文章编号：2096-1472(2016)-12-17-03

基于数据挖掘的网管告警处理方法研究

许劭庆，马 彪，安海英
（国网吉林省电力有限公司，吉林 长春 130021）

摘 要：为了方便电力行业网管人员能够快速的从这些告警中找到有用信息，迅速的定位设备故障。本文设计了由接口告警采集、故障处理系统、告警相关性分析模块等构成处理框架。采用过滤告警、补全缺值数据、去重等方式进行数据预处理。通过告警发现和相关性分析机制实现告警的匹配和识别，利用基于关联规则的方法对告警信息挖掘。采用本文设计的方法，能够有效的实现网管系统的告警信息的挖掘分析，提高网管的效率。

关键词：数据挖掘；告警信息；关联规则
中图分类号：TP311 **文献标识码**：A

A Study on the Network Management Alarm Processing Method Based on Data Mining

XU Shaoqing,MA Biao,An Haiying
(*Jilin Electric Power Co.,Ltd.,Changchun* 130021,*China*)

Abstract: In order to facilitate network managers in the power industry to quickly find useful information from the alarm and rapidly locate the equipment failure,the paper designs an alarm information processing framework,which is composed of the interface alarm collection,the fault processing system,the alarm correlation analysis module,etc..Pre-processing is conducted through filtering alarms,complementing the missing value of data,and removing duplicate data. Alarms are recognized and matched through alarm detection and correlation analysis mechanism.It implements the alarm information mining based on the association principle.The method proposed in this paper can effectively implement the analysis and mining of the alarm information on the network management system and improve the efficiency of network management.

Keywords:data mining;alarm information;association principle

1 引言(Introduction)

随着信息技术的发展，越来越多的行业采用计算机技术来实现管理的自动化和智能化。电力企业也不例外，为了实现对供电线路和设备的管理，建立了自己的监控网络，进一步提高自己的运维保障能力。但随着监控网络的不断扩大，网络结构越来越复杂，很多的网络技术慢慢的走向了融合，这导致了网络发生故障的突然性越来越高，这些因素交融在一块，致使网络的维护、管理、操作变得越来越困难[1-3]。因此，目前的当务之急是找寻一种更加自动智能化、综合化的网络管理办法。当一个运维支撑网络产生问题或故障发生时，就会导致大量的告警产生，根据这些告警信息，就可以方便的分析设备的问题和故障，能够成功的解决网络运维管理中的各类问题。

本文利用这些告警信息，提出了基于数据挖掘的网管告警处理方法，该处理方法是在对比和研究了告警相关性分析理论以及技术之后才提出的。规则引擎可以利用其中已有的规则来动态的实时分析告警信息，规则引擎中保存的规则是在分析那些数据库中的告警信息并自学习之后才获取到的，告警转故障和根告警的分析就是通过对比这些规则来实现的，从而可以为系统故障智能化管理提供一系列的解决方案。不管是从理论研究还是实际工程研究应用方面都有着很强的意义和价值[4]。

2 告警处理总体设计(General design of alarming treatment)

告警信息处理涉及告警数据的采集、分析、过滤等方面工作。一个完整的电力行业网络监控的告警处理由网管子系统、接口告警采集、JMS消息通信、故障处理系统、告警相关性分析、规则专家管理模块和GUI界面等构成，具体的结构如图1所示。

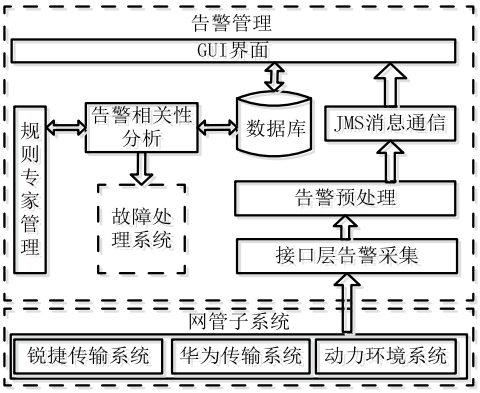

图1 告警信息处理系统结构

Fig.1 Structure of alarm information processing system

其中，接口层告警采集完成对网管子系统(如锐捷传输系统、华为传输系统、动力环境系统)的告警数据采集，接口层会根据不同的网络硬件设备，设计出多种不同的网络接口，从而实现对不同设备告警信息的采集；告警信息采集完毕后由告警预处理模块对告警数据的进行预处理；告警信息处理完毕后通过JMS消息传送到数据库中；告警相关性分析模块会根据数据库中的告警数据结合规则专家系统，确定告警的类型，并交由故障处理系统进行处理。所有信息处理完毕后，用户可以从GUI界面看到告警的状态、处理的过程、故障的定性和解决方案，从而实现对网络设备告警的全面管理。

3 网络告警数据预处理(Alarm data pretreatment)

通过对接口层采集的告警数据观察分析后发现，有些告警只持续短短几秒钟，有些告警的关键字段缺失，有些告警重复多次出现,有些属性冗余等。这些数据会影响数据挖掘的准确度和效率,因此为了将原始数据转换成便于挖掘的形式,通常会对这些告警信息进行预处理，预处理的过程通常会包括过滤瞬断告警、过滤噪声、补全缺值数据、去除重复记录、完成数据类型转换等操作[5,6]。具体的内容为：

(1)过滤瞬断告警：瞬断告警是指那些历时比较短的告警，瞬断告警本身没有价值，所以不需要进行分析，要过滤掉。

(2)过滤噪声：告警库中那些无法识别或者缺失关键字段的告警数据，是无用的“噪音”数据，故此也要将其过滤掉，比如缺失时间或者告警名称的信息等。

(3)补全不完整数据：告警库中存在一些不完整但能根据告警信息的不同字段或者其他的信息来对缺失的信息进行推导，进而补全的信息，比如可以通过电路的路由信息来推导告警的网元信息等。

(4)去除重复记录：将在小段时间内产生的重复告警合并成一条告警，消除对冗余数据的分析。合并后告警的产生时间为最先发生的重复告警的开始时间。

(5)数据转换：数据转换是为了方便数据挖掘，减少数据维度，从告警库中的原始告警信息中抽取与数据挖掘相关的属性如告警的产生时间、告警的名称等。

经过这五步的预处理后，告警信息会被整理成一条独立完整的告警信息，告警信息基本形式详细信息如表1所示。

表1 告警信息内容

Tab.1 Alarm information content

名称	描述
告警ID	告警唯一标识
告警名称	告警名称
告警来源	告警发生的设备
告警类型	告警的具体类型
发生时间	告警发生的具体时间
告警级别	告警的严重程度，通常包括高、中、低

4 告警规则发现与相关性分析(Rule discovery and correlation analysis of alarm)

4.1 告警规则发现

告警规则发现是告警处理系统中的核心模块，具有不可替代的作用。告警规则发现模块实现着告警事件的匹配、过滤等工作，负责告警事件的识别、发现和挖掘等多项任务，规则发现功能的框架如图2所示。

规则发现功能是一个相对独立的模块。从流程框架图不难发现，规则发现功能总体的流程处理为规则发现功能模块首先对数据库中的告警数据进行预处理操作，即对其作加权和时间跨度的划分，然后从预处理后的数据中筛选可以满足最小支持度的频繁项目集，接着再从这些筛选得到的项目集的最大频繁项目集里筛选，找出可以满足最小可信度的规则，最后，再把这些挖掘出的规则保存在数据库中。前台不仅可以根据需求为规则引擎添加其所需大量必要的规则支持到规则引擎的规则库中，方便规则引擎系统对告警数据的相关分析；同时，还可以通过规则库管理来对挖掘出的规则进一步的筛选及增删改操作。该功能需要大量的开销，因为该功能的实现需要对数据库进行多次的扫描和处理。

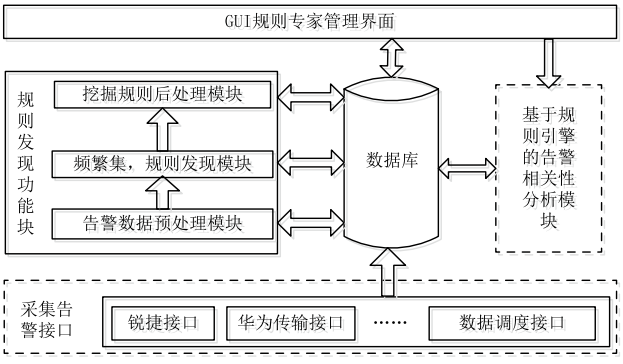

图2 规则发现功能模块框架图

Fig.2 Diagram of rule discovery function module

4.2 告警相关性分析

告警产生的原因及告警之间的内在关联是由告警相关性分析负责分析。如图3所示，即为告警相关性分析的内部处理流程。图3给出了基于规则引擎的相关性分析系统处理上报实时告警的情况。需要特别说明的是，如果有必要告警数据是可以转化为故障信息，并进入故障处理系统进行故障处理的[7]。

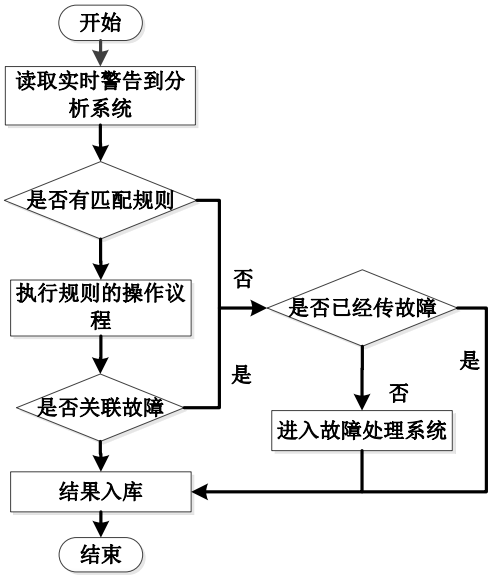

图3 相关性分析处理流程

Fig.3 Correlation analysis process

综合网络管理系统的告警相关性分析模块是由规则发现功能和相关性分析处理功能组成的。在整个系统中，告警转故障后进入的故障处理系统即工单处理系统和配置数据的接口网关或接口上传告警均与告警相关性分析模块有着直接的相关。其中，故障的维修及相关性信息的填写就是在工单处理系统中完成的。填写完成之后，再将故障处理的详情回写数据库，从而一点点的建立故障维修专家库。

5 告警规则挖掘方法(Mining method of alarm rule)

本文将关联规则算法应用在了告警数据挖掘方面，算法的主要目的是为了挖掘出告警数据中的关联规则。

根据告警原因、告警所属网元、告警级别等因素作为比较对象来划分算法中用到的告警分类集[8]。

设第k种告警可用i_k来表示，则各种告警的分类集合为$I=\{i_1,i_2,\cdots,i_m\}$；设告警分类集中第k种告警对应的权值用w_k来表示，则告所有种类的告警对应的权值集合为$W=\{w_1,w_2,\cdots,w_m\}$；告警集t是告警数据被时间跨度interval将按时间顺序划分形成的；L_k的每个元素$\{i_{j_1},i_{j_2},\cdots,i_{j_k}\}(1<j_k<m)$，满足$\sum_{n=1}^{k}w_{j_n}i_{j_n}>$最小支持度且$L_k$中每个元素是由$I$中$k$个不同的告警组成，故$L_k$称为$k$项告警频繁项目集；$L_k$中的元素$\{i_{j_1},i_{j_2},\cdots,i_{j_k}\}$之间存在的告警规则形如$\{i_{j_1},i_{j_2}\}\Rightarrow\{i_{j_3},\cdots,i_{j_k}\}$，且$\{i_{j_1},i_{j_2}\}\cup\{i_{j_3},\cdots,i_{j_k}\}=\{i_{j_1},i_{j_2},\cdots,i_{j_k}\}$，$\{i_{j_1},i_{j_2}\}\cap\{i_{j_3},\cdots,i_{j_k}\}=\varnothing$。

挖掘告警数据关联规则的算法步骤如下：

(1)首先将预处理后的告警数据用时间跨度interval按时间顺序划分为多个告警集t，并去除告警集中的那些重复告警类。

(2)在按照时间顺序划分好的告警集t中，寻找1−项告警频繁项目集L_1。

(3)在划分好的所有告警集合t中，以$k-1$项告警频繁项目集L_{k-1}为基础，寻找第$L_k(k\geq 2)$项告警频繁项目集L_k，直到得到 为空集为止。

(4)合并告警频繁项目集，$L=L_1\cup L_2\cup\cdots\cup L_m(1<n<m)$。

(5)遍历告警频繁项目集的集合L中的每一个元素L_k，计算并在L_k每个元素中寻找满足最小置信度的告警关联规则的元素，如此循环遍历，直到遍历完L中的所有元素为止。

(6)最后，将那些满足要求被筛选出来的告警关联规则放入待处理的规则库里。

关联规则挖掘算法的流程图如图4所示。

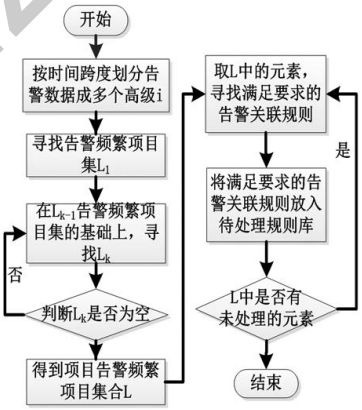

图4 关联规则挖掘算法流程图

Fig.4 Flow chart of association rule mining algorithm

上面给出了应用于挖掘综合网络管理系统的告警数据的关联规则模块的关联规则挖掘算法，该算法是经过了加权和时间约束处理的。算法模型的建立使规则发现系统的实现成为可能。

6 结论(Conclusion)

本文设计了由网管子系统、接口告警采集、JMS消息通信、故障处理系统、告警相关性分析、规则专家管理模块及GUI界面等构成的网管告警处理框架。采用过滤瞬断告警、过滤噪声、补全缺值数据、去除重复记录等方式进行数据预处理。通过告警发现和相关性分析机制对实现告警的匹配和识别，利用基于关联规则的方法对告警信息挖掘。实践证明，本文提出的处理方法能够有效的实现网管系统的告警信息的

(下转第9页)