

# 一种大型数据中心数据校核分析系统

岳 备

(上海宝信数据中心有限公司 上海 200949)

**摘要** 数据中心随着运行时间的增长以及机房环境的日益复杂,数据采集会发生误差。依靠传统的人力并不能对数据完成校验。针对此问题设计一种大型数据中心数据校核分析系统。通过寻找系统环境的温度、湿度以及大型数据中心数据之间的关系,并对相应的联系进行对比、汇总,发现设备运行的规律和趋势。最终为数据中心维护人员提供决策依据。整个系统运行至今已减少了大量的人力,提高了校核准确度。

**关键词** 数据中心 数据校核 自动分析

中图分类号 TP3 文献标识码 A DOI:10.3969/j.issn.1000-386x.2018.12.008

## DATA CHECK ANALYSIS SYSTEM OF LARGE-SCALE DATA CENTER

Yue Bei

(Shanghai BaoXin Database Co., Ltd., Shanghai 200949, China)

**Abstract** With the increase of the running time and the increasing complexity of computer room environment, data acquisition in the data center occurs to error. Relying on traditional manpower cannot check the data. To solve this problem, we designed the data check analysis system of large data center. By searching and acquiring the temperature and humidity of system environment, and the relationship between data in the large-scale data center, system compared and summarized the corresponding relations, and found the rule and trend of the equipment operation. It provided the decision-making basis for the maintenance personnel in the data center. The operation of the whole system has reduced a large number of manpower, and improved the accuracy of verification.

**Keywords** Data center Data check Automatic analysis

## 0 引言

伴随着大数据技术与云计算计算的飞快发展,许多公司都对服务器运维十分重视。一些大型的互联网计算公司的服务器机房数据中心占地面积越来越大,而且还存在多地多数据中心联动工作<sup>[1]</sup>,数据中心的平稳运行关系着公司业务的命脉。在数据中心维护中,公司对数据中心也越来越重视,数据中心的供配电、暖通空调、消防弱电等系统日益变得复杂,大量传感器、智能仪表应用于各系统。通过采集并监控现场各类传感器和智能仪表数据,并加以分析和处理,能够有效降低电力成本,减少事故发生率,提高设备运行效率,保障设备机组安全,提高数据中心的经济效益<sup>[2]</sup>。

由于大型数据中心采集海量数据,机房维护人员往往难以通过传统手工模式及时完成数据校核分析,从而导致故障信息无法提前获取,延误处理时机,给机房管理工作带来运营风险。另外,由于一些测量设备随着运行时间增加而发生测量偏差,也会误导维护人员对故障和业务容量的准确判断,从而影响机房运维质量和成本等。针对这些低概率测量误差设备,采用大面积统一校核不太合适,还会增加运营成本和维护风险<sup>[3]</sup>。机房维护部门通常把数据校核分析业务外包给各类供应商。由于各类供应商主要针对各自设备单独校核,工具方法标准不统一,相互间提供校核数据和分析结果就会出现偏差,达不到维护效果,也给机房运营带来巨额成本。

## 1 研究目的

本文所述校核不是简单核对的意思,而是指针对寻找和发现大型数据中心所产生数据之间的关联依附关系,进行比对、校正、汇总、处理和相关性分析,进而发现系统或设备运行规律和趋势。最终为机房管理者和运行维护人员提供决策依据等系列工作。

大型数据中心现场通过传感器采集温度、压力、真空、热电偶、热电阻、电压、电流、电阻、频率等运行数据。运行设备根据这些初步采集数据进行一些加工成中间数据<sup>[4]</sup>,如功率因数、功率、有功功率、无功功率、电能、流量、冷量等。环控系统将这些不同设备和传感器初步数据和中间数据集中汇总。关于这些不同系统、设备和传感器,本文根据已研究的成果,主要从 6 个维度进行数据的关联,分别为上下级关联校验、同级左右校验和历史前后校验。(1) 上下级校核:通过校核上下级仪表之间数据,进而发现数据是否异常和运行趋势。(2) 同级左右校核:通过校核同级仪表之间数据,进而发现数据是否异常和运行趋势<sup>[5]</sup>。(3) 时间前后校核:通过校核仪表历史前后之间数据,进而发现数据是否异常和运行趋势。

随着自动化和信息化技术的不断发展,目前大型数据中心数据将逐渐由手工校核工作向自动化校核的方向发展。本文根据上文数据关联性研究成果,通过信息化系统自动完成大型数据中心数据统计、汇总、处理和相关性分析。研究如何呈现哪些主要关联性数据及其处理,从而帮助机房维护人员提高工作效率,降低运维风险。

## 2 系统模块

### 2.1 数据校核分析系统

数据中心服务终极目标为用最低成本提供最好服务。根据行业实际调查结果,大部分机房运营目标主要有两点:确保客户业务持续性,即业务和设备运行不得中断;运营成本有效降低,即降低维护成本和实现机房设备运行节能<sup>[6]</sup>。为此,数据校核分析系统作为数据中心信息化管理系统的子系统,其建设主要围绕以上两个目标开展,即通过实施数据校核分析系统可有效帮助机房维护人员确保客户业务连续性,并降低运营成本,数据校核分析系统如图 1 所示。

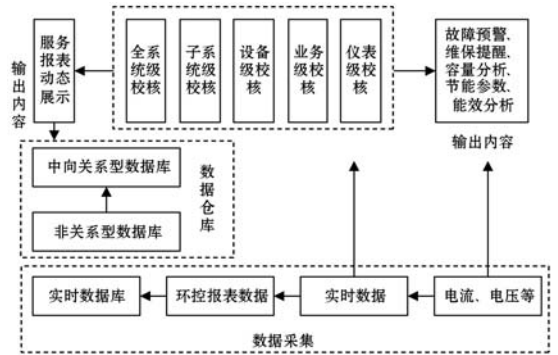


图 1 数据校核分析系统

从图 1 可以看出,数据校核系统可通过不同层面实现以下目标:通过机房全系统数据校核,可发现全系统运行健康程度、系统隐患和能效,进而帮助机房管理者从整体上把控整个机房运行状况;通过子系统数据校核,可发现子系统健康程度、系统隐患和能效,进而帮助专业工程师从子系统层而把控其运行状况,调整和改善子系统运行策略,从而实现机房节能运行;通过设备数据校核,可帮助一线运维人员发现设备健康程度、设备运行隐患和能效,提醒运维人员提前做好设备维护保养、检修准备<sup>[7]</sup>;通过业务数据校核,可帮助客户调整和优化实际业务策略,并提高客户满意度;通过仪表数据校核,及时发现个别仪表误差,避免大面积计划性维护带来成本。根据上述不同层面数据校核,设计了 5 个系统子模块,并与上述目标一一对应。

### 2.2 全系统级校核模块

全系统级校核模块主要从全局视角关注机房全系统运行数据校核,如图 2 所示。涉及机房全系统校核数据主要有两个输入变量,分别为 IT 负载变量和机房环境变量,输出数据主要有 PUE、用电总量、总耗冷量以及之间比较机房 EOP。通过校核输入变量与输出变量变化,进而提取整系统运行规律和趋势,从而可以判断出存在哪些隐患,系统健康程度如何以及能效高低,进而帮助机房管理者做出系统级项目改造和调整。

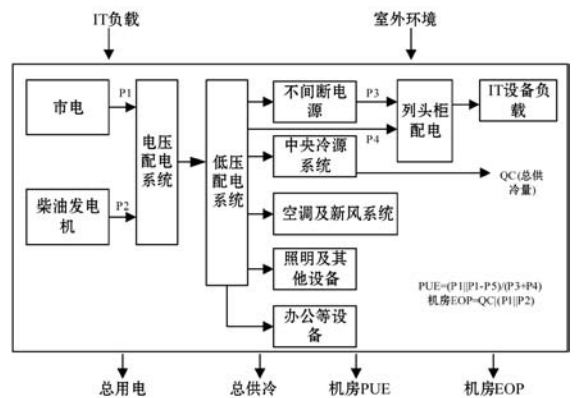


图 2 全系统级校核模块图

### 2.3 子系统校核模块

子系统级模块主要从局部关注子系统运行数据校

核,主要分为供配电子系统、暖通空调子系统、消防子系统和弱电子系统。本文只介绍暖通空调子系统,子系统校核模块图如图3所示。涉及暖通空调子系统校核数据主要有4个输入变量<sup>[8]</sup>,分别为IT负载数据、冷机出水温度数据、末端空调出风温度数据、室外环境干湿球温度数据。输出变量数据主要有:暖通设备总功耗及组成功耗、暖通设备供冷量、冷机EOP等。通过校核输入变量与输出变量变化趋势和规律,帮助专业工程师及时发现子系统存在运行隐患、健康和能效状况,提前做应对防范措施,调整和优化运行设备及其参数,从而避免故障发生,提高系统能效。

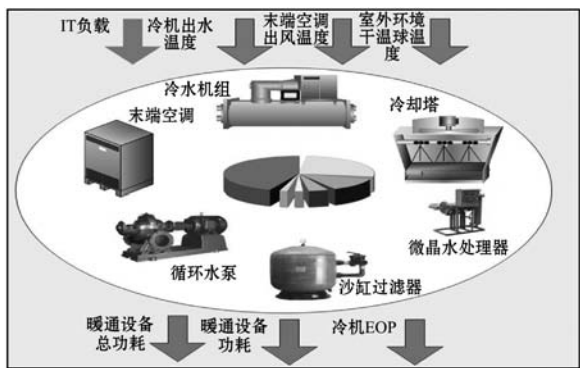


图3 子系统校核模块图

### 2.4 设备校核模块

设备级校核模块主要关注具体设备运行数据,并通过6个维度校核设备之间数据。设备级校核模块图如图4所示。主要数据有:电能、电压、电流、有功功率、功率因数等。通过上下仪表、同级仪表和历史数据校核<sup>[9]</sup>,发现仪表数据是否误差,运行数据趋势突变规律或异常,进而帮助专业工程师或一线运维人员提前发现设备运行隐患和故障情况,做好提前检修和维保准备。可通过不同颜色(本文用灰度表示)显示其偏差,如校核数据按偏差范围设为三个级别,显示三种颜色,分别为红色、黄色、绿色。目前暂定偏差 $\leq \pm 10\%$ 区间显示绿色, $\pm 10\% \leq$ 偏差 $\leq \pm 20\%$ 显示黄色,偏差 $\geq \pm 20\%$ 显示红色。

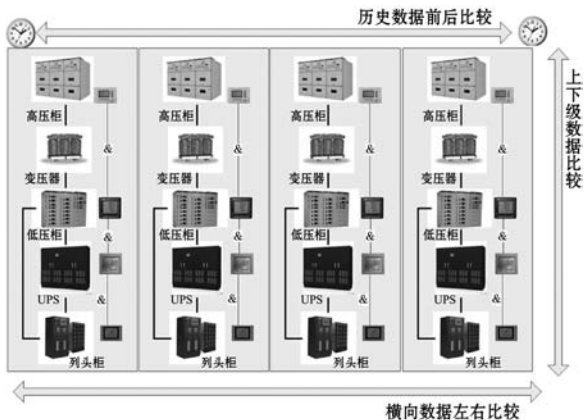


图4 设备级校核模块图

### 2.5 业务级校核模块

业务级校核模块主要关注客户机柜及其列头柜运行数据校核。主要有整个IT机房运行功率、整个冷通道运行功率、机柜电流、电压和功率、精密空调供冷量、机房温湿度、精密空调功耗等校核数据<sup>[10]</sup>。业务级校核模块图如图5所示。通过校核这些数据,并提供相应机柜负载报表,可以帮助客户及时发现IT设备异常状况,精密空调运行健康状况等,进而可以帮助客户合理调整IT设备负载和空调供冷模式。



图5 业务级校核模块图

### 2.6 仪表级校核模块

仪表级校核模块主要关注传感器数据与环控数据之间校核,仪表级校核模块如图6所示。仪表级校核模块主要校核数据有电压、电流、电阻、频率、压力、温度、湿度、流量计等。现场运维人员通过一些校验工具,将测量的数据录入手持终端系统后,系统自动判断数据偏差,进而帮助一线运维人员工作量,节省运维成本。原先校核方法是:两名维护人员,一名在现场测量,另一名在监控大厅通过对讲机联系校核环控数据与传感器,效率较低。

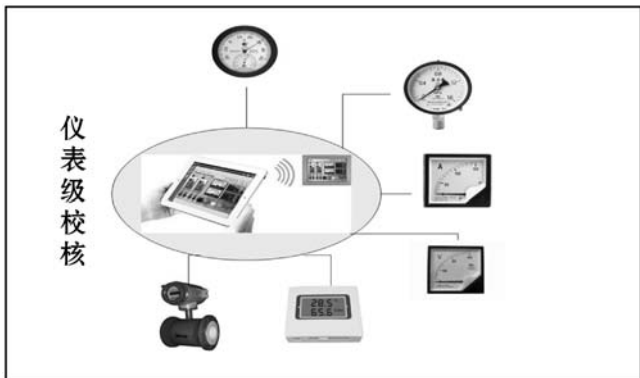


图6 仪表级校核模块图

## 3 实现意义

### 3.1 自动化取代部分人工

大型数据中心数据校核由于工作量较大,校核难

度高,往往外包给维护商。供应商采用方法往往采用手工校核,采集数据时间一致性存在一定误差。供应商技术人员经常出入机房且进行大量接触性操作,给机房带一些不确定因素,也给管理增加一些难度和成本。采用自动化校核不仅能节省维护外包费用,还能避免一些不必要事件发生。另外,自动化校核通过同时关联实时数据,引入许多数学分析模型,比逐点校核效率高,且能发现一些系统层面上的问题,系统自动化校核取代人工校核示意图如图 7 所示。

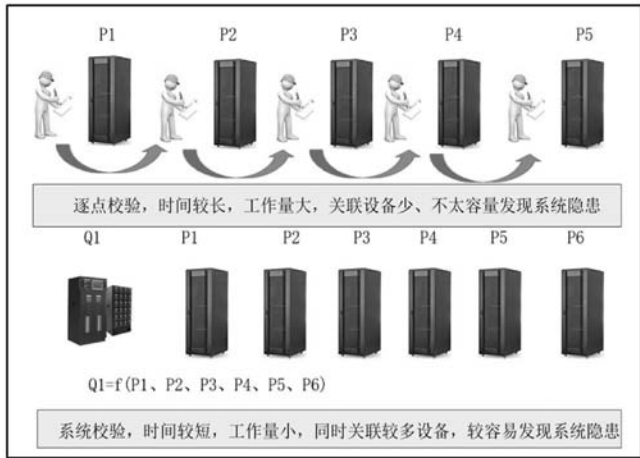


图 7 系统自动化校核取代人工校核示意图

### 3.2 智能预测和诊断

关于故障诊断与故障预测方法的分类,现在许多方法的标准都是不统一的,基于经验、趋势、模型和数据等分类方法也是片面化的方法。本文系统从故障诊断与故障预测的运行趋势数据出发,依据采集数据校核分析,对众多的故障诊断与故障预测方法进行归类分析,形成较为完善故障诊断与故障预测技术。设备健康管理把设备分为三类状态:健康——亚健康——故障。设备使用寿命是一个由健康——亚健康——故障——报废,即设备形态与性能由量变到质变的动态过程。设备现行管理和维修的理论、模式、制度是一种被动式滞后性管理。它以设备的故障管理与维修为核心,重点关注设备的故障阶段,以被动保养、排故诊断、解体换件维修为基本模式,缺乏对设备在“亚健康”阶段的形态与性能的动态劣化和系统平衡紊乱的控制对策。其结果势必造成无可挽回的能源、备件、人力、时间、生产和产品的损失。因此,设备健康管理是监控的重中之重。本文系统根据设备数据校核分析,诊断系统及设备的健康状态,为设备健康管理提供决策信息。图 8 为冷水机组数据校核分析表,其中深色区域数据已超出其基准范围,代表设备运行不健康,而浅色区域在基准范围之内,代表健康。

类型	校核对象	校核基准	冷水机组							
			1#	2#	3#	4#	5#	6#	7#	8#
冷冻泵	进水压力	1~3kg	2.5	2.5	2.5	0	2	2	3	0
	出水压力	2~5kg	3	3	3	0	4	4	5	0
冷冻水	进水温度	2~5℃	3	3	3	0	3	3	3	0
	出水温度	7~12℃	7.5	7.5	7.5	0	7.2	7.2	7.2	0
	出水压力	2~5kg	3	3	3	0	4	4	4	0
冷却水	进水压力	1~4kg	3	3	3	0	2	3	3	0
	出水压力	2~5kg	3	3	3	0	2	3	3	0
	进水温度	20℃ T-25℃	22	22	22	0	22	22	22	0
压缩机	高压 (134g)	8.5kg i / 5.5kg i	7	7	7	0	6	6	6	0
	低压	3.5kg i / 2.2kg i	3	3	3	0	2.5	2.5	2.5	0

图 8 冷水机故障校核表

### 3.3 节能增效

根据校核数据得到的结果,可以给数据中心节能降耗提供直观科学的依据,为机房查找能耗弱点,促进机房管理水平的进一步提高及运营成本的进一步降低,冷冻机设备动态分析图如图 9 所示。通过数据分析,一方面可以使能源使用合理,控制能源浪费,达到节能减排的目的,节能降耗,创造能源效益;另一方面,可以帮助机房对整个系统全面的了解,对各个机房以及主要耗能设备进行实时考核比较,防止能源浪费,并可以帮助机房进一步优化用电措施,以降低单位能耗成本。图 9 中关联了冷却水出水温度和冷凝器饱和温度,当冷凝器饱和温度与冷却水出水温度数据贴合度较近时,表明整个冷机系统换热效率较高,冷机能效较高。当冷凝器饱和温度偏离冷却水出水温度较大时,表明冷机系统换热效率差,能效较低,需要清洗冷凝器。

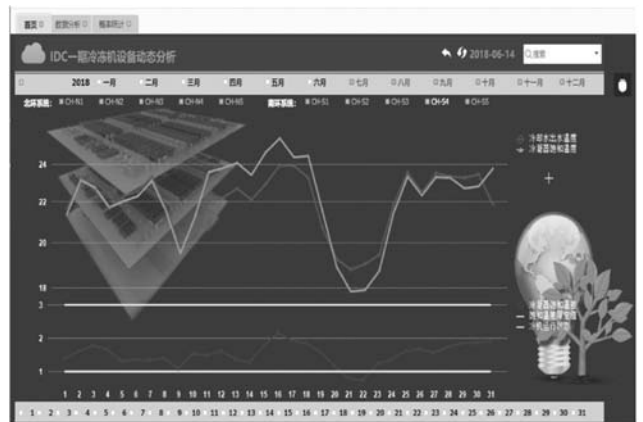


图 9 冷冻机设备动态分析图

## 4 结 语

通过校核分析系统研究并实现,可及时发现系统和设备隐患,提前做好检修和维保工作,从而减少甚至 (下转第 53 页)

### 3 结 语

本文求解了无标度 Sierpiński 网络的匹配数,利用无标度 Sierpiński 网络结构上的规律,建立了匹配数的递推关系,最后得到了匹配数的解析表达式,避免一般求解匹配数方法的高时间和空间复杂度。随着无标度 Sierpiński 网络每一次的迭代构建,网络的结点总数也在奇偶交替变换,且当网络结点总数为偶数时,无标度 Sierpiński 网络存在完美匹配,边覆盖数等于匹配数;当网络结点总数为奇数时,无标度 Sierpiński 网络存在近完美匹配,边覆盖数等于匹配数加 1。本文计算出匹配数最大时可能出现的子图的组合情况,然后结合无标度 Sierpiński 网络的结构特点,建立每种情况最大匹配数的递推关系,最后得到了无标度 Sierpiński 网络最大匹配数的递推关系式。本文采用的方法,不仅对无标度 Sierpiński 网络适用,对其他有自相似性质网络上的计数问题也有重要的借鉴意义。

### 参 考 文 献

- [ 1 ] Gast M, Hauptmann M, Karpinski M. Inapproximability of dominating set on power law graphs[J]. *Theoretical Computer Science*, 2015, 562:436–452.
- [ 2 ] Li H, Zhang Z. Maximum matchings in scale-free networks with identical degree distribution[J]. *Theoretical Computer Science*, 2017, 675:64–81.
- [ 3 ] Vukičević D. Applications of perfect matchings in chemistry [M]//*Structural Analysis of Complex Networks*. Birkhäuser Boston, 2011:463–482.
- [ 4 ] Liu Y Y, Slotine J J, Barabási A. Controllability of complex networks[J]. *Nature*, 2011, 473(7346):167–173.
- [ 5 ] Nepusz T, Vicsek T. Controlling edge dynamics in complex networks[J]. *Nature Physics*, 2011, 8(8):568–573.
- [ 6 ] Liu Y Y, Barabási A L. Control principles of complex systems[J]. *Reviews of Modern Physics*, 2016, 88(3):035006.
- [ 7 ] Valiant L G. The complexity of enumeration and reliability problems[J]. *Siam Journal on Computing*, 1979, 8(3):410–421.
- [ 8 ] Meghanathan N. Maximal assortative matching and maximal dissortative matching for complex network graphs[J]. *Journal of King Saud University—Computer and Information Sciences*, 2016, 28(2):230–246.
- [ 9 ] Zhang Z, Wu B. Pfaffian orientations and perfect matchings of scale-free networks[J]. *Theoretical Computer Science*, 2015, 570:55–69.
- [ 10 ] Hinz A M, Klavžar S, Zemljič S S. A survey and classification of Sierpiński-type graphs[J]. *Discrete Applied Mathematics*, 2017, 217:565–600.

- [ 11 ] Rosvall M, Trusina A, Minnhagen P, et al. Networks and Cities: An Information Perspective [J]. *Physical Review Letters*, 2005, 94(2):028701.
- [ 12 ] Haslinger C, Stadler P F. RNA structures with pseudo-knots: graph-theoretical, combinatorial, and statistical properties [J]. *Bull Math Biol*, 1999, 61(3):437–467.
- [ 13 ] Kabakçioğlu A, Stella A L. Scale-free network hidden in a collapsing polymer[J]. *Physical Review E Statistical Nonlinear & Soft Matter Physics*, 2005, 72(2):055102.
- [ 14 ] Jacobsen J L, Read N, Saleur H. Dense loops, supersymmetry, and Goldstone phases in two dimensions[J]. *Physical Review Letters*, 2003, 90(9):090601.
- [ 15 ] Zhang Z, Zhou S, Zou T, et al. Incompatibility networks as models of scale-free small-world graphs[J]. *European Physical Journal B*, 2007, 60(2):259–264.

(上接第 41 页)

避免事件发生,最终提高机房服务质量和效率。另外,该系统通过数据校核,可及时发现影响系统能效指标因素,通过技改或调整系统和设备运行模式,从而达到节能降耗目标。

### 参 考 文 献

- [ 1 ] 邓彬,江伟,郝蛟,等. SCADA 系统基础模型数据质量校核专家系统[J]. *电子技术应用*, 2016(z1):33–35.
- [ 2 ] 张丹,余维. 基于 Petri 网的电力系统监控数据多点校核方法[J]. *计算机测量与控制*, 2016, 24(4):101–104.
- [ 3 ] 周桂强. 调度集中数据校核系统研究与实现[J]. *铁道通信信号*, 2017, 53(6):8–11.
- [ 4 ] 陶佳燕,李银红,石东源,等. EMS 与继电保护定值校核系统实时数据匹配新方法[J]. *电力系统自动化*, 2012, 36(10):79–85.
- [ 5 ] 吕瑞,马春英. 列车运行控制数据校核及管理系统[J]. *铁道通信信号*, 2017, 53(2):17–19.
- [ 6 ] 刘伟铭,庄岩浩. 一种激光点云数据与视频图像相融合的收费数据校核系统[J]. *公路交通科技*, 2013, 30(8):123–128, 158.
- [ 7 ] 周栓林,吕豪英. 线路基础数据闭环校核管理系统的研究[J]. *中国铁路*, 2010(9):44–48.
- [ 8 ] 方东根,沈雷. 基于 iBeacon 技术的安全性服装设计 [C]//2014 年第四届“申洲杯”全国针织科技大会, 2014: 285–289.
- [ 9 ] 李文跃,苗龙,朱浩月. 工程机械冷却系热状态校核方法研究[J]. *建筑机械*, 2018(2):78–82.
- [ 10 ] 乔冠伦,李梧桐,李福兴,等. 同步发电机励磁系统参数辨识与校核工具开发[J]. *机电工程*, 2017(11):1348–1353.