

非“十五”前兆数据转换方案的研究与应用*

田野¹, 郝臻¹, 王小娟¹, 牛延平¹, 王建军^{1,2}

(1. 甘肃省地震局, 甘肃 兰州 730000; 2. 中国地震局兰州地震研究所科技创新基地, 甘肃 兰州 730000)

摘要: 针对甘肃省前兆台网部分仪器不能并入“十五”前兆数据库的问题, 提出一种兼备灵活性和扩展性的软件解决策略。基于此策略实现的数据转换软件, 采用了动态编译机制, 理论上可以实现对任意格式前兆数据并入“十五”数据库前的处理工作, 具备较高的推广价值。实施中该软件将非“十五”标准数据转换为“十五”标准格式, 顺利完成了非“十五”标准仪器的自动并网工作。

关键词: 前兆台网; 数据转换; 非“十五”前兆数据; 动态编译

中图分类号: P315.6

文献标识码: A

文章编号: 1000-0666(2013)03-0390-05

0 引言

中国地震前兆台网作为我国目前规模最大的地震前兆监测台网, 涉及多项学科, 是国内地震前兆数据共享服务的主要数据来源(周克昌等, 2009; 杨满栋等, 2000)。中国地震前兆数据管理系统实现了地震前兆数据的逐级自动采集、交换、汇集以及野外观测仪器的网络化控制, 为前兆台网统一规范的运行、观测数据的整合与数据共享服务提供了可靠的技术支撑平台。

“十五”前兆数字地震观测网络工程项目和“九五”并网改造工程的建设及运行, 使各省前兆台网的大多数数字地震监测仪器接入地震局行业网, 大大简化了区域台网维护的复杂度, 为观测数据的管理和科学共享提供了极大便利。然而, 部分观测仪器不能并网接入“十五”前兆台网的现象依然普遍存在, 并且随着地震前兆理论研究和探测方法的不断进步, 多种结构各异、类型繁多的前兆数据将源源不断地产生, 这都对“十五”前兆数据管理系统的适用性和稳定性提出了新的挑战。

本文介绍了一种新的数据转换的实施方案, 旨在解决目前存在的部分前兆仪器不能并入“十五”前兆数据管理系统, 以及该类仪器产出的数据无法自动入库的问题, 并对该实施方案的运行

情况及存在的问题进行了分析和展望。

1 运行现状

“十五”前兆台网充分利用计算机通信技术, 将前兆数据系统部署在网络上, 通过基于 web 的采集管理软件和数据处理软件, 实现了前兆数据的统一管理 & 数据资源的共享和开放。“十五”前兆台网信息系统的传控流程如图 1 所示(马文娟等, 2010; 黄晖, 2008)。相对于早期的前兆系统, “十五”前兆数据网络通过以太网、采用基于 TCP/IP 协议的 SOCKET 连接的前兆台网专用设备与客户端进行数据信息交换(中国地震局, 2005a, b), 其传输模式有着明显的优势, 它使观测数据的传输和管理更加高效、安全、便捷, 这种模式是信息网络化的必然要求。

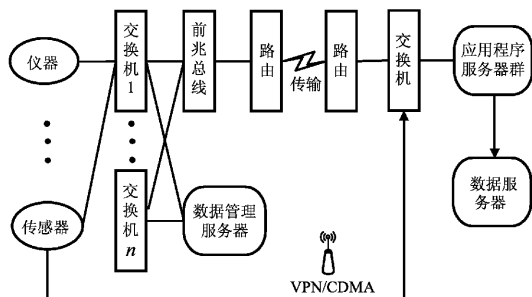


图1 “十五”数据传控流程图

Fig. 1 The flow chart of data transmission and control for “Tenth Five-Years” data

* 收稿日期: 2012-10-31.

基金项目: 甘肃省地震局兰州地震研究所地震科技发展基金(2012Q02)资助.

甘肃省地震局前兆台网目前仍存在部分非“十五”标准（以下简称“非标”）仪器，其数据格式与数据传输等技术指标无法与“十五”前兆数据管理系统兼容，只能由人工完成数据转换及传输等工作，效率低且成本高，安全性和可靠性差。因此将非标数据自动纳入“十五”前兆数据管理系统已成当务之急。

在不更改“十五”模式的运行流程且不影响前兆数据传输稳定性的前提下，较为合理的解决方案主要有两种：① 硬件策略。使用前兆仪器协议转换器等硬件设备完成，其功能在于将“十五”系统所发送的通信规程指令信息和其他通信信息进行双向解析，实现对非标数据的访问、采集和仪器控制等功能。该方案需要在应用程序服务器端或前兆仪器端增加新的协议转换设备（王秀英等，2004；何案华等，2008）；② 软件策略。开发“十五”数据转换专用软件，在数据采集后和入库前，根据相关通信协议完成数据格式的标准化转换。该方案只需在应用程序服务器端部署数据转换软件即可（王军等，2008；崔慧敏等，2004）。

硬件策略存在实现和部署成本过高等缺点；而软件策略具有低成本、方便部署、灵活且易修改等优点。因此本工作采用软件策略完成非标数据并入“十五”前兆数据管理系统的任务。

2 实施方案

非标数据在采集后，以文本文件的形式存储于原始数据服务器中，作为输入数据参与数据转换软件的处理；部署于应用程序服务器上的数据转换软件负责输出“十五”标准格式的数据文件，并将输出文件上传至“十五”前兆数据管理系统服务器中以备入库。整体处理过程如图2所示。

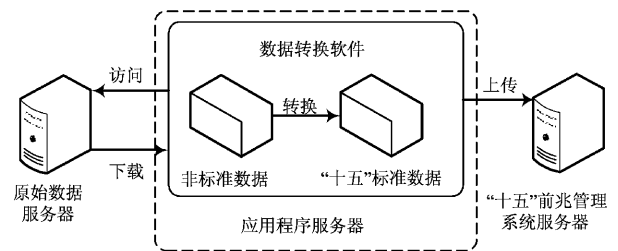


图2 非标准数据到标准数据的转换处理过程
Fig. 2 Transform process from the non-standard data to the “Tenth Five-Years” standard one

2.1 数据同步

如图2所示，数据文件的传输涉及到3台服务器间的文件同步问题，分别是应用程序服务器与原始数据服务器间、“十五”管理系统服务器与应用程序服务器间的数据通讯。文件同步是数据转换软件获得即时输入文件、生成输出文件并及时入库的前提。

鉴于服务器的操作系统各异（原始数据服务器和应用程序服务器均为 Windows，前兆数据管理系统服务器是 Suse Linux），因此考虑使用 Rsync 作为服务器间文件同步的工具。Linux 操作系统中自动集成了 Rsync，而在 Windows 平台上则部署 cwRsync 软件来实现服务器之间的数据文件同步。

2.2 转换流程设计

研究工作的主要环节为将非标数据格式转换为“十五”前兆数据管理系统兼容的标准数据格式，考虑到软件的灵活性和扩展性，数据转换软件应具备以下几个特征：

- (1) 数据转换软件的运行只与逻辑流程有关，和需要转换的数据及其格式无关；
- (2) 软件可以适应需要转换的数据类型的增加、删除、修改等变更操作；
- (3) 数据格式转换算法与软件的编译运行相互独立，算法不会影响软件运行的稳定性。

图3显示了转换软件在应用程序服务器中的工作流程，该工作流程以上述特征为设计前提，其中数据转换算法由用户根据原始数据和标准数据的格式设计，以函数的形式组织编写。数据转换软件具有编译数据转换算法和执行数据转换两个

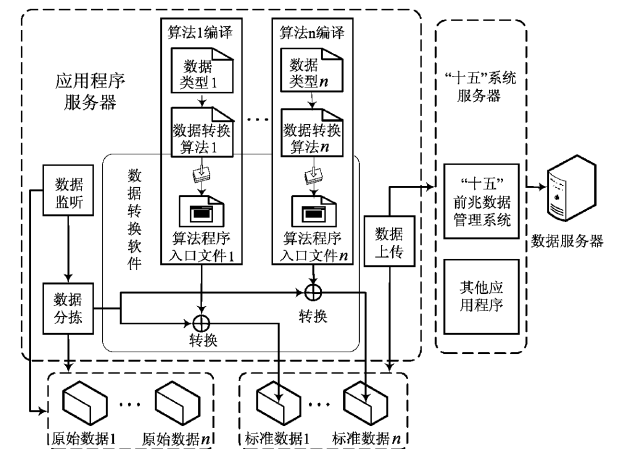


图3 数据转换方案工作流程图
Fig. 3 The working flow chart of data transformation scheme

主要作用。前者将算法编译为程序可识别的入口文件备用；后者通过数据监听模块获知原始数据的接收状况，并通知数据分拣模块实时获得原始数据。程序根据原始数据的类型选择对应的算法程序入口文件，之后运行转换功能将其保存为标准格式的数据，最后通过数据上传模块将转换后的标准数据上传到“十五”前兆数据管理系统服务器以备其他程序使用。

考虑到文件传输的安全性，上述流程中的数据监听、数据分拣和数据上传模块均由数据文件同步完成。

2.3 动态编译与执行

在具体实施中引入动态编译技术以实现转换流程所要求的 3 个特征。动态编译技术允许在程序运行时加入额外的编译信息以完成特定功能（崔慧敏等，2004），其优点在于减少了程序主体的体积，并且增强了程序的可扩展性；但另一方面，该技术也增加了程序框架的逻辑复杂度，使系统运行效能有所降低，并且对程序的内存管理提出较高要求。

程序框架采用 Microsoft Visual Studio 的 .Net Framework 开发平台完成，.Net 在 System. CodeDom. Compiler 命名空间中封装了各种支持指定编程语言源代码的生成和编译管理的类与接口，这些类与接口提供了实现动态编译功能的方法。数据转换软件的动态编译与执行的运行框架如图 4 所示。程序在运行时，首先检测转换算法是否存在变更，如果存在变更，则利用编译模块将以文件形式存在的数据转换算法动态编译为动态链接库（dll）文件；之后程序的执行模块将 dll 文件加载到新开辟的应用程序域中，利用反射机制，远程调用 dll

文件提供的数据转换算法，输入原始数据完成数据转换工作，最后销毁新的应用程序域及其内存占用。这种设计有效改进了动态编译在运行效率和内存管理等方面的不足。

2.4 技术实现

动态编译与执行的运行框架提供了程序中算法编译和数据转换模块的运行逻辑，其具体步骤如下：

（1）构建数据转换代码。获取所有数据转换算法的代码，生成包含完整命名空间以及转换算法类的代码文件。

（2）编译代码文件。.Net 的 CSharpCodeProvider 类提供动态编译功能，利用该类的 CompileAssemblyFromSource 方法编译第 1 步得到的代码文件，生成唯一的转换算法类的动态链接库，保存到程序指定目录。

（3）创建远程访问类和跨程序应用域调用的接口。新建继承于远程访问类 MarshalByRefObject 的子类 RemoteInvoker，实现其创建远程访问程序应用域的功能；新建接口 IRemoteInf，实现其跨应用程序域调用方法，用于执行指定的算法函数。

（4）开辟新的应用程序域。利用 AppDomain 类的 CreateDomain 方法生成一个新的可管理的应用程序域 appDom，以下的步骤皆在此空间中运行。

（5）加载数据转换代码。利用应用程序域类 AppDomain 的方法 CreateInstance 的创建 RemoteInvoker 的一个实例，并加载转换算法的动态链接库。

（6）调用数据转换功能。利用第 5 步的实例，实现接口 IRemoteInf 调用函数的功能，完成数据转换。该步骤使用反射机制，从当前应用程序域远程调用 appDom 域中的相关功能。

在上述步骤中，第 1、2 步只在新建或者更新数据转换算法时执行；第 3、4 步的作用为在主程序应用域之外创建新的应用域空间，用于管理内存中动态执行的数据转换功能，并为主程序提供了跨程序域调用的方法；第 5、6 步在执行转换时将生成的动态链接库加载到内存中的指定应用程序域中，利用反射调用需要的转换方法。结合图 3 可以看出，转换算法是以类似于插件的形式参与到程序的运行中的，更新、增加算法只需要对动态链接库文件进行相关操作即可，无需改动程序，保证了算法和转换程序之间的独立性。

具体编码实现如下：

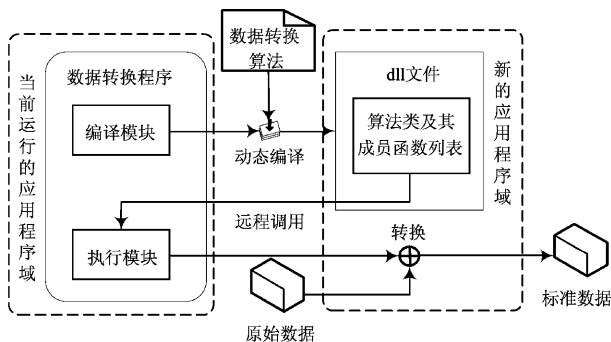


图 4 动态编译与执行的运行框架
Fig. 4 The operating framework of dynamic compilation and execution

```
//创建编译参数实例，加载引用、设置动态
链接库输出路径，arithmeticName 代表算法名称
CSharpCodeProvider objCompPara = new CSharp-
CodeProvider ();
objCompPara.ReferencedAssemblies.Add
(" System.dll");
objCompPara.OutputAssembly = " arithmetic
Name.dll";
//创建 CSharpCodePrivoder 对象 cp，source-
Code 为代码文件（实现步骤 2）
CSharpCodeProvider cp = new CSharpCodeProvid-
er ();
cp.CompileAssemblyFromSource ( objCompPara,
sourceCode);
//创建应用程序域（实现步骤 4），名称为
DataTransformation
AppDomain appDom = AppDomain.CreateDomain
(" DataTransformation", null, objSetup);
//创建远程访问类对象 ri，RemoteInvoker 继承
自 MarshalByRefObject（实现步骤 5）
RemoteInvoker ri = ( RemoteInvoker ) appDom.
CreateInstance ( " RemoteInvoke ", " RemoteInvoke.
RemoteInvoker" ). Unwrap ();
```

```
//加载动态链接库，加载转换算法类（实现
步骤 5）
object obj = ri.Create ( " arithmeticName.dll",
"DynComp. DataTransformation ", null);
//调用转换算法（实现步骤 6），objArithmet-
icParams 为算法所需的输入参数
IRemoteInf iri = (IRemoteInf)obj;
iri.Invoke ( " arithmeticName ", objArithmetic
Params);
```

3 实施结果

目前甘肃省前兆台网共有 17 台（套）仪器产出的非标数据使用本文的数据转换方案进行处理。以 ZKGD-3000 型中科流体仪水温数据为例，表 1 和表 2 分别列出了其原始数据格式和转换后的“十五”标准数据格式。数据转换软件中的算法根据此格式编写，处理后的数据已经完全符合“十五”标准。

在数据转换前，将每台 ZKGD-3000 型中科流体仪以非在线仪器的形式添加到“十五”管理系统中，它们各自具有唯一的仪器 ID 标识。之后运行数据转换软件实现非标数据的自动转换及入库。

表 1 ZKGD-3000 水温原始数据格式

Tab. 1 The original data format of the underground water temperature recorded by the ZKGD-3000 instrument

| 项目 | 格式 | 实例 |
|-----|--------------------------------|----------------------|
| 文件名 | (用户自主命名).TXT | 0701-0705 横梁水温.TXT |
| 内容 | YYYYMMDDmmss (采集时间) + Tab + 数据 | 201207030000 6. 1236 |

表 2 ZKGD-3000 “十五”标准水温数据格式

Tab. 2 The data format of the underground water temperature in “Tenth Five-Year” recorded by the ZKGD-3000 instrument

| 项目 | 格式 | 实例 |
|-----|---|---|
| 文件名 | 台站代码 + 测项代码 + 仪器 ID + YYYYMMDD (采集日期). EPMS | 620534112ZKGD000420120703. EPMS |
| 内容 | YYYYMMDD (采集日期) 台站代码 测项代码 + 仪器 ID 采集周期 通道个数 测项代码 观测数据值…… | 20120703 62053 4112ZKGD0004 01 01 4112 6. 1236… |

4 总结与讨论

本研究工作以非标仪器的数据转换入库为出发点，在总结归纳“十五”前兆台网运行流程的基础上，开展相关数据转换实施方案的工作。处理对象可以涵盖各类非标仪器的地震前兆观测数据。数据转换方案具有一定的拓展前景，同时也

有不足之处，有待在今后的研究工作中进一步完善和改进。

(1) 数据的完整性和一致性校验

完整性和一致性校验是数据转换的后续工作，是转换后数据的准确性和可靠性的保证。数据校验可在数据转换后、上传前完成，数据若通过完整性和一致性校验则正常上传，否则重新实施转换。数据校验涉及到具体的数据内容和格式，因

此亦可采用动态编译的方式实现。

(2) 数据预处理

实施方案设计的数据转换软件, 对处理算法采用了类似于插件的运行模式, 所以对数据对象和处理方法都具有较高的兼容性。各种数据对象的处理算法经过适当改写, 都可运行于软件中。因此本工作研发的数据转换软件, 除了实现数据转换的功能以外, 还可向多种数据预处理功能拓展, 包括异常值识别、仪器工作日志信息提取、文件格式转换等。

(3) 仪器控制功能

数据转换方案的实现首先必须依赖于仪器产生的原始数据已正确传输至数据服务器, 而部分仪器必须由客户端向其发送指令, 才会传输数据。这导致数据转换工作的自动化流程受到该操作延迟的影响, 因此有必要研究仪器客户端的数据采集接口或命令, 并将其集成于数据转换软件中以实现仪器自动控制功能。

参考文献:

- 崔慧敏, 戴桂兰, 王生原, 等. 2004. 动态编译技术研究[J]. 计算机科学, 31(7): 113-117.
- 何案华, 赵刚, 王军, 等. 2008. 串口前兆仪器的因特网接入方案与配套软件开发[J]. 地震研究, 31(3): 279-283.
- 黄晖. 2008. “十五”前兆观测模式之嬗变及其应对[J]. 华南地震, 28(2): 103-107.
- 马文娟, 何案华, 曹开, 等. 2010. “九五”前兆仪器与“十五”前兆管理系统的整合[J]. 地震研究, 33(4): 361-364.
- 王军, 赵刚, 何案华, 等. 2008. “九五”前兆台站与“十五”前兆台网的整合研究[J]. 大地测量与地球动力学, 28(4): 125-130.
- 王秀英, 周振安, 牛从达, 等. 2004. 高精度数据采集器网络通讯功能的实现[J]. 地震研究, 27(2): 203-207.
- 杨满栋, 李国峰, 姜立新, 等. 2000. 中国地震前兆台网中心数据信息网络服务系统[J]. 地震, 20(增刊): 211-216.
- 中国地震局. 2005a. 中国地震前兆台网技术规程[M]. 北京: 地震出版社.
- 中国地震局. 2005b. 中国地震信息服务系统技术规程[M]. 北京: 地震出版社.
- 周克昌, 张崇立, 纪寿文, 等. 2009. 中国地震前兆台网主要问题探讨[J]. 地震地磁观测与研究, 30(1): 76-80.

Research and Application of Transformation Scheme for Non-standard “Tenth Five-Years” Precursory Data

TIAN Ye¹, HAO Zhen¹, WANG Xiao-juan¹, NIU Yan-ping¹, WANG Jian-jun^{1,2},

(1. *Earthquake Administration of Gansu Province, Lanzhou 730000, Gansu, China*)

(2. *Lanzhou Base of Institute of Earthquake Prediction, CEA, Lanzhou 730000, Gansu, China*)

Abstract

We proposed an approach with flexibility and extensibility to solve the problem that parts of instruments cannot incorporate into “Tenth Five-Years” precursory network in Gansu province. The data transformation system was realized based on this approach. Using dynamic compilation technique, the system could realize the goal that any types of precursory data can incorporate into data processing process of “Tenth Five-Years” precursory data library, and it possesses higher popularization value. The data transformation system in developing transforms the non-standard data into the “Tenth Five-Years” standard one, and completes the mission that non-standard instrument parallels in “Tenth Five-Years” precursory station network automatically.

Key words: precursory network; data transformation; non-standard “Tenth Five-Year Plan” precursory data; dynamic compilation