基于大数据的配电网规划数据自动收资处理技术研究

方 兵[1]，宁光涛[1]，俞 悦[1]，张佳艺[1]，王 悦[2]

（1. 海南电网有限责任公司电网规划设计研究中心，海南 海口 570203；
2. 海南电网有限责任公司三亚供电局，海南 三亚 572000）

摘要：通过基于版本化的电气网络图模管理技术的研究，实现对存储于通用关系数据库中的版本化电网模型数据进行有效管理，使得规划系统通过自动化方式获取数据已经逐渐具备了应用条件。基于海南电网GIS平台、海量准实时数据平台和海南电网数据中心等基础数据平台，研究海南配电网规划核心大数据自动化收资管理关键技术，建设配电网规划数据管理模型，实现基础数据自动化或半自动化采集、存储和数据再定义，同时实现基础数据的智能统计分析，为配电网规划提供数据，对配电网规划过程和成果数据实行版本化的全过程管理。

关键词：大数据； 配电网模型；数据清洗；数据融合

Research on automatic data collection processing technology of distribution network planning data based on big data

FANG Bing[1], NING Guangtao[1], YU Yue[1], ZHANG Jiayi[1], WANG Yue[2]

(1. Hainan Power Grid Co.Ltd. Power Grid Planning and Design Research Center, Haikou 570203, China;
2. Hainan Power Grid Co. Ltd. Sanya Power Supply Bureau, Sanya 572000, China)

Abstract: Through the research of version-based electrical network graph management technology, the versioned grid model data stored in the universal relational database is effectively managed, and the planning system has gradually obtained the application conditions through automatic acquisition of data. Based on the Hainan Power Grid GIS platform, massive quasi-real-time data platform and Hainan Power Grid data center and other basic data platforms, the key technologies of Hainan distribution network planning core big data automation collection management are studied, and the distribution network planning data management model is constructed to realize the basic data automation or semi-automated acquisition, storage and data redefinition, to achieve intelligent statistical analysis of basic data, provide data for distribution network planning, and implement versioned whole process management for distribution network planning process and achievement data.

Key words: big data; distribution network model; data cleaning; data fusion

0 引言

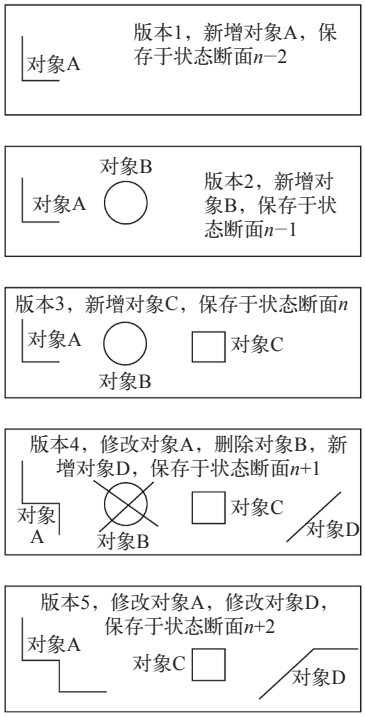
近年来，随着我国电力系统向高度信息化、自动化方向发展，电力数据的采集运用越来越广泛，对电力设备设施数据、用户数据和规划数据等的管理也提出了较高的要求。根据数据来源的不同，可以将配电网数据分为两类[1-2]：一类是配电网内部数据，主要来源于电力信息采集系统、配电管理系统、设备检测和监测系统等；另一类

是外部数据，主要来源于地理信息系统、公共服务部门等。另根据数据内容和采集时间节点的不同，也可以将配电网数据分为两类[3-4]：一类是基础数据，包括规划设计数据（如电力设备位置、回路、建设时间和厂家等）和配电网资源数据（如电缆地下管网相关信息），上述数据均为在配电网建设时产生的历史数据；另一类是状态数据，包括设备运行日志数据、监测数据等。

而对于配电网规划数据的采集，主要来源于调度自动化、生产管理、电网GIS和营销业务(SG186)等多个专业系统，数据量大，来源复杂。只有通过数据采集和智能化处理服务才可以最大限度地将有关信息集成起来，从而为电力系统决策人员提供一个多元化的决策依据[5-6]。而常规的数据采集和智能化处理服务效率较低，存在误差。基于此，本文通过对版本化的电气网络图模管理技术进行研究，实现对存储于通用关系数据库中的版本化电网模型数据的有效管理，使得规划系统通过自动化方式获取数据，并逐渐具备了应用条件。通过对海南电网GIS平台、海量准实时数据平台和海南电网数据中心等基础数据平台的建设和分析，进一步研究海南配电网规划核心大数据自动化收资管理关键技术，构建大数据管理模型，对基础数据进行智能化分析，从而通过大数据技术实现，对配电网数据的自动收资，为海南配电网的规划提供理论依据。

1　基于版本化的电气网络图模管理技术

研究基于版本化的电气网络图模管理技术，其核心是实现对存储于通用关系数据库中的版本化电网模型数据的有效管理，并能够获得快速加载的方法。在该方法中，首先使用建立电网状态断面的定义，电网模型的修改历史过程从初始状态断面1开始，每次对电网模型的修改都被记录在一个状态断面中，依次为状态断面2、3、…、$n-3$、$n-2$、$n-1$、n、$n+1$、$n+2$。例如在状态断面$n-1$的基础上，电网模型修改集合n就保存在状态断面n中。电网模型的修改是采用增量记录的形式保存在电网模型数据表中。电网模型处理示意图如图1所示。



（a）在状态断面中新建、修改和删除对象的示意图

电网模型数据表

对象内码（主键1）	对象内码（主键2）	修改标志	对象属性
A	$n-2$	新建	
B	$n-1$	新建	
C	n	新建	
A	$n+1$	修改	
B	$n+1$	删除	
D	$n+1$	新建	
A	$n+2$	修改	
D	$n+2$	修改	

（b）在数据库表中的存储方式的示意图

图1　电网模型处理示意图

Fig.1 Schematic diagram of grid model processing

基于电网模型数据库的版本化存储和加载优化方法进行设计，其示意图如图2所示。研究基于预先生成的基线主键记录集和动态生成的增量主键记录集结合的方法，实现快速版本化的电网模型数据。使用该机制，系统可以保存所有电网版本断面的完整数据，保证每个版本的网络拓扑、设备参数和相关属性都是通过历史版本链增量新建、修改和删除获得的和可追溯的。

通过电网图模的版本化管理，软件能够实现规划的多年滚动修编以及规划管理中的多年度多方案管理方式。通过系统底层的版本化数据管理，程序能够实现不同方案间的网架对比、规划计算结果对比等功能。

目前海南电网GIS平台建设已经完成，中压配电网数据录入工作也已经完成，数据质量提升工作逐步开展，数据可用性正在提高。规划系统通过自动化方式获取数据已经逐渐具备了应用条件，通过电网GIS平台同步网架及设备参数数据，支持省级规模的工程化应用。网架数据自动获取

框图如图3所示。规划系统定期自动或人工启动任务，获取电网GIS平台中的网架数据作为现状电网断面，经过自动化简、美化、人工确认和数据二次定义形成面向规划业务的现状电网数据模型。

2 基于大数据的配电网规划数据自动化收资技术

通过在线数据采集、不良数据辨识与清洗以及多源数据融合，实现深度提取PMS、计量自动化和调度SCADA等系统的海量电网数据，解决数据分散、异常数据异构等问题，构建统一的规划业务大数据库。图4为电网规划大数据收资、应用

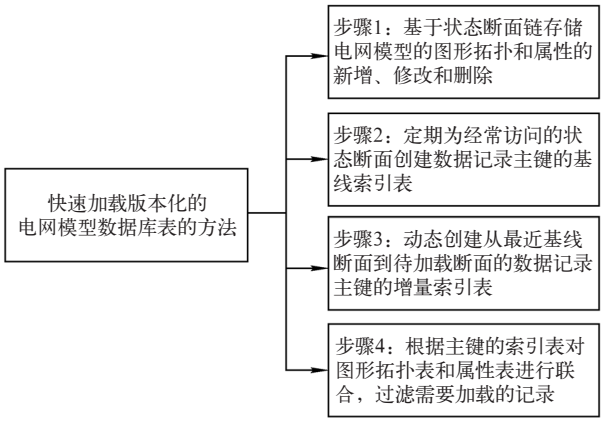


图 2 电网模型版本化示意图

Fig.2 Schematic diagram of grid model versioning

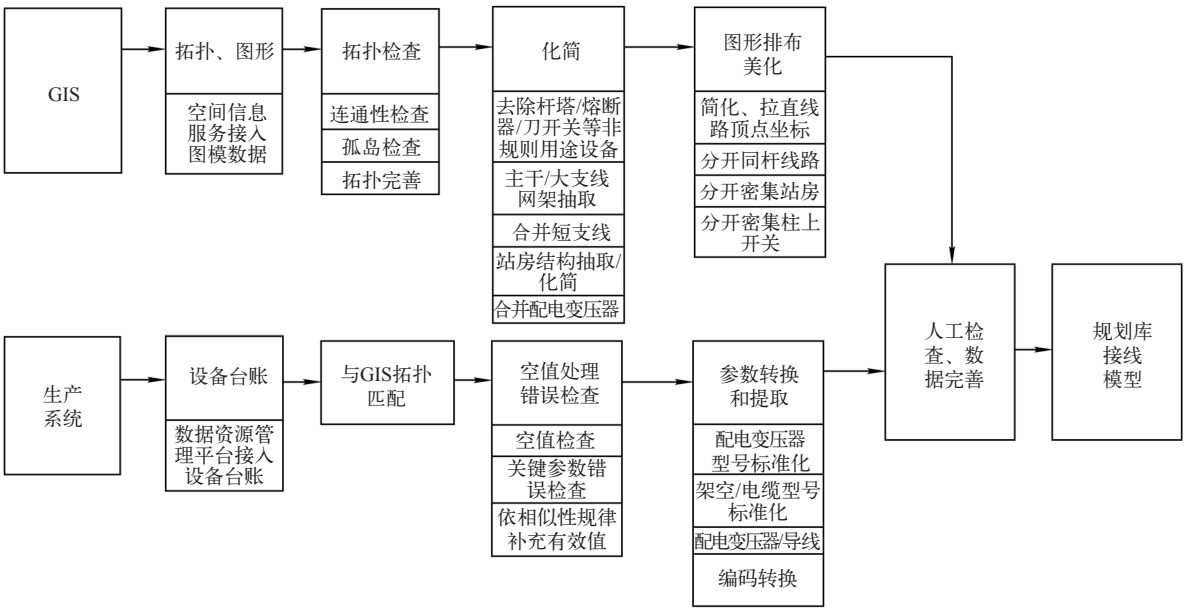


图3 网架数据自动获取框图

Fig.3 Grid data automatic acquisition block diagram

示意图。基于大数据技术的数据自动化收资技术为配电网规划业务应用提供了数据支撑。

2.1 数据校核与清洗

按照配电网数据集成规范，将数据转换成格式统一的结构化数据，并进行数据识别与纠错。数据识别与纠错方法如图5所示。

首先将转换后的数据和预定义的校验规则内存化，然后交由基于Disruptor和EDA为核心的校验引擎检查数据的正确性、一致性和逻辑性。数据校验分为3个环节。

2.1.1 完整性校验

依据数据集成规范，验证已抽取数据中的关键数据项是否有数据。以某35 kV变电站数据校验为例，变电站名称、设备编码、电压等级、供电企业、主变压器台数、主变压器容量和最大负荷等字段均为关键数据项。若存在为空的情况，则返回修改原数据后再进行校验；若均不为空，则校验通过，转入下一步正确性校验。

2.1.2 正确性校验

检查已抽取的数据项、数据类型、准确度和

图4 电网规划大数据收资、应用示意图

Fig.4 Schematic diagram of big data collection and application of power grid planning

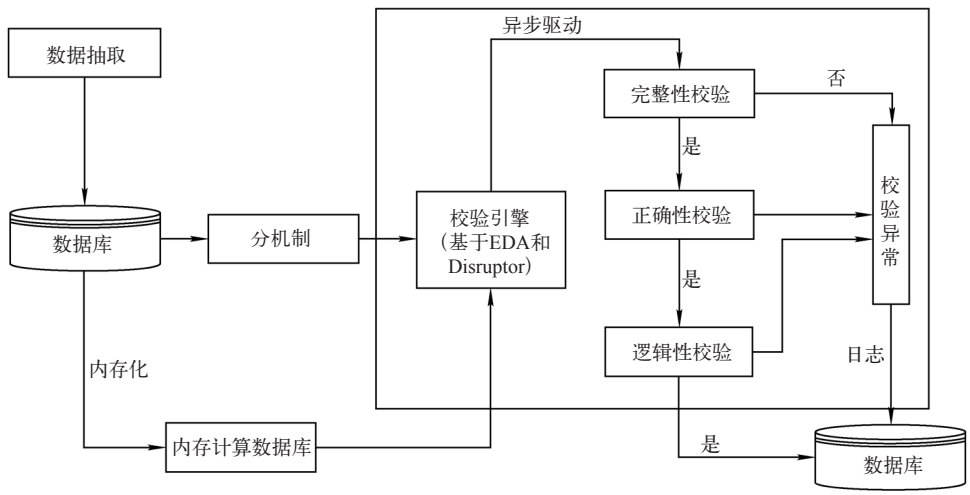


图5 基于Disruptor框架的数据识别与纠错过程

Fig.5 Data recognition and error correction process based on Disruptor framework

范围是否与集成规范要求一致。校验单元格内容类型符合要求，包括字段类型为数字实际值是否为数字、字段类型为时间实际值是否为时间。

1）单元格内数字是否在设置范围内，判断类型包括大于、小于、等于、不等于、大于等于、小于等于、介于和未介于。

2）单元格内容不可与其他单元格内容重复，包括设备ID及设备名称。

3）单元格内类型为枚举的，其内容是否在枚举范围内，如用户类别：01 高压，02 低压非居民，03 低压居民。

2.1.3 逻辑性校验

依据集成系统之间的业务关系，检查数据项是否符合业务逻辑关系。如某主变压器最大负荷超出额定容量的200%，则可判定其最大负荷和额定容量两个数据必有一个错误，需返回修改原数据后重新进行校验。具体规则如下。

1）设备从属关系校验。包括主变压器其所属

变电站、馈线所属变电站、站房内设备（负荷开关、母线）、所属站房（配电室、环网柜）和柱上设备（柱上变压器、故障指示器）所属杆塔。

2）分项值小于累加值，包括售电量表中月售电量≤累计销售电量；单项值小于总值，包括发电量中全口径水电发电量≤全口径发电量。

2.2 多源数据融合

以电网拓扑中资源设备为核心，通过设备生产ID与设备台账参数、变电站、线路及配电变压器的运行数据进行融合，形成规划业务综合数据库。图6为多源数据融合示意图。

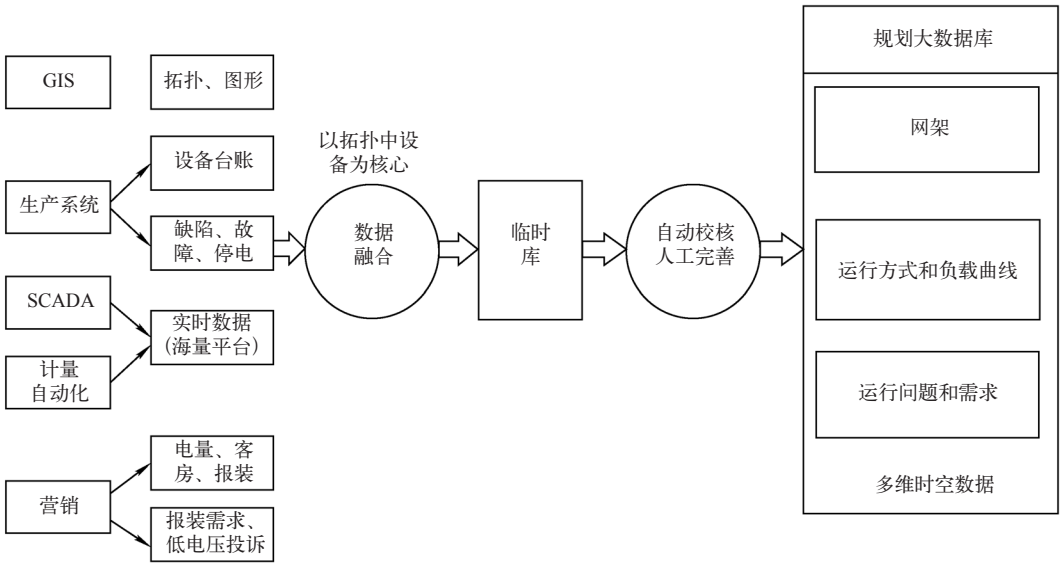


图6 数据融合示意图

Fig.6 Schematic diagram of data fusion

运行数据主要通过海量数据平台定期在线采集所需的基础数据，经过分析处理自动形成全年最大、最小以及一年四季的平日、周末和节假日典型日运行方式下的线路及配电变压器负荷曲线、负荷分布，作为配电网规划业务应用的历史运行基础数据。运行数据自动获取框图如图7所示。

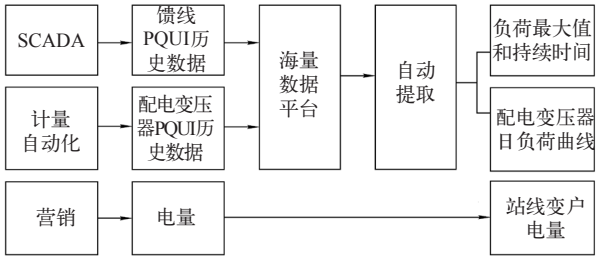


图7 运行数据自动获取框图

Fig.7 Operation data automatic acquisition block diagram

3 结束语

本文分析了基于版本化的电气网络图模管理技术，实现对存储于通用关系数据库中的版本化电网模型数据的有效管理，使得配电网规划系统数据的自动化采集成为了可能。基于原有的海南电网GIS平台、海量准实时数据平台和海南电网数据中心等基础数据平台，对基于大数据技术的配电网规划数据自动化收资技术进行总结，分析数据校核与清洗、多源数据融合的方式，从而更好地确保了配电网规划数据自动收资的准确性，为海南电网的长远规划奠定基础。

参考文献

[1] 王喆，归三荣，李扬，等．电改新形势下配电网规划不确定因素分析[J]．电力系统及其自动化学报，2017，29(4)：110-114．
WANG Zhe，GUI Sanrong，LI Yang，et al．Analysis of uncertain factors in distribution network planning under the condition of electric power reform[J]．Journal of Electric Power Systems and

Automation，2017，29(4)：110-114.

[2] 陈敬德，盛戈皞，吴继健，等．大数据技术在智能电网中的应用现状及展望[J]．高压电器，2018，11(1)：35-43.
CHEN Jingde，SHENG Gehao，WU Jijian，et al. Application status and prospect of big data technology in smart grid[J]. High Voltage Apparatus，2018，11(1)：35-43.

[3] 谭明丽．电力行业大数据应用建设及配网停电优化的研究[J]．自然科学(全文版)，2016，24(9)：115-115.
TAN Mingli. Study on big data application construction and distribution network power loss optimization in power industry[J]. Natural Science(Full-text version)，2016，24(9)：115-115.

[4] 刘苑红，张伟，崔艳妍，等．大数据在用户供电可靠性预测评估中的应用[J]．电力信息与通信技术，2016，35(3)：55-59.
LIU Yuanhong，ZHANG Wei，CUI Yanyan，et al. Application of big data in user power supply reliability prediction and evaluation[J]. Electric Power Information and Communication Technology，2016，35(3)：55-59.

[5] 郭壮志，徐其兴，洪俊杰，等．配电网快速高容错性故障定位的线性整数规划方法[J]．中国电机工程学报，2017，54(3)：786-795.
GUO Zhuangzhi，XU Qixing，HONG Junjie，et al. Linear integer programming method for fast fault-tolerant fault location in distribution network[J]. Proceedings of the CSEE，2017，54(3)：786-795.

[6] 张沈习，程浩忠，邢海军，等．配电网中考虑不确定性的分布式电源规划研究综述[J]．电力自动化设备，2016，34(8)：1-9.
ZHANG Shenxi，CHENG Haozhong，XING Haijun，et al. Summary of distributed power supply planning considering uncertainty in distribution network[J]. Electric Power Automation Equipment，2016，34(8)：1-9.

作者简介

方兵（1983－），安徽宿州人，男，工程师，本科学历，主要研究方向为智能电网和能源规划。

(收稿日期：2018－07－03)EA

方兵