

一种基于BI的电网停复电事件分析方法

介绍了一种借助商业智能BI提供的ETL工具抽取电网一次设备如线路、负荷等的相关量测信息的方法，结合一次设备量测变化情况和相关遥信变位变化情况判断设备的停复电情况。此方法充分利用了电网调度主站系统中的海量历史数据，通过BI数据抽取分析结合电网运行经验，给出了基于历史数据的停复电事件分析结果。此外，还分析了影响统计结果的环节和因素，给出了解决问题的办法和建议。

杨启京¹ 孟勇亮¹ 杨立波² 季堃¹ 尹寿垚¹ 翟毅¹ 张留留¹/1.国网南瑞科技股份有限公司 2.国网河北省电力公司



杨启京/工程师

关键词/Keywords

商业智能BI ·
ETL ·
停复电事件 ·

0 引言

随着国网大运行体系建设的不断深化和电网的不断扩大,对电网调度管理规范化、一体化及精益化提出了更高的要求^[1],在这样的大背景下,各省调控中心陆续开展了省地一体化的系统的建设。这些系统的建设实现了区域35 kV以上全电网设备的实时监控,同时也为后续基于全电网数据分析应用打下了良好的基础。在过去很长一段时间,调度主站系统的实时采集的量测数据、告警信息等信息更多是为了实时监控、汇总统计及简单展示所用,对于存储在数据库中的海量历史信息的深层利用较少。面对这一发展趋势,本文结合河北省地一体化模型中心系统建设应用实际需求和模型中心省地海量历史数据,引入了商业智能BI的概念^[2],借助商业智能BI提供的ETL^[3]工具对电网历史数据的抽取,过滤,分析开发出了一种基于海量历史数据的电网停电事件分析方法。

1 商业智能BI与电网数据分析

1.1 BI概念

BI是Business Intelligence的英文缩写(中文解释为商务智能)是用来帮助企业更好地利用数据提高决策质量的技术集合,也是从大量的数据中钻取信息与知识的过程。BI的关键是从许多来自不同的数据源的数据中提取出有用的数据并进行清理,以保证数据的正确性,然后经过抽取、转换和装载,合并到一个的数据仓库里,从而得到数据的一个全局视图,在此基础上利用合适的查询和分析工具、数据挖掘和OLAP工具等对其进行分析和处理,最后将知识呈现给管理者,为管理者的决策过程提供支持^[4]。

在电力系统行业外,商业智能BI系统越来越多地被应用于各个热门领域,特别像互联网行业、电信及银行等行业商业智能BI技术得以大量的应用,在商业智能BI应用中,预测、分摊、假设模拟及数据挖掘等交互式技术成为未来的发展方向,越来越被业界所重视。从行业外的经验来看,企业商业智能BI所应用、依赖的信息和数据大部分来源于企业内部业务系统中的长期积累的海量数据材料,并借此进行分析加工、做出决策得出分析结果^[5-8]。

1.2 ETL概念

ETL,是英文 Extract-Transform-Load 的缩写,用来描述将数据从来源端经过抽取(extract)、转换(transform)和加载(load)至目的端的过程。其目的是将分散、零乱及标准不统一的数据整合到一起,为决策提供分析依据^[9-10]。ETL是BI中重要的一个环节。

1.3 大型电网数据分析

大型电网由于规模大的其数据分析一直是一个难题。以河北南部电网为例，全网拥有500 kV变电站15座，220 kV变电站179座，110 kV变电站757座，35 kV变电站1 585座。每天几千座变电站所采集的历史数据就有上百GB的容量。如果针对这种大型电网对全网设备进行统计分析，如采用常规接口方式效率低，系统负荷重，难以进行有效的数据分析^[11]。通过引入BI，将BI在数据分析领域的优势与电网数据分析相结合，通过ETL从海量模型和历史数据中抽取有效信息进行分析，可大大提高数据分析效率。

2 基于历史数据的设备停电判定方法

2.1 基于历史数据的判定规则

在电网一次设备运行监视中，判定设备是否带电的一个重要依据就是看设备的电流值是否归零。在数据采集系统正常运行的情况下，设备电流如果为零，那么即判定该设备处于失电状态，如果设备电流值非零则判定该设备处于带电状态。根据这个重要的判断条件，如果将设备一段时间内连续的电流值加以分析，可以获得设备停电时刻和复电时刻，在本文中将设备停电时刻发生的事件定义为停电事件，将复电时刻发生的事件定义为复电事件。

2.2 数据状态的连续性判定

考虑到调度系统中数据抖动现象很难避免，在一段时间内获取的电流量测历史数据可能会有数据抖动的问题，因此在判定设备由一个状态变化为另一个状态时，就不能简单地由两个相邻测点的值来进行判断（这样判断的结果是会产生大量误报信息），因而需要通过多个连续点的变化情况来进行判断。

如图1、图2所示的分别是停复电情况下电流随时间变化的曲线。其中当前时刻是以 T_1 表述，无论是停复电，其后续 $T_1 \sim T_5$ 点的数据状态必须是连续的，即对于复电来说是连续大于零，对于停电事件是连续的零值。通过数据状态的连续性判定停复电事件的有效性。依据上述判断依据从给定事件段中可分析出符合条件的有效停复电事件。

此外还有一种情况连续的数据状态——电流

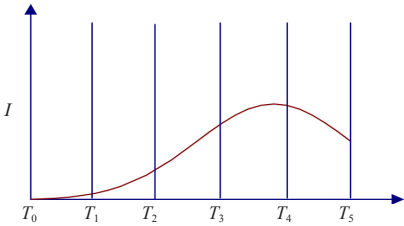


图1 复电曲线

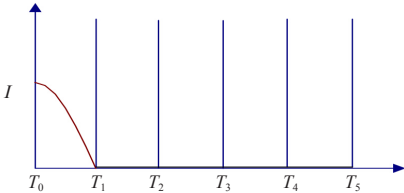


图2 停电曲线

恒为0，此状态下主要是一些长期停电的设备所具备的状态。在判定时通常根据统计所用的时间周期，获取周期内的电流最大值，根据最大值是否为0，判定设备状态是否在该周期内恒为0。在进行数据状态的连续性判定时，通常先进行是否电流恒为0的判定，再进行复电和停电状态的判断。

2.3 前—后五原则

上文描述了对于停复电事件判断要求当前时刻 $T_1 \sim T_5$ 五个点的数据状态必须是连续的。对于当前时刻前状态的判定，考虑到停复电事件发生前如果也根据数据状态连续性进行判断，对于事件发生前出现数据抖动的情况，则可能会出现漏掉事件发生前有数据抖动的有效事件的问题，鉴于上述情况，对于事件发生前的数据状态可仅用一个点的状态即 T_0 来判定， $T_0 > 0$ 且后续状态连续为0，则为停电事件， $T_0 = 0$ 且后续状态连续大于0，则为复电事件。因此，每一个事件判定都是包含当前时刻在内的前一个点和后五个点数据判断而来，该数据选取规则成为前—后五原则。

2.4 停复电事件关联性判定

设备的停电复电是有规律的，同一个设备量测由零变为非零（复电），在这之前必然有量测有非零变化为零（停电）的过程。依据上文介绍的事件判定原则，可分析出大量的停复电事件，而这些事件之间是有着关联关系的即同一设备停、复电事件是交替进行的。也就是说在理想情况下对于同一个设备本次如发生了停电事件，那么该设备上一次发生的事件应该是复电事件，反之亦然。当然考虑到判定统计的时段以及历史数

据的采集总是有时间限制的，因此对于判定统计中不可能所有事件都找到上次对应状态，特别是统计计算初期，但对于一个较长统计时间如一年以上，设备经历了一个停复或复停状态后，这种关联性判定还是非常准确和有效的。此外此种关联性判定还可以获取同一设备的停电，送电起始终止时刻，从而统计出停电时长。

2.5 遥信变位过滤原则

在实际电网运行过程中，一次设备（如线路，主变）的停复电必然伴随着其相关的开断设备的分合变化。对应着主站量测信号，即必然有相关的遥信变位信号产生。根据此原则，结合上文判断出的停复电事件，根据对应设备相关遥信在事件发生前后是否有遥信变位这一判据可增加事件发生的有效性。

基于遥信变位过滤，还可以继续沿着这一思路对一次设备停复电事件发生做扩展性判断，即根据一次设备关联遥信判断该遥信是否在事件发生前后有遥控操作。获得上述信息对于丰富停电事件分析和展示有非常重要的帮助。

3 实施方案

3.1 河北模型中心系统概述

河北模型中心系统采用的是现阶段国内应用范围最广的调控系统D5000系统，通过模型服务汇集了省内各地调的一次设备模型，通过消息转发获取地调实时数据，在模型中心一区通过分钟级采样存储全省的历史数据，并通过隔离装置将模型和数据同步到三区，在三区通过BI服务器，对历史数据进行分析统计并将结果写入OMS展示结果库。系统中历史数据存储采用的是国产时序数据库海迅数据库，模型存储采用的是国产数据库金仓库，BI分析统计结果存储采用的是ORACLE数据库。数据流程如图3所示。

从上文可了解河北模型中心系统结构复杂，设计的数据库类型和系统众多。针对停复电统计分析传统的实现方法会带来较多的编码工作量，或针对不同数据库消耗大量的系统资源进行历史数据抽取及模型分析工作。本项目中采用开源BI软件，结合电网调度系统特性，在开源软件上进行了二次开发。本项目所用BI软件支持多种数据

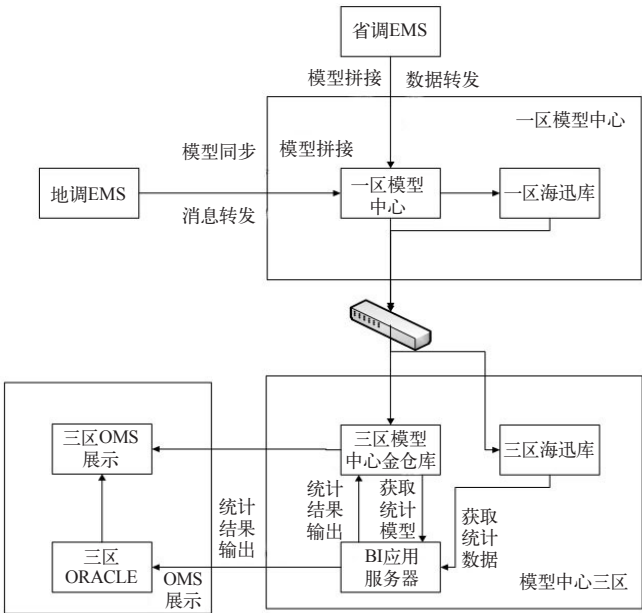


图3 河北模型中心数据流程图

源，能够以开放数据库互连（ODBC）、Java数据库连接（JDBC）等标准接口读写金仓、达梦、海迅及Oracle等数据库，能够处理多种文件格式，如E格式^[12]、Text和Excel等常规文件，能够实现关系型模型数据到多维数据模型的转换，并且支持图形方式创建和浏览ETL过程，能够直观地监视ETL全过程，并能快速定位发生错误的位置^[13]。

3.2 BI停复电设备模型设计

在进行设备停复电统计前需要对统计模型进行设计，以方便后续统计分析、展示。在构建停复电设备模型时，针对设备的厂站、区域、电压、间隔、参数及采样定义等存储在不同关系表中的属性根据设备ID进行了ETL抽取，将这些属性融合在停复电设备统计模型表中，该模型表中主要属性如图4所示。

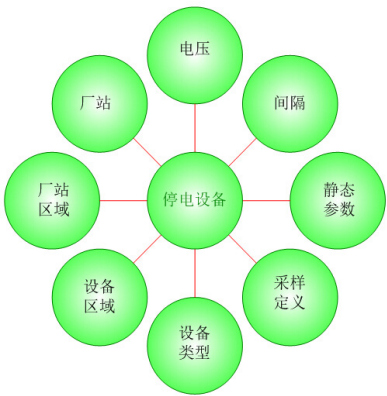


图4 停复电设备模型

通过ETL工具可以方便地从金仓模型库中将上述需要用到的属性抽取出来，后续扩展新的属性也非常的方便，只要在界面中定义相关数据字段即可。

3.3 BI停复电事件统计

根据上文中介绍的停复电事件判定方法，分别设计了停复电事件表、停复电事件关系表用以存储BI分析时生成的中间结果。停复电事件表用来存储根据停复电模型表中定义设备量测基于停复电事件判定规则分析出的结果，凡是符合判定规则停复电事件都存储在该表中。停复电事件表中存储了设备ID，事件发生时间，事件类型（停电、复电或恒为0），是否有变位信息，是否有遥控信息。停复电事件关系表是基于停复电事件表生成的，根据停复电事件表中记录按设备ID扫描

停复电事件表，查找与设备本次事件对应的上次事件并记录下上次事件ID，事件发生时间等。

通过BI中的ETL软件根据“模型→数据→事件”这一链路从海迅数据库和金仓数据库中抽取数据并将数据串联起来形成内存中的数据流，数据流在BI提供的各转换环节流动，获取通过转换得到的数据属性，最后将有效的事件结果数据写入ORACLE数据库结果表中，供后续查询展示使用。

3.4 BI停复电结果分析

根据停复电事件统计结果，可以很容易地得到许多有用的信息，如某一时刻的停电设备数（见图5）、某一时刻发生停电的事件数（见图6）以及某一时刻停电设备信息（见图7）等。通过曲线，列表等展示方式可以很直观地将一天中停电设备的数据以不同角度展示出来。

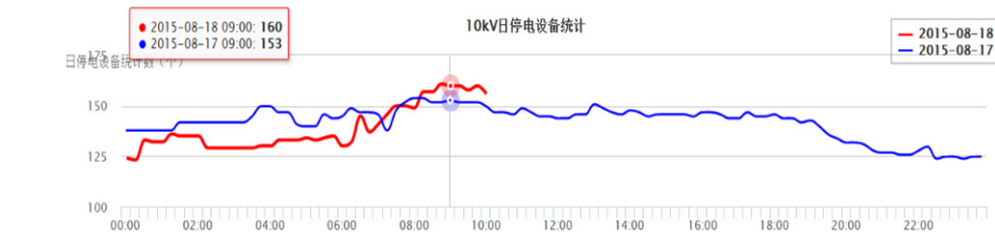


图5 停电设备数曲线

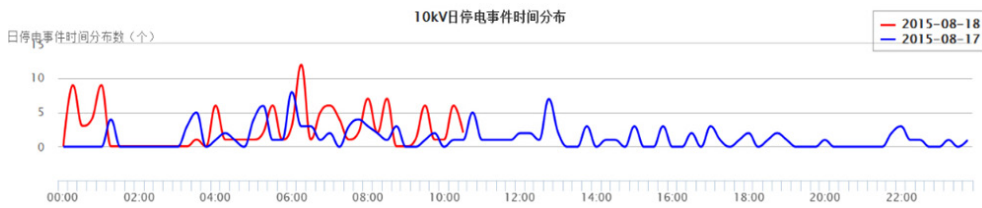


图6 停电事件数曲线

序号	地区	单位	厂站	设备名称	类型	停电时间	送电时间	停电时长 (小时)
1	衡水	武强县	35kV.李屯站	10kV.李吉线587线路	线路	2015-08-18 08:42	2015-08-18 09:42	1.0
2	保定	雄县	35kV.双堂站	10kV.程岗线513线路	线路	2015-08-18 08:38	2015-08-18 09:58	1.3
3	石家庄	井陉县	35kV.雪花山站	10kV.西河线595	线路	2015-08-18 08:32		3.1

图7 停电设备信息

4 影响统计结果因素分析

4.1 遥测跳变

某一时刻遥测数据超过定义的合理范围称为遥测跳变^[14]。在项目测试过程中发现在统计设备停复电所用的历史数据时，经常会遇到因遥测跳变造成统计结果不准确的情况。出现这类情况的原因主要是因为站端采集装置的问题，站端采集装置存在故障，送出的数据就会经常发生异常

变化。还有就是远动传输方面的因素，远动网络通道稳定性对于遥测采集也有一定的影响。因此对于调度系统来说想要百分百杜绝数据跳变的现象，基于现有技术难度还很大。那么通过BI中的数据清洗过滤功能，通过设定阈值过滤掉不合理的遥测数据，可对遥测跳变的情况做出适当的处理。

4.2 零漂值设定

零漂，是零点漂移^[15]（Zero drift）的简称，对

于一次设备电流零漂值来说,可能由于温度的变化、电源电压的不稳定、采集装置精度等原因造成设备失电情况下电流不为零的情况。因此在判定设备停复电时通常比较的零值不是绝对零值,而是零漂值。零漂值的设定是否合理对于统计电网设备停复电事件的结果影响是非常大的。因此对每类乃至每个设备设置其合理的零漂值对于提高统计的准确性是非常必要的。

4.3 特殊厂站

测试中还发现依据上文中所述的停复电判定规则,在某些特定厂站(诸如用户站、高铁站),一天中会发生多次停复电的情况。结合这些厂站中设备电流曲线分析以及厂站运行情况分析,此类厂站负荷变化特性与本文介绍的判据恰好吻合。比如某高铁站,在没有列车经过该站控制区时负荷很低,电流很小接近于零;当有列车经过时,负荷瞬间拉升,电流变大;当列车驶离厂站控制区域时,负荷变小,电流归零。某些用户站负荷变化情况与高铁站类似。这些厂站负荷变化情况与本文中描述的设备停复电规则基本吻合,一旦发生了此类负荷变化情况,就有可能发生误判事件。面对这一问题,结合BI中大数据分析的优势,对于此类特殊厂站可以通过增加对设备关联遥信在事件发生前后是否动作,设备关联设备如主变对应绕组量测在事件发生前后是否产生波动等条件来判定停复电事件是否有效。

5 结束语

本文介绍的基于BI的电网停电事件分析方法充分利用了电网系统中海量历史数据,通过与商务智能BI相结合,可从海量数据中获得有用信息,并能对影响统计结果的因素做出有效处理,为电网监控运行分析提供了新的思路和方法。本文研究成果已经在河北模型中心系统得到具体应用,并取得了很好的效果。基于此研究成果可以开展更多基于BI的电网运行统计分析功能的开发研究。

参考文献

- [1] 陈树勇,宋书芳,李兰欣,等.智能电网技术综述[J].电网技术,2009,33(8):1-7.
- [2] 夏国恩,金炜东,张葛祥,等.商务智能在中国的现

状和发展研究[J].科技进步与对策,2006,23(1).

- [3] 张宁,贾自艳,史忠植,等.数据仓库中ETL技术的研究[J].电网技术,2002,38(24).
- [4] 北京大学光华管理学院.商务智能的理论和应用研究综述[J].计算机系统应用,2007,(1).
- [5] 荀挺,张珂珩,薛浩然,等.电网调控数据综合智能分析决策架构设计[J].电力系统保护与控制,2015,(11).
- [6] 张艳杰. BI技术在国华电力企业信息门户系统中的应用[J]. 电力技术,2010,19(4):22-25.
- [7] 黄杰,孙莉.商业智能在电力公司决策支持系统的研究与应用[J].计算机应用与软件,2010,27(1):221-223.
- [8] 李秋丹,迟忠先,金妮,等.基于数据仓库的地区电网调度决策支持系统[J].电力系统自动化[J],2004,28(12):86-89.
- [9] 段成,王增平,吴克河,等.一种轻量级电网实时数据ETL系统的设计与实现[J].电力系统保护与控制,2010,38(18).
- [10] 宋杰,郝文宁,陈刚,等.基于MapReduce的分布式ETL多维数据模型研究[J].计算机科学,2013,40(z2).
- [11] 宋亚奇,周国亮,等.智能电网大数据处理技术现状与挑战[J].电网技术,2013,37(4):927-935.
- [12] 辛耀中,陶洪铸,李毅松,等.电力系统数据模型描述语言[J].电力系统自动化,2006,30(10):48-51.
- [13] 宋鑫,郭骏,尹寿垚,等.商务智能在电网调度控制系统数据分析中的应用[J].电力系统自动化,2015,12(11):43-46.
- [14] 张鹰,楼平,卢黎明,等. SCADA系统遥测数据"跳变"现象的原因分析与对策[J].华东电力,2009,37(12):2112-2114.
- [15] 李振华,王忠东,卢树峰,等.基于空心线圈互感系数自校验原理的大电流校验系统[J].高电压技术,2014,40(8):34-38.

作者简介

杨启京(1982-),男,本科,工程师,主要从事IEC 61970标准在电力系统的应用开发工作。

(收稿日期:2015-12-10) **EA**

(编辑:杨得润)