

光伏数据可视化系统研究与实现

王 明, 高 翔, 文 峰

(中车株洲电力机车研究所有限公司, 湖南 株洲 412001)

摘 要: 随着光伏电站智能化水平的不断提升, 现有光伏智能监控系统可视化方法已无法满足高实时性、高并发的数据可视化需求。为此, 文章探讨了实现智能化光伏电站的关键技术, 并针对大数据的光伏数据可视化方法进行了分析与研究, 提出了一种 Kafka, Spark 与 WebSocket 相结合的数据处理方法, 可满足高并发数据请求下光伏实时数据的可视化需求。利用该数据可视化方法, 可有效监控光伏元件的运行状况, 有利于光伏电站的智能化管理、提升发电效率、节省维护管理所需要的时间与人力成本。

关键词: 光伏; 大数据; 数据可视化; 数据推送; 智能化

中图分类号: TP319; TM615

文献标识码: A

文章编号: 2096-5427(2018)02-0052-04

doi:10.13889/j.issn.2096-5427.2018.02.012

Research and Implementation of Data Visualization System for Photovoltaic Power Plants

WANG Ming, GAO Xiang, WEN Zheng

(CRRC Zhuzhou Institute Co., Ltd., Zhuzhou, Hunan 412001, China)

Abstract: With the continuous improvement of intelligent level of photovoltaic power station, existing visualization methods of photovoltaic intelligent monitoring system can not satisfy the near real-time and concurrency data visualization needs. In this paper, photovoltaic data visualization method based on big data was analyzed and studied. Meanwhile, a method of data processing which combines Kafka, Spark with WebSocket was proposed and it can implement the visualization of photovoltaic real-time data under high concurrent data request. The method can effectively monitor operation status of photovoltaic components, and is conducive to manage photovoltaic power plants intelligently, improve power generation efficiency, save maintenance time and labor cost.

Keywords: photovoltaic; big data; data visualization; data pushing; intelligence

0 引言

随着全球环境污染的加剧以及资源的不断锐减, 发展新能源已经是一种不可避免的趋势^[1]。光伏发电因其资源用之不竭、清洁环保且不受地域限制, 加之建设周期短等优点, 在长期的能源战略中具有极为重要的地位^[2]。光伏电站的平稳运行, 不仅仅需要各个光伏模块的正常、高效运行, 也需要各模块协调作业^[3-4]。

如何实时监控各个光伏模块的运行状态, 以便运维人员及时作出维护、优化响应, 从而保证光伏电站的正常运行, 成为亟待解决的问题。传统光伏电站基于网络监控系统与人工运维相结合的设备管理与监控方式^[5], 对于设备故障的处理存在着一定的延时, 且在数据并发量大的情况下无法保证数据传输的稳定性, 影响系统监控管理的效率^[6-7]; 其智能管理系统受限于既有的计算及处理能力, 一般只能对光伏电站部分关键参数进行实时监控, 无法对多个光伏电站所有采样参数进行实时监控及可视化展现。为此, 本文提出了一种基于大数据的光伏

收稿日期: 2017-12-14

作者简介: 王明(1989-), 男, 工程师, 主要从事大数据平台及应用开发工作。

智能管理系统，其采用基于大数据的光伏智能管理系统架构以及实时数据处理和存储方式，可实现光伏电站高并发量数据的可视化展现。

1 光伏数据处理

该光伏智能管理系统的总体架构如图 1 所示。系统中，光伏电站的所有运行数据通过配套的 4G 无线路由器经安全可靠的 IPsec 隧道回传给大数据平台进行解析，并对光伏实时数据进行可视化处理，且支持处理的结果在台式机、笔记本电脑上或手机及 Pad 客户端展现。



图 1 光伏智能管理系统总体架构
Fig. 1 General framework diagram of the intelligent photovoltaic power plant management system

光伏数据可视化主要分为实时数据可视化与历史数据可视化。本文将从实时数据处理以及历史数据存储 2 个方面分别阐述可视化数据的收集和处理原理与过程。

1.1 实时数据处理

光伏数据可视化结果的完整性以及实时性极大程度上依赖于大数据平台接收到数据后处理数据的方式^[8]。

传统的数据处理方式是光伏电站实时数据通过 Socket 通信将数据传回至服务器端，客户端以 AJAX（异步 JavaScript 和 XML）通信方式向服务器请求数据并在 Web 浏览器或者手机 APP 中进行可视化展示^[9-10]。这种方式不能及时处理客户端多并发、多连接情况，且容易造成数据的丢失，影响可视化数据的完整性与精确性；同时由于 AJAX 通信方式的原理是不断向服务器轮询新数据，当多客户端大批量请求数据时，会导致服务器端的压力剧增，影响其服务性能。

本文所提出的光伏智能管理系统采取 Kafka（高性能跨语言分布式发布 / 订阅消息队列系统）、Spark（开源的类 MapReduce 的通用并行框架）和 WebSocket（下一代客户端 - 服务器的异步通信）相结合的方式（图 2）。为了防止客户端发生多并发、多连接情况，数据接入服务器时会将实时数据先写入支持快速持久化、高吞吐、完全分布式的 Kafka 消息队列；从 Kafka 消息队

列获取到实时数据后，为了加快数据预处理过程，采用可提供海量数据交互式查询的 Spark 分布式计算框架^[3]；为了避免客户端不断轮询服务器所产生的压力，采用 WebSocket 数据推送方式，所推送的数据将作为 Web 浏览器或手机 APP 客户端可视化数据源的实时数据。

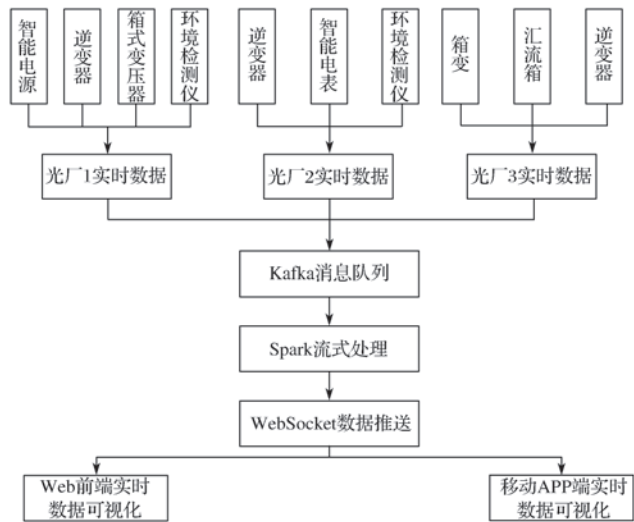


图 2 光伏智能管理系统实时数据处理流程
Fig. 2 Real-time data process flow of the intelligent power plant management system

1.2 历史数据存储

对历史数据进行可视化操作的前提是获取历史数据源。考虑到后期光伏电站扩容及存储所有实时数据的需求问题，传统的 Mysql 数据库（关系型数据库）存储数据的方式已不能满足应用要求，为此该光伏智能管理系统采取 Mysql 与 Hbase（非机构化数据库）相结合的方式进行数据存储，其数据存储流程如图 3 所示。

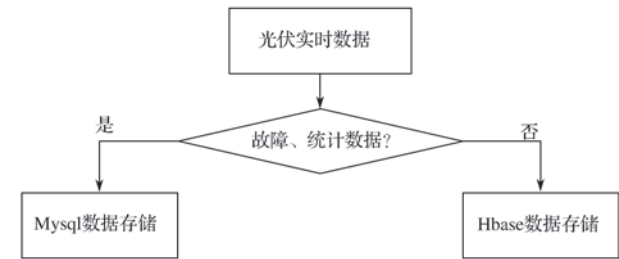


图 3 数据存储流程
Fig. 3 Real-time data storage flow of the system

如图 3 所示，服务器端接收的光伏实时数据将全部存储至 Hbase 中进行备份和高可靠性安全存储。针对非海量数据查询，Mysql 的速度比 Hbase 的快，故光伏故障与统计数据将在 Spark 中计算后存储到 Mysql 数据库中；因数据查询速度会随存储容量的增加而降低，故本系统只在 Mysql 中存储近 3 个月的故障数据，更长时间段的历史数据则需在 Hbase 中进行查询操作。

为了便于历史数据的可视化，在 Mysql 数据库中存储了所有光伏电站的发电企业信息表、设备信息表、故障码表、电站元件拓扑结构表和用户信息表等相关静态信息表。

良好的数据存储方式将直接影响后期数据查询的速度和效率。对于所有数据在 Hbase 中的存储备份，该系统采取以光伏电站编号与时间戳相结合的方式生成每个 Cell（数据存储单元）的 RowKey（Hbase 数据检索依据）值。

2 光伏数据可视化实现

所设计的光伏智能管理系统支持 Web 浏览器及手机 APP 客户端图表展现，并支持用户在图表中进行交互性操作以快速获取光伏电站各个元器件相互关联关系。

2.1 实时数据可视化

2.1.1 光伏电站关键信息实时数据可视化

系统运维人员可在 Web 浏览器端选择查看光伏电站关键信息请求，服务器端将最新的实时数据通过 WebSocket 数据通信方式推送至对应前端页面并进行可视化图表绘制。其关键信息可视化数据处理流程如图 4 所示。

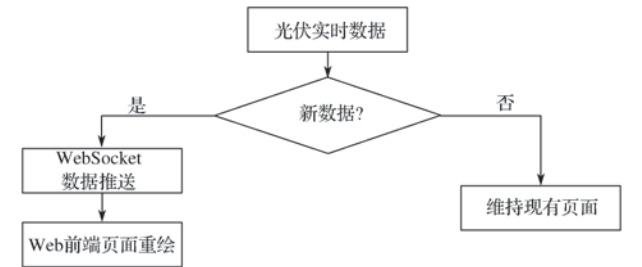


图 4 关键信息可视化数据处理流程
Fig. 4 Flow chart of key information visualization data process

光伏电站关键信息可视化效果如图 5 所示。通过该系统，运维人员可快速获取电站的日发电量、功率因数和发电效率等实时关键信息。



图 5 关键信息可视化效果
Fig. 5 Visualizations effect of key informations

2.1.2 光伏组件拓扑实时数据可视化

光伏电站各组件拓扑实时数据可视化实现过程如图 6 所示。

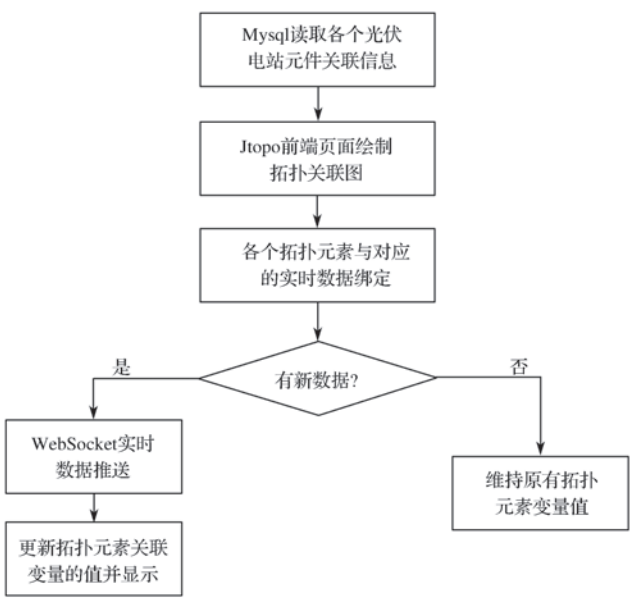


图 6 光伏组件拓扑实时数据可视化过程
Fig. 6 Real time data visualization flow chart of photovoltaic module topology

以某光伏电站的拓扑关系为例，在 Web 浏览器端获取的光伏组件拓扑实时数据可视化效果如图 7 所示。运维人员将鼠标移至所关心的拓扑元素上时，该元素的关键实时信息会以图表形式进行可视化展示；通过双击拓扑元素，可以获取该元素的所有实时变量信息值。当某个光伏组件发生故障时，其在拓扑图上所对应的元素将变成红色并闪烁，运维人员可通过双击该元素来获知故障发生的详细描述，并可作为故障处理的参考依据。

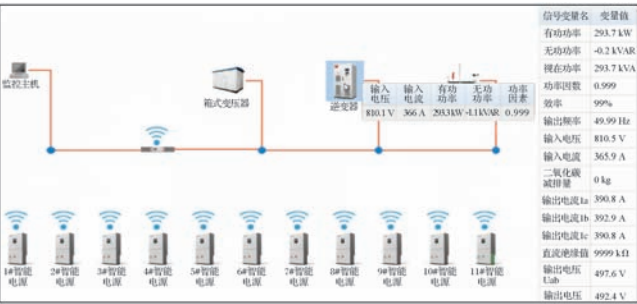


图 7 光伏组件拓扑实时数据可视化效果
Fig. 7 Real time data visualization effect of photovoltaic module topology

2.2 历史数据可视化

2.2.1 故障统计数据可视化

光伏电站历史故障统计数据对于针对性维护光伏元件正常工作有着重大意义，利用这些数据，运维人员可以集中各类资源与精力在高频故障上。故障统计可视化数据处理流程如图 8 所示，前端页面接收到服务端返回的结果后，可以将查询的结果以饼图、柱状图和折线图等方式进行多维度展示。

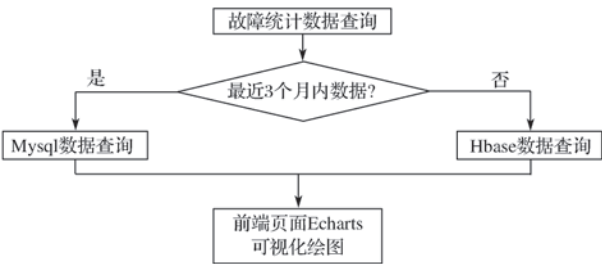


图 8 故障数据可视化处理流程
Fig. 8 Flow chart of fault data visualization process

故障统计数据可视化效果如图 9 所示，其支持用户的可交互性操作，用户可改变图中的筛选条件来查看感兴趣时间区间的故障数据，并通过设置图表展示类型来从不同维度审查故障统计可视化结果。



图 9 故障统计数据可视化效果
Fig. 9 Visualization picture of the fault statistical data

2.2.2 运营统计数据可视化

运营统计数据可直观展示各个光伏电站在不同时间的发电量，是评估光伏电站平稳运行的另一重要指标。运营统计数据可视化过程与故障统计数据可视化过程类似，判断所请求的数据是否为近 3 个月内的数据，分别从 Mysql 与 Hbase 数据库中查询获取数据，并根据应用需求进行可视化展示。其运营统计数据可视化效果如图 10 所示。用户可以根据重点关注的维度来选择希



图 10 运营统计数据可视化效果
Fig. 10 Visualization picture of the management statistical data

望查看的运营统计数据。该系统支持按日、周、月、年来对运营数据进行可视化展示，并支持用户依据不同展示维度来展现运营统计数据。

3 结语

本文所提光伏可视化系统可实现多个光伏电站高并发数据请求条件下实时数据监控及可视化展示。作为运维人员智能化管理光伏电站的有利工具，其能准确提供设备的实时运行状况，极大节省了管理和维护光伏电站所需时间及人力成本。目前，该系统的工作重点在于解决光伏产品关键数据的接收、存储及展示，后续将充分利用已有的历史数据并结合大数据相关技术对光伏电站各个元件的健康状况进行预测，以实现光伏电站的智能化管理。

参考文献：

[1] 孙李平, 李琼慧, 黄碧斌. 分布式光伏发电现状及走势[J]. 中国电力教育, 2013(6): 90–91.

[2] 张磊, 陈国恩, 王跃强. 基于相似日重组的分布式光伏电源发电功率预测[J]. 电气技术, 2016, 17(2): 62–66.

ZHANG L, CHEN G E, WANG Y Q. Power Generation of Distributed Photovoltaic Power Generation based on Similar Days of Restructuring[J]. Electrical Engineering, 2016, 17(2): 62–66.

[3] 王锋, 张永强, 才深. 西安城区灰尘对分布式光伏电站输出功率的影响分析[J]. 太阳能, 2013(13): 38–40.

[4] 马龙. 分布式光伏电站监控系统的设计与研究[D]. 南京: 南京航空航天大学, 2014.

[5] 王清龙. 光伏发电系统发电功率预测研究[D]. 天津: 天津大学, 2013.

[6] 李辉. 光伏发电系统监控与发电预测模型研究[D]. 上海: 东华大学, 2013.

[7] 徐一卜. 分布式光伏发电应用与推广初探[J]. 科技创业家, 2013(14): 121.

[8] 沈才良. 光伏电站故障监测系统的设计与实现[D]. 哈尔滨: 哈尔滨工业大学, 2015.

[9] 韩晶. 大数据服务若干关键技术研究[D]. 北京: 北京邮电大学, 2013.

[10] 卓安. 基于 P2P 可伸缩架构的大数据分析平台研究与实现[D]. 北京: 清华大学, 2012.

(上接第 51 页)

[5] 刘源, 蒋建锋. 基于 IPSec 协议的多 VPN 互通技术在工程实践中的应用研究[J]. 电子世界, 2016(14): 154–155.

LIU Y, JIANG J F. An applicable Studies in Project Practice Base on Multi IPSec VPN Connection[J]. Electronics World, 2016(14): 154–155.

[6] 中华人民共和国工业和信息化部. 中华人民共和国网络安全法

[EB/OL]. (2016–11–08) [2017–06–12]. http://www.npc.gov.cn/npc/xinwen/2016–11/07/content_2001605.htm.

[7] FAN W, LI J, TANG N, et al. Incremental Detection of Inconsistencies in Distributed Data[J]. IEEE Transactions on Knowledge & Data Engineering, 2014, 26(6): 1367–1383.

[8] 全国信息技术标准化技术委员会. 大数据安全标准化白皮书(2017)[R]. 北京: 全国信息技术标准化技术委员会大数据安全标准特别工作组, 2017.