引用本文：王峰，李兴华，许伟军，等. 一种基于多维属性的光传输网络告警压缩方法[J]. 光通信技术，2021，45(5)：7–10.

一种基于多维属性的光传输网络告警压缩方法

王 峰[1]，李兴华[1]，许伟军[2]，赵永利[3*]，左颖敏[3]

（1.国网宁夏电力有限公司 电力科学研究院，银川 750002；2.国网石嘴山供电公司，宁夏 石嘴山 753000；
3.北京邮电大学 信息光子学与光通信国家重点实验室，北京 100876）

摘要：光传输网络中每个月可能出现数百万个光层告警，给网络的运行、管理和维护带来了巨大的挑战。针对这一问题，提出了一种基于多维属性的光网络告警压缩方法。首先，根据告警发生的时间、网络位置提取告警事务集；然后，通过神经网络得到告警属性的权重用于计算告警相似度，对告警信息进行压缩。实际网络数据验证结果表明：该方法能够有效提高告警压缩率约 17%；通过采用关联挖掘算法对压缩后的告警集进行关联规则挖掘，得到的强关联告警事务增加了 55%，有利于进一步分析告警信息之间的相关性和定位故障源。

关键词：光传输网络；网络拓扑；告警压缩；告警关联

中图分类号：TN929.11　**文献标志码**：A　**文章编号**：1002-5561(2021)05-0007-04

DOI：10.13921/j.cnki.issn1002–5561.2021.05.002　**开放科学(资源服务)标识码(OSID)**：

Alarm compression method for optical transmission network based on multi–dimensional attributes

WANG Feng[1], LI Xinghua[1], XU Weijun[2], ZHAO Yongli[3*], ZUO Yingmin[3]

(1. Electric Power Research Institute of State Grid Ningxia Electric Power Co., Ltd.,Yinchuan 750004, China; 2. State Grid Shizuishan Power Supply Company, Shizuishan Ningxia 753000, China; 3. State Key Laboratory of Information Photonics and Optical Communications, Beijing University of Posts and Telecommunications, Beijing 100876, China)

Abstract: Millions of optical layer alarms may appear in optical transmission networks every month, which brings huge challenges to network operation, management and maintenance. Aiming at this problem, this paper proposes an optical network alarm compression method based on multi-dimensional attributes. First, the alarm transaction set is extracted according to the time when the alarm occurs and the network location. Then, the weight of the alarm attribute is obtained through the neural network to calculate the alarm similarity. Finally, the alarm information is compressed by alarm similarity. The actual network data verification results show that this method can effectively improve the alarm compression ratio by about 17%. By using association mining algorithm to mine association rules of compressed alarm set, the strong association alarm transaction increased by 55%, which is helpful to further analyze the correlation between alarm information and locate the fault source.

Key words: optical transmission network; network topology; alarm compression; alarm correlation

0 引言

随着光传输网络规模的扩大，光层报警次数可能在一周内就超过 100 万次，给网络的运行、管理和维护带来很大的挑战[1–2]。如果能够合理、自动地对这些告警进行压缩，那么告警分析就会容易得多，这将有助于网络管理员找到故障的根本原因，及时解决光传输网络故障，使业务得以恢复。实现告警压缩的一个有效手段是进行告警关联分析。现有的大多数研究都是基于关联规则挖掘、相关系数或互相关信息来分析告警相关性[3–6]；同时，基于关联规则挖掘的报警关联

收稿日期：2020–11–06。

基金项目：2020 年度第二批宁夏自然科学基金项目(2020AAC03487)资助；国网宁夏电力有限公司 2020 年研究开发项目计划（第一批）(5229DK19004W)资助。

作者简介：王峰(1989—)，男，汉族，硕士，高级工程师，曾长期从事电力通信网运行检修、工程建设和技术监督管理工作，现任职国网宁夏电力有限公司电力科学研究院，主要从事电力信息通信技术、网络与信息安全技术和电力物联网等方面的研究工作。

*** 通信作者**：赵永利(E-mail: yonglizhao@bupt.edu.cn)。

分析以其压缩报警量和发送高频链报警的优点得到了广泛的关注。典型的规则挖掘方法是 Apriori-like 算法,它通过重复扫描数据库来定位频繁的报警事务,并将其作为规则呈现[7-8]。然而,光传输网络的结构非常复杂,存在着复杂的拓扑连接关系和时钟拓扑关系等,如今的故障告警规则无法应对复杂的应用和服务。以前主要采用的告警关联分析都是基于时间的,对在同一时间段产生的告警进行关联分析[9],忽略了网络的拓扑因素。在实际的光网络中,某些告警类型存在上下游的关系,即一个节点的告警将引发相邻节点的告警。

为此,本文提出一种光传输网络告警压缩方法,对告警数据进行切分、聚合、分类和压缩。

1 光传输网络告警数据分析

针对包含 46 个光传输网络网元设备的某省级骨干网,本文采集了 20 万余条的告警数据,该网络中主要的告警类型及对应的告警级别如表 1 所示。涉及到的告警共有 23 种,其中紧急告警有 10 种,重要告警有 6 种,次要告警有 7 种。

本文对 23 种主要告警进行频次统计,统计结果如图 1 所示。可以看出,告警频次的分布不均匀,其中

表 1 网络中主要告警类型和告警级别

序号	告警名称	告警级别	告警类型
1	R_LOS	紧急	通信
2	ODU_SNCP_STA_INDI	次要	通信
3	MUT_LOS	紧急	通信
4	OSC_LOS	紧急	通信
5	LINK_ERR	紧急	通信
6	SUM_INPWR_LOW	重要	通信
7	R_OOF	紧急	通信
8	ODU_SNCP_PS	重要	通信
9	R_LOF	紧急	通信
10	REM_SF	次要	通信
11	OPUk_CSF	次要	通信
12	OTUk_LOF	紧急	通信
13	IN_PWR_LOW	重要	设备
14	BD_STATUS	重要	设备
15	COMMUN_FAIL	重要	设备
16	BEFFEC_EXC	次要	设备
17	LOOP_ALM	次要	设备
18	IN_PWR_HIGH	紧急	设备
19	RMON_ALM_INBADOCTS_OVER	次要	设备
20	CHAN_LOS	紧急	设备
21	OTUk_LOM	重要	处理
22	DB_RESTORE_FAIL	紧急	处理
23	SYNC_FAIL	次要	处理

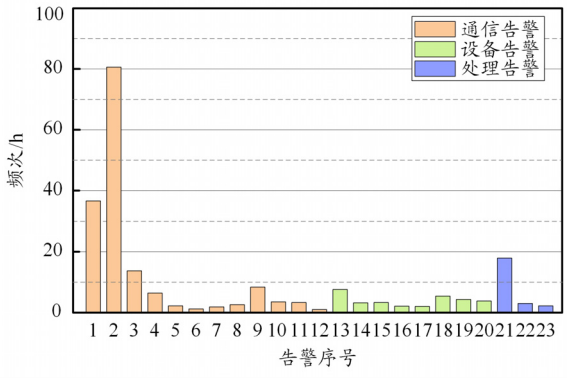

图 1 告警频次统计

发生最频繁的 3 种告警类型为 R_LOS、ODU_SNCP_STA_INDI 和 OTU2_LOM。

本文首先采用时间窗方法提取告警事务,然后采用 Apriori 算法对网络中的告警数据进行关联分析,结果如图 2 所示。可以看出,告警的关联关系较复杂,这是由于告警数据量大,充斥着冗余信息。因此,需要在关联分析之前对告警数据进行整理和合理的划分。

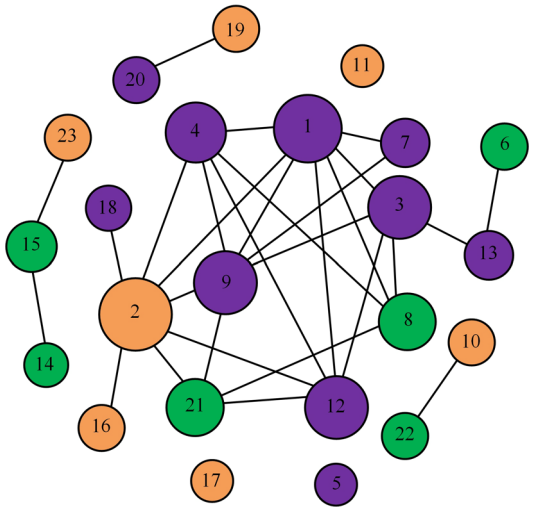

图 2 告警的关联关系

2 基于多维属性的光传输网络告警压缩方法

本文提出了一种基于多维属性的光传输网络告警压缩方法,其流程图如图 3 所示。该方法在提取告警事务时,除了考虑时间维度还考虑网络拓扑的连接关系,对告警集合进一步地划分。详细步骤如下:

①按时序相关性对原始告警数据集进行切分。采集到的告警数据集中每一条数据包含告警发生的时间、告警发生的网络位置、告警名称和告警级别等信息,将告警数据按告警发生时间进行排序,将原始告警划分成若干告警序列,使得同一告警序列中的告警

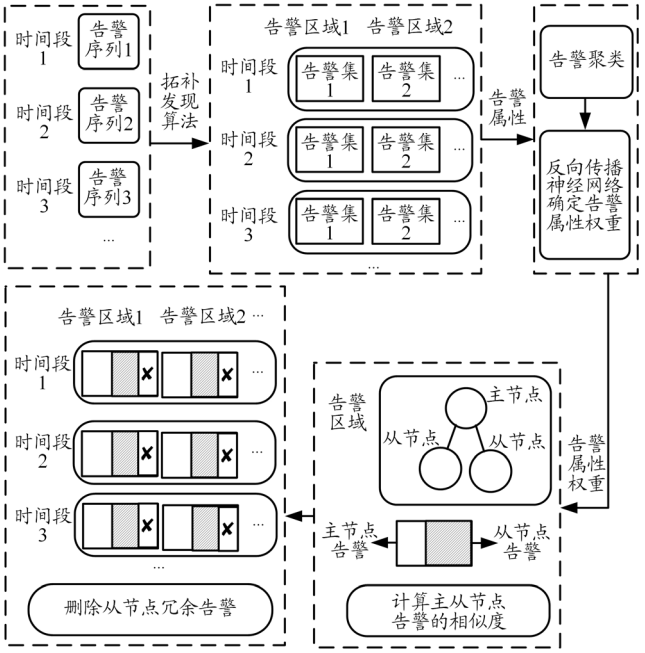

图 3 基于多维属性的光传输网络压缩方法流程图

数据在发生时间上是相近的，不同序列的告警数据在发生时间上是不相关的，即不同序列中的告警不会相互影响。

②基于拓扑发现算法确定告警传播范围。本文提出的拓扑发现算法过程如下：将网络拓扑信息以邻接表的形式存储，即每个网络节点存储与它相邻的节点信息。对每个告警序列，记为 $V=\{v_1,v_2,\cdots,v_i,\cdots\}$ 表示该告警序列所涉及到的网络节点集合，对每一个未访问的节点 v_i 先创立一个空的集合 S_i，将节点 v_i 放入到该集合 S_i 中，依次查找节点 v_i 的每个邻接节点是否在 V 中。如果存在，将该邻接节点加入到集合 S_i 中，并标记为已访问；如果不存在，越过该节点。接着选取节点集 S_i 中的下一个节点，查找该节点的邻接节点是否存在 V 中，重复上述过程。最后，该节点集会划分成一个或多个子节点集，每个子节点集称作一个节点域，在一个节点域中，它们的告警存在内在的联系，而不同的节点域，可以认为它们的告警是不会相互影响的。

③针对每个告警序列，将属于同一个节点域的告警数据进行聚合。因此，每个告警序列根据所在的节点域可进一步划分不同的子告警集。经过步骤①和步骤②，每一个子告警集中的告警数据具有时序和空间的相关性。

④根据告警属性采用 mean-shift 聚类算法对每一个子告警集进行告警分类。告警属性包括告警类型、告警级别和设备类型。将告警属性作为输入，mean-shift 算法可自动根据告警数据的特征将告警数据分成 K 类。

⑤采用神经网络算法计算出告警属性的权重。将告警属性作为输入、该条告警所属的类别作为标签，用神经网络进行拟合训练。在训练过程中，神经网络的权重会不断变化，直到训练结束，神经网络的权重会趋于稳定。根据神经网络的权重可计算出每个告警属性的权重。

⑥确定主、从节点，根据告警相似度压缩告警数据。步骤②得到了每个子告警集的节点域，对每一个节点域，本文先随机选取一个主节点并标记已访问，依次遍历该节点域的其它节点作为从节点，计算主节点与从节点的告警相似度，若从节点的告警相似度超过阈值，则删除该条告警；然后选取下一个未访问的节点作为主节点重复上述过程，直到该节点域中的所有节点都被访问完毕。本文提出的告警相似度计算方式如式(1)所示。m、n 分别代表主节点和从节点的一条告警数据，e_i 代表告警数据集 e 的第 i 个元素，其中 $e=\{\text{al\_type},\text{al\_level},\text{eq\_type}\}$，al_type、al_level 和 eq_type 分别代表告警类型、告警级别和设备类型。通过累加 2 条告警每个属性的差值来定义 2 条告警的相似度，$f(m,n)$ 值越大，说明 2 条告警越相似，相似度高的告警可以认为是冗余的，因此删除从节点中与主节点相似度高的告警。

$$f(m,n)=\sqrt{A_1(m_{e1}-n_{e1})+A_2(m_{e2}-n_{e2})+A_3(m_{e3}-n_{e3})} \quad (1)$$

其中，A_1、A_2 和 A_3 分别表示第 1 个、第 2 个和第 3 个告警属性的权重。

3 仿真结果分析

根据本文提出的光传输网络告警压缩方法，对某省级骨干网采集到的 20 万余条的告警数据进行压缩。采用的时间窗口为 5 min，利用第 2 节的步骤①、步骤②划分告警集。在每个告警集进行告警压缩之前，计算每条告警之间的相似度。本文选取了与故障最密切相关的 3 个属性(即告警类型、告警级别和设备类型[2])来计算告警相似度。每个属性在计算过程中占有一定的比重。本文利用 mean-shift 算法对告警进行分类。mean-shift 算法不需要事先指定聚类个数，根据数据特征自动将告警集分成 4 类，接着利用第 2 节中的步骤⑤得到每个属性的权重。告警属性随迭代次数的变化如图 4 所示。经过 140 次左右的迭代，3 个告警属性的权重分别稳定在 0.50、0.35 和 0.15。

本文提出的方法在划分告警集时考虑了网络的节点连接关系，与只考虑时间维度的方法对比，2 种方法的压缩率如图 5 所示。可以看出，考虑了网络的拓扑因

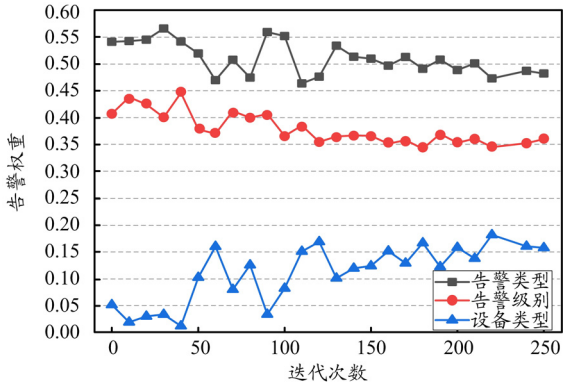

图 4 告警属性随迭代次数的变化

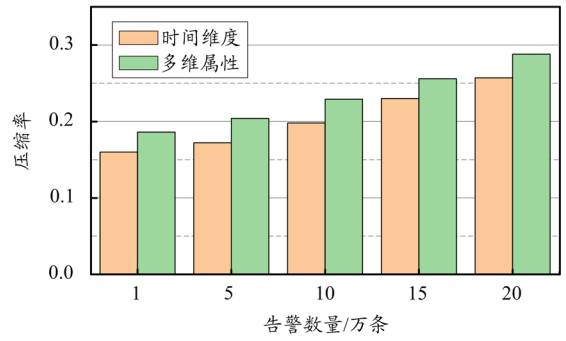

图 5 采用 2 种不同方法得到的告警压缩率

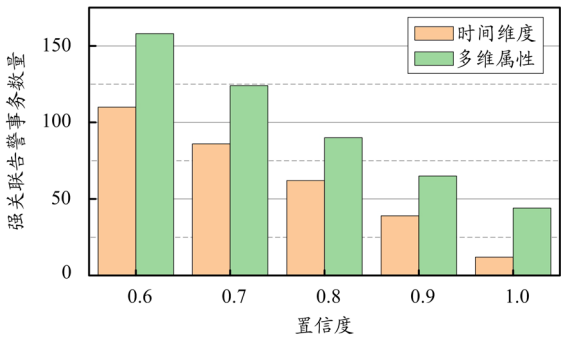

图 6 采用 2 种不同方法得到的强关联告警事务数量

素能进一步提高告警的压缩率,整体提高约 17%。

对 2 种方法(只考虑时间维度的压缩方法和本文提出的多维属性压缩方法)压缩后的告警数据采用 Apriori 算法进行关联分析,最小支持度设为 0.1,采用不同的置信度,得到最终的强关联告警事务,结果如图 6 所示。通过对比可知,采用本文提出的方法得到的强关联告警事务的数量明显增加,平均增加 55%。这是因为告警的压缩率提升,删除了大量的冗余的告警信息,使得有效的告警信息的比例增加。挖掘到的部分强关联事务如表 2 所示,经过本文提出的方法对告警数据进行处理后,得到的告警关联关系更加清晰。

表 2 部分强关联事件表

序号	强关联告警事件表
1	R_LOS、R_LOF、
2	OSC_LOS、OTU2-LOF、
3	MUT_LOS、ODU_SNCP_PS、IN_PWR_LOW
4	IN_PWR_HIGH、ODU_SNCP_STA_INDI
5	COMMUN_FAIL、BD_STATUS

4 结束语

本文提出了一种基于多维属性的光传输网络告警压缩方法。该方法在提取告警事务阶段从时间特征、空间特征等多个维度,充分考虑了告警之间的相关性,并提出了一种告警相似度的计算思路,基于告警相似度对告警事务集进行压缩。经过实际的网络数据验证,该方法能有效提高告警压缩率 17%,并通过关联挖掘算法能得到的强关联告警事务的数量增加了 55%,进一步提升后续告警关联分析的效率。

参考文献:

[1] 袁静,李大伟,陆绍雯,等. 智能监控应用平台告警关联大数据分析算法研究[J]. 电信工程技术与标准化,2019,32(5):80-84.

[2] 楼丽琪. 基于机器学习的光网络告警智能分析技术研究[D]. 北京:北京邮电大学,2019.

[3] AMANI N, FATHI M, DEHGHAN M. A case-based reasoning method for alarm filtering and correlation in telecommunication networks [C]// Canadian Conference on Electrical and Computer Engineering, May 1-4, 2005, Saskatoon, Canada. Piscataway: IEEE, 2005: 2182-2186.

[4] KLEMETTINEN M, MANNILA H, TOIVONEN H. Rule discovery in telecommunication alarm data[J]. Journal of Network and Systems Management, 1999, 7(4): 395-423.

[5] WU Jian, LI Xingmin. An effective mining algorithm for weighted association rules in communication networks[J]. Journal of Computers, 2008, 3(10): 20-27.

[6] YAN L, LI C. Incorporating pageview weight into an association-rule-based web recommendation system[C]// AI 2006: Advances in Artificial Intelligence, December 4-8, 2006, Hobart, Australia. Berlin: Springer Verlag, 2006: 577-586.

[7] 吴广伟. 通信设备告警关联规则挖掘探索—Apriori 算法的应用[J]. 价值工程,2013,32(10):183-184.

[8] 胡雪,封化民,李明伟,等. 数据挖掘中一种增强的 Apriori 算法分析[J]. 信息网络安全,2015(11):77-83.

[9] 徐前方,肖波,郭军. 一种基于相关度统计的告警关联规则挖掘算法[J]. 北京邮电大学学报,2007(1):66-70.