

电网信息系统状态评价技术研究与应用

谢晓娜¹, 常政威², 陈亚军³

1. 成都信息工程学院控制工程学院, 四川 成都 610225;
2. 国网四川省电力公司电力科学研究院, 四川 成都 610072;
3. 华北电力大学, 河北 保定 071003)

摘要: 针对电网信息系统状态检修提出了一种电网信息系统状态评价及短期预测方法。通过充分探讨建立了信息系统状态评价指标体系,并根据各级指标重要程度赋予其相应权重。实现对B/S构架信息系统整体状态的监控和评价,并以信息系统历史状态作为训练样本预测其短期状态,成果将作为重要部分应用于电网信息系统状态检修。

关键词: 信息系统; 状态评价; 短期预测; 状态检修

Abstract: Focusing on condition-based maintenance of grid information system, the state evaluation and short-term prediction methods for grid information system are proposed. The state evaluation index system of information system is established through full discussion, and the corresponding weight is given according to the important degree of each index. The monitoring and evaluation for the overall state of information system with B/S architecture are realized, and its short-term state is predicted taking the history of information system as training samples. The results will be served as an important part of condition-based maintenance for grid information system.

Key words: information system; state evaluation; short-term prediction; condition-based maintenance

中图分类号: TM764.1 文献标志码: A 文章编号: 1003-6954(2015)04-0024-05

DOI:10.16527/j.cnki.cn51-1315/tm.2015.04.006

0 引言

“十二五”期间,国家电网公司全面建设坚强智能电网,推动公司和电网创新发展、集约发展、安全发展,信息化工作已成为国家电网公司实现战略发展目标的基础工程和重要保障^[1]。国家电网公司SG186信息系统、SG-ERP系统等业务信息系统^[2]的大量投入运用,在对公司业务工作支撑力不断提升和主营业务大幅度融合的同时,也使信息系统的规模急剧膨胀,而新系统、新技术的不断应用,使系统出现故障的可能性也随之加大,对系统稳定性和系统管理者的要求也随之提高。如何提高信息系统检修的工作效率、工作质量、安全、稳定等问题也随之突显。因此,采用创新的管理思想和现代计算机技术手段,保障信息系统的安全、降低生产成本、提高企业竞争力,成为电网企业的当务之急。正是在这种背景下,信息系统状态检修已然成为电网检修模式改变的一种必然趋势。

信息系统状态检修的基础是对于信息系统状态的监控与评价,国家电网公司企业信息系统现阶段主要采用B/S构架,B/S模式信息系统主要结构包括应用服务器和数据库服务器,因此对于应用服务器和数据库服务器内部状态和外部状态的监控与评价成为信息系统状态评价研究中最为重要的环节。

信息系统检修是信息运维人员工作的重要内容,信息系统的状态检修也是必然发展趋势^[3-4]。目前国内对于信息系统状态检修的研究还处于起步阶段,大部分研究者对于信息系统状态的监控和评价仅停留在整体的设计思路上面,并没有继续向下延伸到各具体部分状态的研究。也正是基于此,给出了一种对于B/S信息系统状态全面监控与评价的方法,并给出对于信息系统状态短期预测的方法,为信息系统状态检修提供更为具体的研究思路 and 实现技术参考。

1 评价系统总体设计

国家电网公司企业信息系统现阶段主要采

基金项目: 国网四川省电力公司研究开发项目(52199713504A)

用 B/S 构架,下面主要针对 WebLogic 服务器和 Oracle 数据库服务器搭配技术路线的大型信息系统开展研究。图 1 为信息系统状态评价系统的总体设计框架。

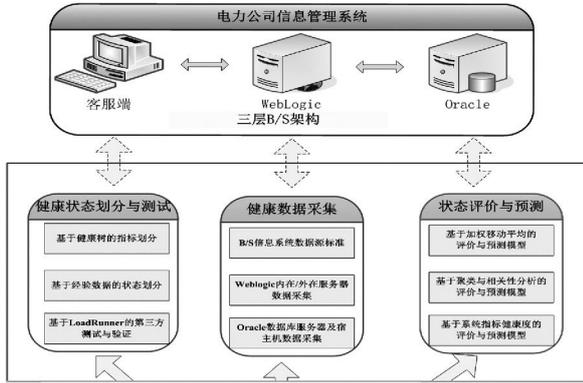


图 1 信息系统状态评价系统的总体设计框架

由图 1 可知,所述系统主要包括健康状态划分与测试、健康数据采集、状态评价与预测和健康状态存储系统这 4 个部分。健康状态划分与测试主要是建立所述评价系统的评价指标体系以及对于信息系统健康度的合理划分;健康数据采集模块是实现 WebLogic 服务器和 Oracle 数据库服务器内在及外在状态数据的监控,并将监控数据存储于健康状态存储数据库中,为状态评价模块提供实时数据来源;状态评价与预测模块实现对信息系统当前状态的评价,并结合信息系统历史状态数据给出对信息系统状态的短期预测;健康状态存储系统即建立的存储数据库,负责所述评价系统涉及的所有数据的存储。

2 健康状态划分与测试

2.1 健康状态划分

根据 B/S 架构系统的特性,将整个系统状态划分成两部分,即 WebLogic 服务器状态和 Oracle 服务器状态。这两个服务器又有各自不同的子状态,这些状态量根据 WebLogic 和 Oracle 数据库的官方文档、书籍介绍和运维人员的建议,对采集到的指标进行了状态划分,每一种状态有相应的阈值,这些阈值也是参考的文档、书籍以及运维人员的经验,得到每一个状态量都划分为 4 种状态:正常状态、注意状态、异常状态、危险状态,各状态对应给出分数如表 1 所示。

表 1 状态划分表

各级指标	正常状态	注意状态	异常状态	危险状态
得分	75 ~ 100	50 ~ 75	25 ~ 75	0 ~ 25

2.2 基于 LoadRunner 的测试实验

LoadRunner 是一种预测系统行为和性能的负载测试工具^[5]。测试分析是根据 LoadRunner 自动生成的各项指标测试结果统计图,如测试简述图、事务平均时间响应图、系统资源图等,从中了解和判断被测系统在不同压力测试下的运行状况以及系统资源的使用情况,对它们进行关联分析,从而确定系统指标不同状态的阈值。通过模拟上千万用户实施并发负载和实时性监测的方式来确认和查找问题,将测试得到的性能指标值与定义的指标状态阈值进行对比,从而验证阈值的正确性。同时,经过测试实验,定义了各级指标权重如表 2 所示。

3 状态数据采集

根据 B/S 系统的分层,将状态监视采集系统也分成相应的子模块,分别对每层进行监视。状态监视采集系统由 3 个子状态监视采集系统构成^[6-7]。分别是 Weblogic 服务器数据采集程序、Oracle 数据库数据采集程序、Oracle 宿主机数据采集程序。它们对相应的模块按照设定的频率进行采集,将采集到的数据持久化后,由主程序进行读取,然后按照状态定义分别对数据进行相应状态的划分,最后得出当前系统状态的评价。

3.1 Weblogic 服务器数据采集程序

Weblogic 服务器数据采集程序^[8]能够获得服务器内 JMS 信息、进程信息、线程信息、JVM 内存信息、CPU 使用率等一系列服务器数据。在获得信息状态数据之后,可以对数据进行分析,从而得出服务器现阶段状态。

逻辑流程如图 2 所示。程序启动后,首先检测 Weblogic 服务器是否启动,如果没有启动,则进行报错,提示服务器没有正常启动。如果 Weblogic 服务器正常启动后,则启动系统数据采集子程序以及 Weblogic 服务器数据采集子程序,相应的子程序按照相应的时间频率刷新重新采集相应指标,主程序按照相应的时间频率获得相应数据后,程序将数据写入到 MySQL 数据库中。

表2 指标权重定义

一级指标	二级指标	权重	三级指标	权重	四级指标	权重	
系统状态	Oracle 服务器	40	Oracle 数据库	25	数据库后台进程	7	
					数据库表空间	7	
					数据库命中率	5	
					数据库响应时间	6	
	Oracle 服务器外在状态	30				cpu	8
						硬盘	4
						交换区	7
						内存	6
						网络质量	5
						WebLogic server 线程	8
	WebLogic 内在状态	25				server 队列	6
						WebLogic 内存	5
						jvm 堆	6
						cpu	8
WebLogic 服务器外在状态	30				硬盘	4	
					内存	6	
					交换区	7	
					网络质量	5	

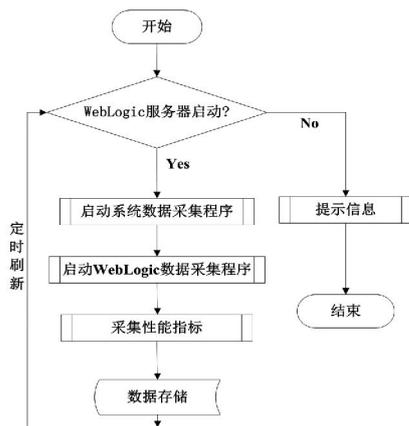


图2 Weblogic 服务器数据采集程序运行流程

3.2 Oracle 数据库数据采集程序

Oracle 数据库数据采集程序能够获得服务器内执行队列、等待执行队列、数据库命中率、SQL 读写数、数据库使用率等一系列数据库服务器状态。在获得信息状态数据之后,将数据写入到 MySQL 数据库中。

流程逻辑如图3所示。程序启动后,首先实例化3个子数据采集程序,并设置相应的采集频率。设置完毕后,则按照相应的频率各子程序采集数据,

并将采集得到的数据存储到数据库中。

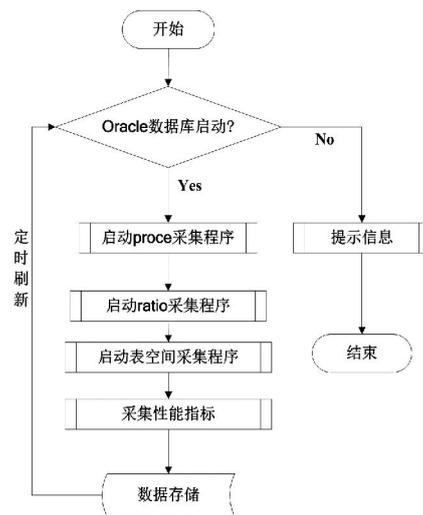


图3 Oracle 数据库监控程序运行流程

3.3 Oracle 宿主机数据采集程序

Oracle 宿主机数据采集程序能够获得客户端服务器的CPU 状态信息、交换区信息、内存信息、硬盘信息等一系列客户端服务器状态信息。在获得状态信息数据之后,可以将数据储存到 MySQL 数据库中。

流程逻辑如图4所示。程序启动后,首先实例化

4个子程序监控系统,并设置相应的采集频率。设置完毕后,则按照相应的频率各子程序采集数据,并将采集得到的数据反馈到服务器上。在获得相应数据后,程序对数据进行分析并进行相应的状态提醒。

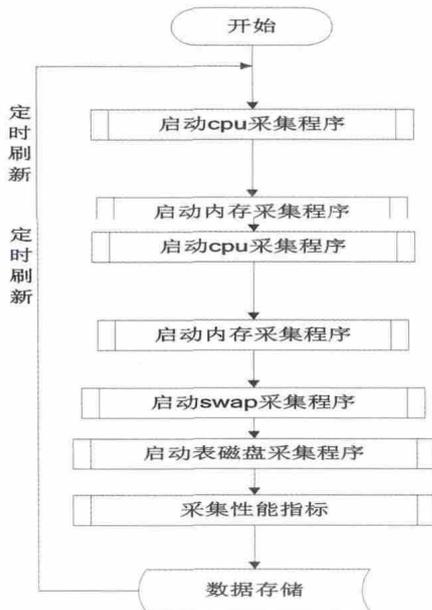


图4 Oracle 宿主机数据采集程序运行流程

4 状态评价与预测

4.1 基于聚类回归的状态评价模型

一般研究认为低负载的系统健康度高于高负载的系统,然而在实际系统运行过程中,一个系统正常态可能是适中的负载,也可能是高负载的。正是考虑到这样的情况存在,专门设计了一个基于系统历史状态聚类与回归的评价模型。

实际系统运行过程中,对于稳定运行的系统而言,可以认为多数情况下系统是正常的,因此利用聚类的方法,将训练集合或者采集到的历史数据划分成 K 类,取数量最多的前 N 类作为系统正常态的代表。选取每一个指标中心点,各指标偏离中心点一定范围,依靠线性扣分,超出范围,以指数扣分,分数扣完为止。依据指标得分健康数目,划分等级,每个指标得分不到其分配得分的 β (β 默认取 0.6),则认为指标异常,结合利用 M5Rules 回归模型,计算各个指标相互间的回归关系,以偏离回归方程的 4 级指标前 n 个,且偏离比例超过 0.3 (计算方式为 $(|实际值 - 回归值|) / 回归值$) 的指标作为异常指标。如果有 n 个指标,所有指标正常则为正常,如果超过 $0.1 * n$ 个指标出现异常为整个系统为注意状态,超过 $0.2 * n$ 个指标异常为异常状态, $0.3 * n$ 个指标出

现异常则为危险状态。

这种模型不仅可以避免对于本身就是高负载系统的状态评价误差,而且不仅是通过由低级指标直接线性组合成为高级指标判断系统状态,其对每一个底层指标聚类分析,指导判断每一个指标的健康度,充分考虑了影响信息系统状态的因素,从指标的健康度来度量系统整体的健康度,避免误报、漏报。

4.2 基于隐马尔科夫的短期状态预测模型

隐马尔可夫模型(hidden Markov model, HMM)是马尔可夫链的一种,它的状态不能直接观察到,但能通过观测向量序列观察到,每一个观测向量都是通过某些概率密度分布表现为各种状态,每一个观测向量是由一个具有相应概率密度分布的状态序列产生。所以,隐马尔可夫模型是一个双重随机过程,即具有一定状态数的隐马尔可夫链和显示随机函数集。在充分研究各经典预测模型并结合电网信息系统状态特性的基础上,选择引入隐马尔可夫模型建立了信息系统状态短期预测模型。

这种模型以大量系统状态历史数据作为训练样本,对模型进行充分训练。对于当评价模型正常运行一段时间后,按照时间顺序记录系统四级指标的得分,便可以通过隐马尔可夫模型预测将来的一次或多次得分,最终预测三级、二级和一级指标状态。为了保证预测的实时性和有效性,周期性更新隐马尔可夫模型。

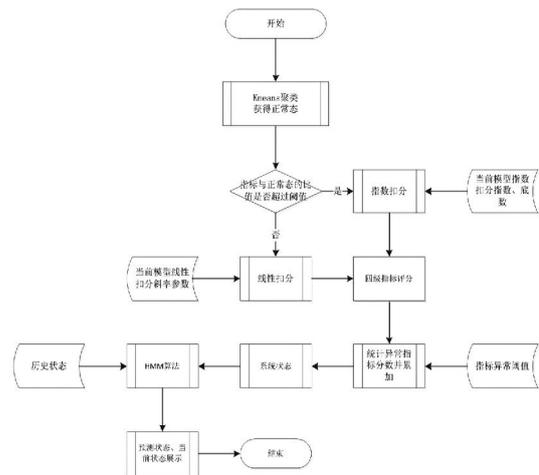


图5 评价预测整体模型流程图

5 实验验证

将所述状态评价系统应用于四川省电力公司 SG186安全监督与管理业务应用系统,它的应用服

表 3 测试场景

	基本用户数	执行时间	测试
场景 1	10	35 min	3 个用户,每 4 min 加载 1 个用户,运行 15 min,每 5 min 退出 1 个用户
场景 2	20	35 min	6 个用户,每 4 min 加载 2 个用户,运行 15 min,每 5 min 退出 2 个用户
场景 3	40	35 min	12 个用户,每 4 min 加载 4 个用户,运行 15 min,每 5 min 退出 4 个用户

表 4 数据采集功能测试结果示例

	Oracle 服务器				Weblogic 服务器			
	cpu 使用率 /%	内存使用率 /%	swap 利用率 /%	User Transaction Per Sec	cpu 使用率 /%	内存使用率 /%	swap 利用率 /%	Weblogic thread total
场景 1	<55	<70	62	1.8	<52	<68	51	29
场景 2	<80	<75	65	1.96	<74	<73	59	37
场景 3	<98	<78	67	2.1	<83	<77	62	60

表 5 评价预测功能测试结果示例

	1 min		2 min		17 min		18 min	
	当前状态	预测状态	当前状态	预测状态		当前状态	预测状态	当前状态	预测状态	
场景 1	正常	正常	正常	正常	正常	正常	正常	正常
场景 2	正常	正常	正常	正常	注意	正常	正常	正常
场景 3	正常	正常	正常	正常	注意	注意	注意	注意

务器和数据库服务器分别是 Weblogic 和 Oracle。为测试所述评价系统数据采集以及评价预测功能,设置如表 3 所示测试场景,可得到结果如表 4、表 5 所示。

从表 4 数据分析可以看出,Oracle 服务器采集程序能够采集此宿主机的硬件的动态信息。从场景 1 到场景 3,cpu、内存、swap 的使用率是递增的,这与测试用户数成倍增长相符,说明采集的数据是正确的,User Transaction Per Sec 的最大值从场景 1 的 1.8 到场景 2 的 1.96,再到场景 3 的 2.1 依次递增,与场景的用户数递增相符,说明采集程序的 Oracle 数据库采集模块能够正常工作。同样,Weblogic 服务器采集程序正常工作。从表 5 数据分析可以看出,评价系统充分评价信息系统当前状态,并能有效预测 1 min 后信息系统状态。在充分训练预测模型的前提下,预测时间可以更长。

6 总结与展望

提出了一种电网信息系统状态评价方法。通过充分探讨建立了信息系统状态评价指标体系,并根据各级指标重要程度赋予其相应权重。实现对 B/S 构架信息系统整体状态状态的监控和评价,并以信

息系统历史状态作为训练样本预测其短期状态,成果将作为重要部分应用于电网信息系统状态检修。

参考文献

- [1] 李向荣,郝悍勇,樊涛,等. 构筑数字化电网建设信息化企业[J]. 电网技术,2007(17):1-5.
- [2] 曾德君. SG186、ERP、IPSS 三者关系论述[J]. 华东电力,2009(9):1442-1444.
- [3] 刘贤杰,刘旭生. 信息通信状态检修系统中状态评价体系的探索与实践[J]. 电力信息化,2012(5):64-69.
- [4] 张涵. 开展信息系统状态检修提高系统运维水平[J]. 农村电气化,2014(6):32-34.
- [5] Xiaojiao Yan, Fuan Wen, Chunmei Fan, et al. Performance Testing of Open Laboratory Management System Based on LoadRunner[C]. Proceedings of the IEEE International Workshop on Intelligent Motion Control, 2011.
- [6] 陈伟. 基于网络的服务器状态监控系统的研究与实现[D]. 保定:华北电力大学,2014.
- [7] 吴岩. 电力信息系统智能巡检系统的设计与实现[D]. 成都:电子科技大学,2012.
- [8] 周卫华,马如军,谭成翔. WebLogic Server 安全技术概述及其在电子政务应用中的实现[J]. 计算机应用研究,2003,20(9):79-81.

(收稿日期:2015-05-14)