

基于电厂运行数据库的专业智能搜索引擎

赵海波¹, 黄勇理¹, 陈 胜¹, 向隆刚², 郑楚光¹

(1. 华中科技大学 煤燃烧国家重点实验室, 湖北 武汉 430074 ;

2. 华中科技大学 达梦数据库多媒体技术研究所, 湖北 武汉 430074)

摘要: 电厂运行数据库中包含着大量的专业数据, 从这些表面上孤立的专业数据中得出某些参数之间的相互关系是比较困难的。介绍了所设计的专业智能搜索引擎的基本架构, 它是利用数据库中的数据挖掘和知识发现技术, 结合专业算法开发的。可以从电厂运行数据库中挖掘出指定运行参数之间的一些特定关系, 并可对挖掘的专业规则进行管理, 最后, 针对具体的电厂运行数据库, 给出了专业智能搜索引擎的应用示例。

关键词: 电厂运行数据库; 数据挖掘; 知识发现; 搜索引擎

中图分类号: TM 6; TP 311.13 文献标识码: A 文章编号: 1006-6047(2003)08-0025-04

0 引言

电厂运行数据库的信息非常丰富, 这些参数之间表面上孤立, 实际上某些参数之间隐藏着强烈的相关关系, 找出这些相关关系对于研究分析电厂运行状况有着重要作用。

在电厂实时运行数据库^[1,2]的基础上, 利用数据挖掘和知识发现技术, 开发了专业智能搜索引擎, 专门用于找出电厂运行参数之间的相关关系, 经过广泛性验证之后, 作为专业规则, 存入专业规则库中, 这些专业规则能可信地应用于故障诊断专家库, 决策支持和多维分析等系统。

1 专业智能搜索引擎系统构架

基于电厂运行数据库的专业智能搜索引擎具有典型的浏览器/服务器/数据库 3 层结构模式。三部分(表示、应用逻辑、数据访问)分割明确, 逻辑上各自独立, 可单独实现。客户端的功能简洁清晰, 大部分的应用逻辑部分被移植到专业应用服务器上, 但简单的专业应用逻辑处理和数据库访问仍然可在客户上实现, 以获得较高的效率。专业应用服务器(包括 Web 服务器)是应用逻辑处理的核心, 它是专业具体的实现, 智能搜索引擎和最关键的专业算法安装于其上。客户端将请求信息发送给专业应用服务器, 应用服务器返回数据和结果, 向数据库服务器发送 SQL 请求, 数据库服务器^[1-3]将数据访问结果返回给应用服务器。

专业智能搜索引擎根据电厂运行数据库中的历

史数据库, 利用专业算法, 进行专业规则挖掘。用户首先观察某个参数的变化趋势, 发现了感兴趣的变化后, 尝试从一些相关的参数中发现具有某种相关模式(同相、反相、提前、滞后等)的参数, 并将这种相关性以图形的方式显示出来。若观察的参数可能形成一个专业规则, 则将这种相关性参数保存, 作为专业规则挖掘的一个输入源。用户可以验证两个参数之间的相关性, 即设定相关模式, 选择合适的数据集及专业算法进行验证。如果通过验证(需要采用多个数据集验证参数之间的相关性), 则将这两个参数之间的相关性作为专业规则存入规则库中。

1.1 专业智能搜索引擎基本构架

专业智能搜索引擎基本构架如图 1 所示。

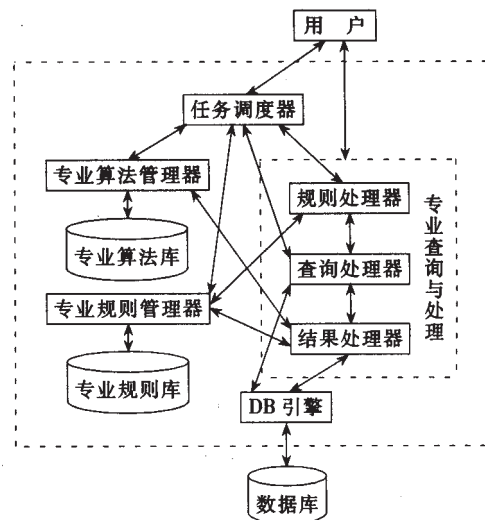


图 1 专业智能搜索引擎体系结构
Fig.1 The framework of professional intelligent search engine

1.2 系统流程图

系统流程如图 2 所示。

收稿日期 2002-11-28; 修回日期 2003-05-15
基金项目: “现代远程教育工程” 2001 年中央财政专项项目资助(教技司[2001]215)

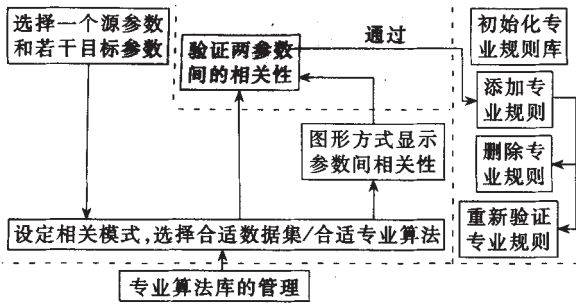


图 2 智能搜索引擎模块系统流程图
Fig.2 The system flowchart of professional intelligent search engine

2 专业智能算法设计 [1]

对于复杂系统, 变量之间的相互关系有些通过设备的运行原理分析是已知的, 但有些仍然是模糊不清或尚未完全了解的, 甚至可能是未知的。其中, 同相与反相、超前与滞后是经常作为评价若干个变量之间相互作用或相互影响关系的典型模式中的几个。

2.1 模型

对于 x, y 两个变量, 假设为平稳随机过程, 相关函数为

$$R_{xy}(\tau) = E(x(t)y(t+\tau)) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyf(x, y, \tau) dx dy$$

对于有限数据元组, 离散化后相关系数计算式为

$$R_{xy}(\tau) = \left(\frac{N \sum_{k=0}^N x_k y_{k+\tau} - \sum_{k=0}^N x_k \sum_{k=0}^N y_{k+\tau}}{\left[\left(N \sum_{k=0}^N x_k^2 - \left(\sum_{k=0}^N x_k \right)^2 \right) \left(N \sum_{k=0}^N y_{k+\tau}^2 - \left(\sum_{k=0}^N y_{k+\tau} \right)^2 \right) \right]^{1/2}} \right)$$

式中 N 为数据窗口宽度 W_i ; $\tau=0 \pm 1 \pm 2 \dots \pm W_i$ 。

逐渐前后平移 y (或 x) 的数据窗口 (即 $\tau=mT$, $m=0 \pm 1 \pm 2 \dots \pm W_i$), 可获得一系列相关系数, 如果两个变量具有相关特性, 相关系数可获得一个较为明显的极大值变化过程, 从相关系数为极大值时间点可判别两个变量的超前或滞后; 从相关点附近的数据变化可判别两个信号的特性是同相相关还是反相相关。

为了保证计算分析的准确性, 提高判别的可信度, 对于用户选区的数据搜索时间区间, 要充分利用搜索集内数据, 还要适当考虑信号的频谱特性, 相关计算的窗口尺寸应与信号的频谱范围相适应。初步设定如下:

$$W_i = n / [2 + \text{mod}(n/1200) - \text{factor}]$$

式中 W_i 为相关计算窗口宽度, 以时序点数计算, n 为集内数据点数, $\text{mod}(\cdot)$ 为取模运算, factor 为调整因子, $\text{factor}=0, 1, 2, 3$ 。

2.2 算法描述

为了提高计算的精度, 防止计算溢出, 要求程序

计算变量均采用双精度浮点数据类型。同相与反相、超前与滞后的相关规则验证算法简述如下:

- a. 依用户设置的时间区段, 计算数据时间跨度;
- b. 标准化主导信号和相关信号在计算时间区段内的数据;
- c. 计算确定相关分析计算窗口宽度和整个数据时间跨度内所需进行的相关验证的次数;
- d. 确定主导信号计算数据单元起点, 取 $\tau=-W_i \sim W_i$ 进行相关系数的计算;
- e. 寻找极值点, 若存在极值点, 确定其所在的时间坐标, 并确认其相关性;
- f. 若不存在极值点, 逐步放宽计算窗口宽度, 重新反复计算若干次, 找到极值点, 返回 e, 若找不到则认为此轮计算未获得相关性验证;
- g. 按照相关计算的时间坐标跨度遍历整个选定的数据区段, 每轮计算相关系数的极值超过设定的阈值, 则认为此段相关, 超过次数大于验证次数的半数以上才认为验证变量相关。

2.3 算法流程设计

程序编码采用动态链接库的形式实现, 算法流程图见图 3。

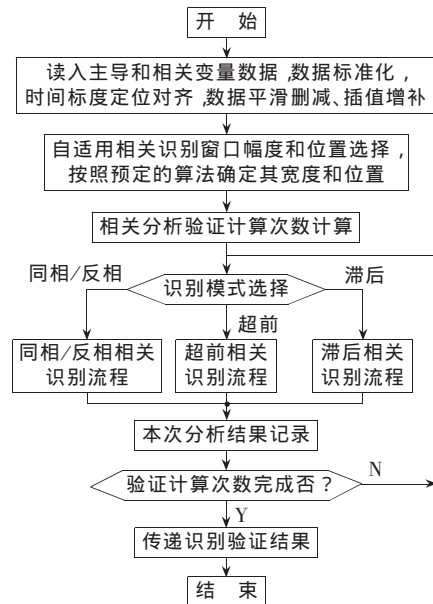


图 3 专业智能搜索引擎相关判别算法流程图
Fig.3 The flowchart of correlation discovery arithmetic in professional intelligent search engine

3 专业搜索示例

本套系统应用于湖北鄂州电厂 2 台 30 MW 机组。这里选择源参数为鄂州电厂 1 号机组主系统的汽包压力, 选择可能的相关参数为鄂州电厂 1 号机组主系统的主蒸汽压力、省煤器入口给水压力、烟气含氧量、炉侧主蒸汽温度、主蒸汽流量、炉膛压力, 时间窗口选定为 2002-01-10T00:00:00/2002-01-20T00:20:00。智能搜索引擎调用编制的专业算法

表 1 实际参数数值及其变化趋势分析表
Tab.1 The real parameter values and their variation tendencies

参数及变化	01-10	01-10	01-12	01-13	01-14	01-15	01-16	01-17	01-19
	T00 :02 :07	T20 :03 :15	T12 :03 :27	T08 :09 :35	T04 :05 :45	T20 :11 :42	T18 :13 :16	T12 :14 :19	T04 :18 :20
p_{qb}/MPa	0.01	0.01	18.13	14.41	13.61	17.55	18.46	17.26	13.31
p_{qb} 变化趋势		平稳	显著上升	下降	下降	上升	上升	下降	下降
p_{zzq}/MPa	0.03	0.04	17.15	13.86	12.94	16.33	16.93	15.94	12.75
p_{zzq} 变化趋势		基本平稳	显著上升	下降	下降	上升	上升	下降	下降
$p_{smqrkgs}/\text{MPa}$	-0.11	-0.10	18.30	14.50	13.82	17.70	18.59	17.39	13.48
$p_{smqrkgs}$ 变化趋势		平稳	显著上升	下降	下降	上升	上升	下降	下降
$y_{yqhy}/\%$	10.01	10.00	2.76	2.74	3.37	2.15	1.59	3.35	4.55
y_{yqhy} 变化趋势		基本平稳	显著下降	下降	上升	下降	下降	上升	上升
$t_{lczq}/\text{°C}$	39.23	34.30	535.23	531.26	535.89	534.98	540.29	531.15	540.89
t_{lczq} 变化趋势		基本平稳	显著上升	基本平稳	基本平稳	基本平稳	基本平稳	基本平稳	基本平稳
q_{zz} 主蒸汽流量/ $(\text{kg}\cdot\text{h}^{-1})$	2 780.62	2 780.62	884 320.00	540 512.00	582 144.00	911 056.00	997 888.00	907 888.00	557 808.00
q_{zz} 变化趋势		平稳	显著上升	下降	上升	显著上升	上升	下降	显著下降
p_{ll}/MPa	-0.47	-0.91	-1.28	-0.89	-0.59	-1.21	-0.75	-1.18	-0.77
p_{ll} 变化趋势		下降	下降	上升	上升	下降	上升	下降	上升

进行验证,通过专业算法验证的相关参数显示在搜索结果图上,并保存这种相关关系,以便作为专业规则添加到专业库中。搜索结果用不同颜色的曲线表示各个参数在选定时间窗口的变化趋势。智能搜索结果表明,汽包压力和主蒸汽压力、省煤器入口给水压力之间存在强同相相关,与烟气含氧量存在强反相相关关系。各个参数在历史数据库中的实际值及其变化趋势分析结果列于表 1,表中时间为 2002-01-10/2002-01-19。 p_{qb} 为汽包压力实际值, p_{zzq} 为主蒸汽压力实际值, $p_{smqrkgs}$ 为省煤器入口给水压力, y_{yqhy} 为烟气含氧量, t_{lczq} 为炉侧主蒸汽温度, q_{zz} 为主蒸汽流量, p_{ll} 为炉膛压力。从表 1 可以看出,在锅炉运行的过程中,即使在锅炉负荷显著加大(时间点为 2002-01-12T12:03:27 附近)和锅炉负荷波动的过程中,汽包压力的变化趋势与主蒸汽压力、省煤器入口给水压力的变化趋势都是一致的,存在着明显的同相相关关系;而汽包压力的变化趋势与烟气含氧量的变化趋势则是相反的,存在着明显的反相相关关系;汽包压力的变化趋势与炉侧主蒸汽温度、主蒸汽流量、炉膛压力的变化趋势在所考虑的时间段内并不一致或者相反,也不存在提前或者滞后的相关关系(虽然在某一小段时间窗口内可能存在这些相关关系),它们之间挖掘不出指定的相关关系。于是,把发现的此相关关系作为专业规则保存到专业规则库中。为了广泛性验证此专业规则,又选择了另外一个时间窗口,即 2002-01-20T00:00:00/2002-01-30T00:20:00 对此进行了验证,验证通过,认为此专业规则继续有效。

4 小结

本系统的开发为基于电厂运行数据库、利用先进的数据挖掘和知识发现技术的专业智能搜索引擎搭建了一个基础框架,能够投入实际运行,并能广泛

运用,所取得的结果具有指导电厂运行、为科研服务等作用。本系统仍然有许多工作需要进一步完成,首先,需要进一步丰富专业算法,包括各种异常情况下的相关算法,也可以引入一些专业模型作为专业算法;另外,可以试图实现后台无值守的专业智能搜索;最后,如何利用专业规则库服务电厂运行,如故障诊断、决策支持等也是一个重要课题。

参考文献:

- [1] 吴明光,张玉润,金建祥,等.工业控制实时数据库的研究[J].控制与决策,1995,10(3):238-243.
WU Ming-guang, ZHANG Yu-run, JIN Jian-xiang et al. Research of industrial control real time database[J]. *Control and Decision*, 1995, 10(3): 238-243.
- [2] 程正群,钱积新.一种分布式过程实时数据库的设计[J].化工自动化及仪表,1999,26(5):36-38.
CHENG Zheng-qun, QIAN Ji-xin. A design of distributing process real time database[J]. *Control and Instruments in Chemical Industry*, 1999, 26(5): 36-38.
- [3] 尹朝万,祝中华.远程数据库访问系统客户/服务器体系结构模型的设计与实现[J].软件学报,1995,6(8):455-462.
YIN Chao-wan, ZHU Zhong-hua. Design and implementation of remote database access client/server architecture model[J]. *Journal of Software*, 1995, 6(8): 455-462.
- [4] 江胡桢.现代工程数学手册(第三卷)[M].武汉:华中工学院出版社,1986.
JIANG Hu-zheng. Handbook of modern mathematics in science and engineering (Vol. III) [M]. Wuhan: Huazhong University of Science & Technology Press, 1986.

(责任编辑:汪仪珍)

作者简介:

赵海波(1977-),男,湖南宁乡人,助教,博士研究生,主要研究方向为电厂自动化与控制技术和多相湍流反应流体力学(E-mail: kelisiman@mail.hust.edu.cn);

黄勇理 (1963 -),男,湖南长沙人,副教授,博士,主要研究方向为连续过程和离散事件系统的检测、仿真和控制;

陈 胜 (1977 -),男,湖北武汉人,博士研究生,主要研究方向为电厂仿真及 LB 方法;

向隆刚 (1976 -),男,湖北武汉人,博士研究生,主要研究方向为数据挖掘和数据库中的知识发现;

郑楚光 (1945 -),男,湖北武汉人,博士研究生导师,研究方向为燃烧流体力学及电厂控制。

Professional intelligent search engine based on running database of power plant

ZHAO Hai-bo ,HUANG Yong-li ,CHEN Sheng ,XIANG Long-gang ZHENG Chu-guang
(Huazhong University of Science & Technology ,Wuhan 430074 ,China)

Abstract : The running database of a power plant includes a great deal of professional data. It is difficult to discover the correlation between some parameters from those data ,which appear independent. The framework of a PISE (Professional Intelligent Search Engine) is designed and introduced ,which is developed using data mining and knowledge discovery technologies with professional arithmetic. PISE can mine out the correlation between given parameters from running database of power plant and administer these professional rules discovered. An application example based on the actual running database of a power plant is given.

This project is supported by Special Case of Central Finance in 2001 of "Contemporary Long-distance Education Project " (Document [2001] 215 of Technology Department ,Ministry of Education)

Key words : running database of power plant ; data mining ; knowledge discovery ; intelligent search engine

国电南京自动化股份有限公司新产品鉴定信息 (一)

2003 年 4 月 20 日 ,南京市科技局受江苏省科技厅委托 ,主持召开了 RMS 601 继电保护与故障信息管理系统鉴定会 ,与会委员经认真讨论研究 ,一致同意通过鉴定 ,并认为 RMS601 继电保护与故障信息管理系统在系统功能、稳定性、可扩充性等方面采用了许多独特的设计方法 ,提高了系统的整体性能。技术条件符合现行国家标准 ,提交鉴定审查的技术文件完整正确 ,试验项目齐全 ,经现场实际运行 ,证明系统运行稳定可靠。鉴定委员会一致认为 :系统整体技术达到了国内领先水平。

2003 年 5 月 24 日 ,江苏省电力公司受国家电力公司科技环保部委托 ,主持召开了 PSI 200 变电站电气设备状态实时检测系统鉴定会 ,与会委员听取了产品研制报告、科技查新报告、测试报告及试运行报告 ,审查了相关技术文件 ,经认真讨论研究 ,一致同意通过鉴定。鉴定委员会一致认为 :PSI 200 变电站电气设备状态实时检测系统以“零磁通”微电流传感器技术、高精度数字采集与信号处理技术、故障诊断专家系统技术和嵌入式 CAN 网同步相应技术、多层组合式电磁屏蔽技术为核心 ,提供变电站电气设备状态检测的综合解决方案 ,整体性能达到国内领先水平。其中“零磁通”微电流传感器技术、数字信号处理算法达到国际先进水平。南京市科技局受江苏省科技厅委托 ,主持召开了 PSV 601 电压无功综合控制系统鉴定会 ,与会委员在听取了产品研制报告、科技查新报告、测试报告及试运行报告 ,审查了相关技术文件后 ,经讨论研究 ,一致同意通过鉴定。鉴定委员会一致认为 :PSV 601 电压无功综合控制系统在系统功能、控制原理、系统稳定性等方面采用了独特的设计方法 ,提高了系统的整体性能。系统整体技术达到国内先进水平 ,其中各时段的动作策略实现逻辑可编程的技术达到国内领先水平。

(国电南京自动化股份有限公司技术部)