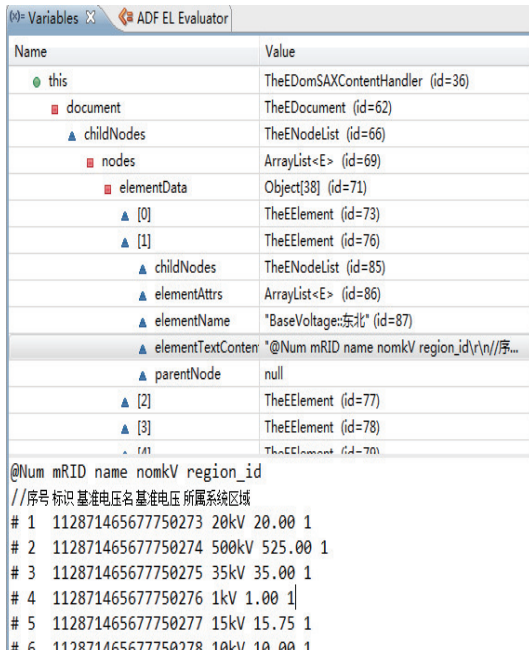


图 8 构建的自定义 Document 文档

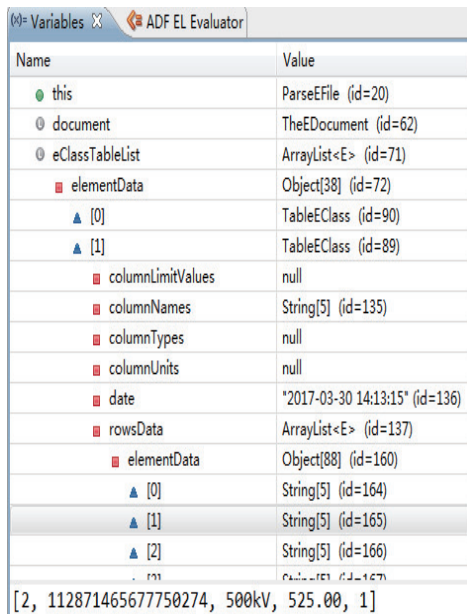


图 9 存储数据的 ArrayList 集合对象

5 结束语

本文首先介绍了公共信息模型 (CIM) 和两种电网模型交换标准, 然后分析了 XML 文件解析的方法及电网模型文件 CIM/E 的文件结构, 之后描述了一种解析 CIM/E 文件方法的具体过程及将解析的数据入库, 并通过电网模型数据文件 CIM/E 进行了验证, 该解析方法对电力系统中模型数据的解析具有一定的指导和参考意义, 可以在此方案的基础上完善各种扩展形式的文件的解析、设计数据结构, 提高解析效率等工作。

Num	mRID	name	Parent	p	q	region_id
1	113715890607882241	国调	{null}	{null}	{null}	1
2	113715890758877185	华北	113715890607882241	{null}	{null}	10
3	113715890926649345	东北电网6	{null}	{null}	{null}	20
4	113715890926649346	T接线6	113715890926649349	{null}	{null}	20
5	113715890926649347	蒙东	113715890926649349	{null}	{null}	20
6	113715890926649348	全局量	113715890926649349	{null}	{null}	20
7	113715890926649349	东北	113715890607882241	{null}	{null}	20
8	113715890926649350	蒙东省调	113715890926649351	{null}	{null}	20
9	113715890926649351	东北公司	{null}	{null}	{null}	20

图 10 ControlArea 表结构数据

参考文献

- 辛耀中. 新世纪电网调度自动化技术发展趋势. 电网技术, 2001, 25(12): 1-10. [doi: 10.3321/j.issn:1000-3673.2001.12.001]
- 周孝信. 第三代电网技术正在向我们走来. 国家电网报, 2011-05-05(005).
- 王民昆. 电力系统 CIM 模型描述及数据交换实现的研究 [硕士学位论文]. 成都: 四川大学, 2004.
- Draft IEC61970 Energy management system application program interface (EMS-API)-Part I: Guidelines and general requirements. Revision 5, 2002.
- Draft IEC61970 Energy management system application program interface (EMS-API)-Part 301: Common information model (CIM) base. Revision 3, 2004.
- 谢淑菁. 基于 CIM 的电力系统模型拼接的研究和设计 [硕士学位论文]. 济南: 山东大学, 2010.
- 辛耀中, 陶洪铸, 李毅松, 等. 电力系统数据模型描述语言 E. 电力系统自动化, 2006, 30(10): 48-51, 92. [doi: 10.3321/j.issn:1000-1026.2006.10.010]
- 米为民, 辛耀中, 蒋国栋, 等. 电网模型交换标准 CIM/E 和 CIM/XML 的比对分析. 电网技术, 2013, 37(4): 936-941.
- 刘芳, 肖铁军. XML 应用的基石: XML 解析技术. 计算机工程与设计, 2005, 26(10): 2823-2824, 2839. [doi: 10.3969/j.issn.1000-7024.2005.10.083]
- 朱晓娟. 基于关系的 XML 数据存储技术研究. 计算机与现代化, 2009, (12): 51-54.