

面向边缘计算的电力通信网告警归并技术研究^{*}

李霁轩¹,吴子辰¹,郭 煦¹,朱鹏宇²,吴季桦³

(1.国网江苏省电力有限公司信息通信分公司,江苏 南京 210000;

2.国网电力科学研究院有限公司,江苏 南京 210012;3.北京邮电大学网络与交换国家重点实验室,北京 100876)

摘要: 电力通信网的覆盖范围及复杂程度逐渐增大,为电力通信网带来巨大的运维压力。通过部署边缘节点在边缘侧完成数据采集和信息过滤,提供计算支持,能够极大程度上缓解电力通信网管理侧压力。告警分析是运维当中的重难点问题,传统的告警分析先使用规则对于告警进行归并从而减少后续处理的工作量,但是规则的完备需要大量专家知识和人力资源的投入且存在局限性。将无监督聚类引入到电力通信边缘云部署架构下的告警归并流程当中,提出了一个新的轻量级算法,将基于密度的聚类方法(Density-Based Spatial Clustering of Applications with Noise,DBSCAN)与现有的归并规则进行结合,结果表明加入无监督学习能够显著提高告警归并的效果,切实提高了电力通信网运维中缺陷定位的准确性和完备性。

关键词: 无监督学习;DBSCAN;告警归并;边缘计算

中图分类号: TN915

文献标识码: A

DOI:10.16157/j.issn.0258-7998.201269

中文引用格式: 李霁轩,吴子辰,郭煦,等. 面向边缘计算的电力通信网告警归并技术研究[J].电子技术应用,2021,47(4):17-23.

英文引用格式: Li Jixuan, Wu Zichen, Guo Tao, et al. Research of alarm correlation technique for edge cloud computing in power communication network[J]. Application of Electronic Technique, 2021, 47(4): 17-23.

Research of alarm correlation technique for edge cloud computing in power communication network

Li Jixuan¹,Wu Zichen¹,Guo Tao¹,Zhu Pengyu²,Wu Jihua³

(1.Information and Communication Branch of State Grid Jiangsu Electric Power Co., Ltd, Nanjing 210000, China;

2.State Grid Electric Power Research Institute Co., Ltd, Nanjing 210012, China;

3.State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China)

Abstract: The demand for electric power and the coverage of power communication network has been gradually expanding, which brings the challenges that need to be addressed in the operation and maintenance of the domain. The deployment of edge nodes provides the availability of data collection, information filtering and computational support, which can heavily alleviate the pressure of management. Alarm analysis is a key and difficult problem in operation and maintenance. The traditional alarm analysis first uses rules to merge alarms, so as to reduce the workload of subsequent processing. However, the completeness of rules requires a lot of expert knowledge and human resources investment, and there are limitations. This paper proposed a novel and lightweight algorithm. In this paper, Unsupervised clustering is introduced into the alarm merging process of power communication network edge cloud computing, and the density based clustering method is combined with the existing merging rules. The experimental results show that the effect of alarm merging can be significantly improved by adding unsupervised learning, which is helpful to improve the accuracy and completeness of subsequent defect location.

Key words: unsupervised learning;DBSCAN;alarm correlation;edge computing

0 引言

我国电力行业的高效平稳发展是保证经济安全、快速、稳定发展的能源保障。人工智能时代对电力通信领

域提出了新的要求,也为电力通信管理系统(Telecom Management System,TMS)的发展提供了新方向^[1]。TMS 作为电力领域信息化产物,为整个电力系统中的电网调度、自动化、继电保护、安全自动控制、电力市场交易以及企业信息化等工作提供了坚实的基础,同时也为电力通信

* 基金项目:2020 年国家电网公司科技项目(5700-202040367A-0-0-00)

中的异常检测、路由优选等智能化应用提供支撑。

随着特高压电网、各级电网协调的统一发展,智能网的建设的需求也逐渐加强,对支撑电网信息化基础TMS系统提出了更高要求。在电力通信信息化、智能化建设和应用实践过程中,电力公司积累了海量的实时数据和运行数据,传统基于规则的缺陷处置方法难以满足智能化的需求,尤其缺乏一种对拓扑复杂、设备类型繁多的缺陷数据进行智能分析的方法^[2]。

电力通信网在信息化过程中产生了大量的数据,然而这些数据的海量增长,促使了数据归并技术(即告警归并技术)的发展。目前国内外主要使用基于规则匹配的方法进行告警归并^[3]。具体而言,就是操作员根据系统实时情况结合专家知识动态地调整告警归并规则。同时,也有基于规则匹配方法上的改进。例如,加入数据预处理和数据过滤等方法辅助告警归并^[4]。上述方法在告警数据规模较小、告警延迟低、告警类别固定等情况下,能达到很好的归并效果。但随着告警数据的海量增长,上述方法及其相关改进方法难以适应当前的数据环境。MADZIARZ A 在移动通信网领域提出了基于 K-MEANS 聚类的告警聚类方法^[5],尝试引入无监督聚类以摆脱对规则的依赖。虽然该方法无须大量人力资源的投入,但实际归并效果差强人意,且需要业务专家参与预测缺陷的数量,有着极大的局限性。

5G 技术、边缘计算、人工智能新技术的到来给电力通信领域带来了新鲜血液。新技术与电力通信领域的有机结合,对于构造电力通信新生态,解决遗留问题,节约人力资源,面对新的挑战至关重要。

本文介绍了一种基于密度聚类(Density-Based Spatial Clustering of Applications with Noise, DBSCAN)^[6]结合人工规则进行告警归并协助通信缺陷诊断的无监督学习算法。该算法具有良好的鲁棒性、轻量性,支持边缘云部署,将算法在 TMS 系统提供的数据中进行实验,结果显示算法达到了较好的效果。

1 电力通信领域相关数据介绍和归并系统方法原理

1.1 电力通信领域相关数据介绍

本文依托国网电力科学研究院建设的国家电网通信管理系统(SG-TMS)进行实验数据采样、模型调试并进行算法验证。SG-TMS 系统在国家电网通信处 2019 年通信专业重点工作部署项目中,依据人工规则制定缺陷自动派单规则,实现了传输网监视优化及缺陷自动派单功能^[7]。但是该方法依然存在局限性,具体表现在:(1)基于人工规则的方法依赖于专家经验知识积累,只能总结整理常见的影响较大的缺陷;(2)随着工作的推进,积累的规则逐渐丰富时,给运维人员带来了巨大的维护压力,同时也容易产生规则间的冲突,时间一长则难以继。

SG-TMS 提供海量的电力通信网核心数据,主要涉及

SDH、OTN 等设备,覆盖光缆(传输光路)、传输设备、通信电源和机房环境告警^[8]。数据由边缘节点进行数据采集和过滤清洗。在系统中,数据主要分为以下两部分:

(1)原始告警数据,告警数据由传输设备直接产生,经北向接口采集后存入系统数据库。但是由于系统中同时存在不同制造商生产的设备,同样的缺陷现象产生的告警存在差异,设备告警名称和实际问题的映射缺乏统一定义。

(2)缺陷单数据,缺陷单数据描述了缺陷的具体位置、实际原因和实时告警的对应关系,由具备相关专业知识的运维人员总结整理而来,是电力通信运维领域知识的结晶,同样也是人工智能算法中天然的标签数据。但是相对于庞大的原始告警数据,缺陷单数据较少,其次,缺陷单和告警由于现场设备的复杂性以及人为检修的干预可能不会直接呈现一对一的映射关系。

综上所述,TMS 系统在一定程度上实现了信息化自动化,但是仍然依赖专家干预。由于原始数据数量巨大,标注数据完备性不足,结合电力通信设备中同种缺陷导致相似的告警模式频繁出现的基本前提,考虑使用无监督学习来协助完成告警归并,采用基于密度的 DBSCAN 聚类来捕获告警簇。

1.2 归并系统方法原理

DBSCAN 算法作为经典的密度聚类算法,其在无监督密度聚类中的得到了广泛的应用。算法将点分类为核心点和非核心点,定义 1~定义 6 描述了该算法^[9]。

定义 1 邻域内的点的集合(Eps-neighborhood of a point): $N_{Eps}(p)$ 表示从 p 点出发,邻域内所有点的集合,即:

$$N_{Eps}(p)=\{q \in D \mid dist(p, q) \leq Eps\}$$

其中, D 表明全体点的集合, $dist(p, q)$ 表示 p 点和 q 点之间的欧氏距离, Eps 表示邻域半径。

定义 2 直接密度可达(Directly density-reachable):点 p 被称为从点 q 直接密度可达,当且仅当:

$$p \in N_{Eps}(q), |N_{Eps}(q)| \geq MinPts$$

其中 $MinPts$ 为给定的使 q 成为核心点的邻域内最小点数。

定义 3 密度可达(Density-reachable):如果有一系列的点 $p_1, \dots, p_n, p_1=q, p_n=p, p_{i+1}$ 到 p_i 直接密度可达,那么称点 p 从 q 密度可达。

定义 4 密度相连(Density-connected):如果点 p 和点 q 都从点 o 密度可达,则称点 p 和点 q 密度相连。

定义 5 簇(Cluster):对于集合 D ,簇 C 是 D 的一个满足以下条件的子集:

(1) $\forall p, q: \text{if } p \in C, \text{ 并且 } q \text{ 从 } p \text{ 密度可达,那么 } q \in C$ 。

(2) $\forall p, q \in C: p$ 和 q 密度可达。

定义 6 噪声(noise):噪声为不属于任意一个簇 C_i 的离散点,即:

$$\text{noise}=\{p \in D \mid \forall i: p \neq C_i\}$$

具体而言,在划分簇时,对于给定的边界距离 Eps 和最小核心节点数 $MinPts$,和非空节点集 D ,簇 C 构建时

首先检测其密度直达性。首先将核心点中具有密度直达关系的点分类给簇 C ,之后检测相连性,对剩下的点检测其与簇内任意一点的密度相连性,如果密度相连则归入簇 $C^{[10]}$ 。

在分类完成后,对于不属于任何簇的孤立点 p ,将其视为噪声^[11]。

2 告警归并系统实现

2.1 数据处理和特征构建

面对海量、复杂、标准不一的原始告警库,需要做一定的数据处理工作^[12]。规则部分根据专家知识和运维经验,将告警种类划分为:根告警、伴随告警、未被分类的告警。基于告警之间衍生关系和业务、通道的关系,系统制定了一套归并规则作为基准参考。缺陷单数据存在人工填写部分,不是规范化文本,无法简单地进行文本匹配。因此利用自动化文本分析方法^[13],对缺陷单进行信息提取。为了后续的有效性评价,系统对于缺陷单抽取的信息进一步进行分类,以此间接获取告警和缺陷的分布信息。

DBSCAN 是基于密度的算法,意味着输入的特征应当是对应空间的坐标点,或者是点之间的距离矩阵。在实际背景当中告警是连续的文本信息,因此告警的向量化过程应该体现为特征提取和特征向量之间的权重分配。

基于一个缺陷可能会引起一个或多个设备在一段时间内持续输出相似告警的特性,聚类目标是将属于拥有这种特性的属于同一缺陷的告警聚为一个簇^[14]。对告警而言,有两方面的信息较为重要:告警本身的相关参数(告警种类、发生位置、设备类型、设备位置等)以及告警时间^[15]。其中告警本身的相关参数反映了告警之间的相关程度以及告警在空间上的相近程度,告警时间是当前告警产生的时间,蕴含了缺陷发生的时间信息。对于告警本身的相关参数,使用 one-hot 方法^[16]将其映射为特

征向量,对于没有制定权重的 one-hot 来说,告警之间任意一个特征的差距映射在空间上面距离相同,在 DBSCAN 算法当中作用相同,而通过调整各个特征的权重可以反映不同特征的重要性。对于时间信息而言,显然不能简单运用 one-hot 方法,则关键在于统一时间和其他特征在特征向量上的距离关系。告警集合为 D ,告警具有 1 个时间特征和 m 个其他经过 one-hot 转换的相关参数特征。则第 i 条告警对应特征向量表述为 $D_i=(T_i, X_{i,1}, \dots, X_{i,m})$,其中 T_i 为告警 i 的一维时间特征, $X_{i,j}(1 \leq j \leq m)$ 表示告警 i 的第 j 个相关参数对应的向量。则假设告警本身相关参数的每个特征都被赋予同样的权重,如果要使得任意变化一个特征带来的空间坐标之间的距离变化等价于两个相邻告警 a 和 b 的时间差超过了时间窗口的距离变化,设变化的特征为第 t 个特征,特征维度为 n ,时间窗口为 WINDOWS,则可以得到:

$$\beta \times \text{WINDOWS} = \sum_{k=1}^n (X_{a,t,k} - X_{b,t,k})^2 \quad (1)$$

对于单纯的 one hot 向量特征,显然 $\beta=2/\text{WINDOWS}$ 。

进一步,可以得到:

$$\beta \times \text{WINDOWS} = \sum_{j=1}^m \sum_{k=1}^{n_j} \beta_j (X_{a,j,k} - X_{b,j,k})^2 \quad (2)$$

其中 n_j 表示第 j 个特征对应的维度, β_j 表示第 j 个特征的权重。可以通过调整 β 和 $\beta_j(1 \leq j \leq m)$ 的值来调整时间和其他参数之间的权重。

不同样本的距离综合考虑了告警本身相关参数距离和时间距离。以此对所有告警进行聚类,则最后得到的聚类结果应该是使得时间上较为聚集的相似告警或者是时间上极为聚集的较相似告警成为同个簇^[17]。

2.2 系统架构设计

如图 1 所示,使用基于边缘计算的电力通信告警归

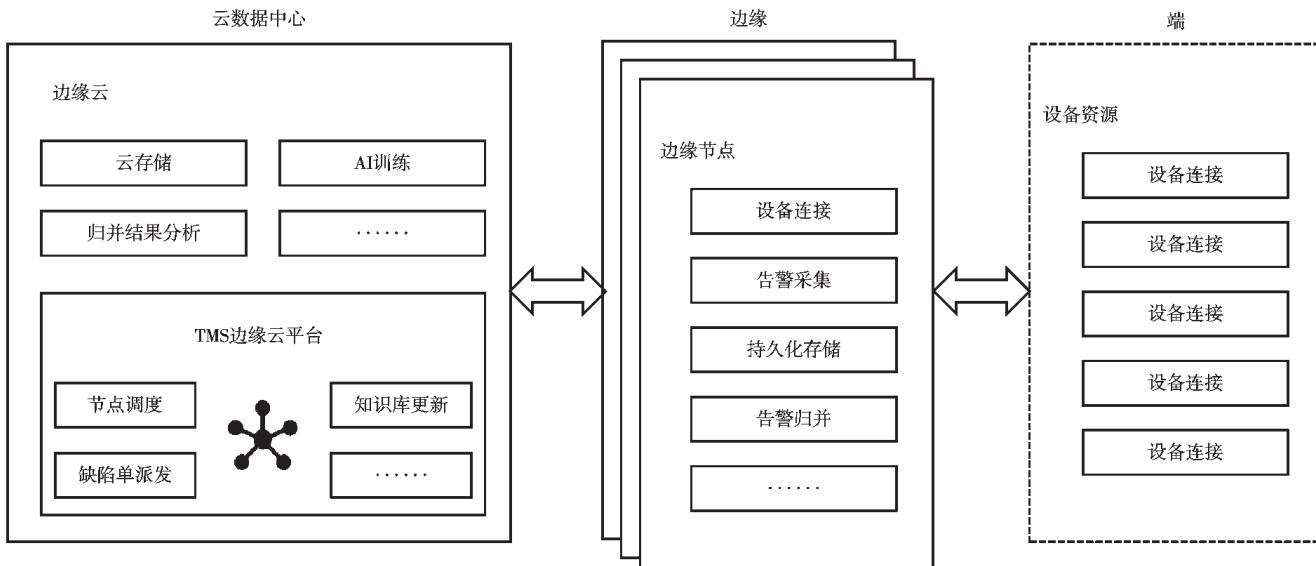


图 1 电力通信网自动派单边缘计算系统

并架构构建的计算平台,可以由边缘节点对端设备进行告警采集,在边缘节点进行告警归并,完成边缘计算资源生命周期管理,最后将告警归并结果返回云侧,支持后续缺陷定位定级、可靠性评估、迂回路由优选、缺陷单自动派发等功能,实现知识库的自动化迭代更新,助力建立电力通信行业边缘计算新生态。

图2介绍了单边缘设备上告警归并方法的基本流程。缺陷单数据经过自动化文本分析和分词得到缺陷单以及对应告警的关系,保留数据以支持后续有效性评估^[18]。

原始告警经过数据清洗、数据预处理完成特征分解,进一步对特征进行向量化,细节在2.1小节中已经讨论。然后对特征向量进行DBSCAN聚类,对得到的簇进一步使用人工规则进行告警归并^[19]。对于输出的告警归并数据使用缺陷单得到的缺陷-告警关系来进行有效性验证,根据效果调节前面聚类部分的参数。

3 系统有效性验证

3.1 算法参数 Eps 和 MinPts 的确定

对于本系统中应用的聚类算法而言,预先设定参数的选择会最大程度上影响归并表现效果,其中预先需要设定的参数为邻域半径 Eps 和密度判断阈值 MinPts。算法细节在第1小节中已经给出,显然,邻域半径 Eps 设定过大或密度判断阈值 MinPts 设定过小会使得噪声点或者多个不同的类被并入一类,邻域半径 Eps 设定过小或密度判断阈值 MinPts 过大会使得原先的同一类被划分到多个簇中。

为了确定来自单边缘节点所收集设备的本告警数据集中不同参数对算法结果的影响,表1列出了不同参数(即邻域半径 Eps 和密度判断阈值 MinPts)的不同设置对应的仅基于 DBSCAN 方法的告警归并效果以及对应的训练时间,评价指标为 V-measure 对应的 h-score、c-score 和 v-score,指标含义在3.2节中进一步阐述。

从表1的比较中可以发现,在 Eps=5 以及 MinPts=2 时,基于 DBSCAN 的告警归并方法可以获得最优,在这种情况下 c-score 也同时达到了最优。由于本系统中的

表1 不同参数对基于 DBSCAN 的告警归并性能和效率影响

实验	参数设置		性能			训练时间/min
	Eps	MinPts	h-score	c-score	v-score	
1	2	2	0.99	0.87	0.93	52.70
2	2	3	0.85	0.87	0.86	52.08
3	2	4	0.85	0.87	0.86	52.65
4	3	2	0.95	0.91	0.93	56.77
5	3	3	0.94	0.91	0.92	56.03
6	3	4	0.94	0.91	0.92	56.68
7	4	2	0.94	0.94	0.94	72.56
8	4	3	0.93	0.94	0.94	71.87
9	4	4	0.93	0.94	0.94	72.43
10	5	2	0.93	1.00	0.96	93.77
11	5	3	0.93	0.96	0.94	93.16
12	5	4	0.92	0.95	0.93	93.70
13	6	2	0.93	0.96	0.94	76.17
14	6	3	0.93	0.96	0.94	75.59
15	6	4	0.93	0.96	0.94	66.84
16	7	2	0.92	0.96	0.94	71.60
17	7	3	0.91	0.96	0.94	69.19
18	7	4	0.92	0.95	0.93	71.03
19	8	2	0.91	0.95	0.92	66.59
20	8	3	0.9	0.95	0.91	65.76
21	8	4	0.9	0.94	0.91	66.40

规则应用于基于 DBSCAN 生成的簇内,并在后续进一步细化簇,因此簇的完整性指标也就是 c-score 是应当被首要保证的。在本系统实例中,选用参数为 Eps=5 和 MinPts=2。

3.2 结果分析

本文利用 SG-TMS 已有的缺陷单作为检验聚类效果的标准,因为缺陷单由具有相关专业知识和从业经验的人员进行归并,反映了告警与缺陷的真实关系。

在实际对数据进行处理时,考虑到 SG-TMS 库中缺陷单相对有限,对原始告警进行切分,使其与缺陷单的时间范围一致,保证能对 DBSCAN 算法的聚类效果进行有效评估。由于原始告警数据的复杂性,需要结合人工规则进行辅助划分。针对不同制造商的设备对于同一种

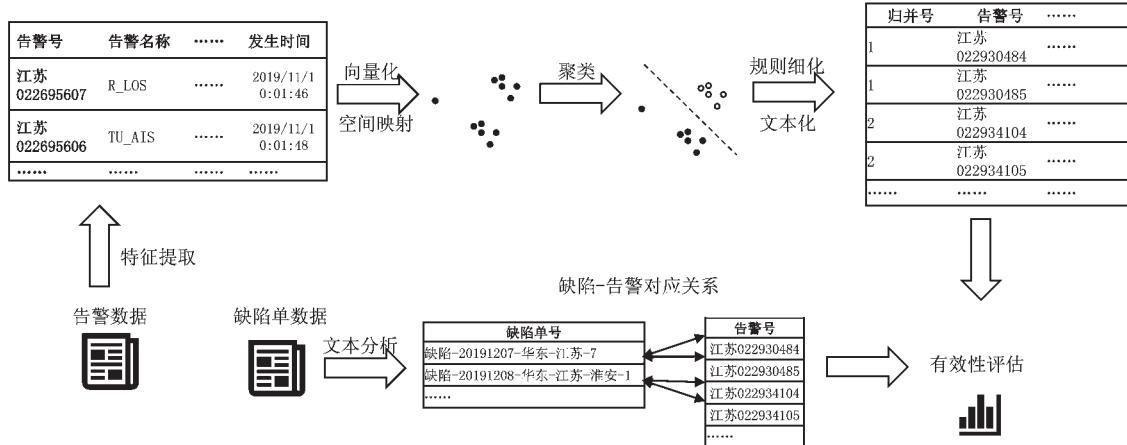


图2 单边缘设备上基于无监督聚类结合规则的告警归并流程

告警现象制定的告警文本信息不一致的情况,通过建立专家知识库^[20]实现了不同厂商设备告警的关联映射。

归并结果的有效性验证^[21]借鉴了聚类方法的评估指标,聚类方法的评价指标^[22]分为外部指标和内部指标,内部评价聚类的估计趋势,体现数据的非均匀分布程度。在电力通信系统中,比起数据的非均匀程度更加关注告警与实际场景的一致性(告警归并结果直接影响后续缺陷处理),因此借助缺陷和告警簇的分布情况通过外部指标来评价归并结果是否准确且完备。

原始告警进行聚类归并之后,选取缺陷单的对应告警与得到的告警归并结果进行比对,要求同一缺陷单对应的告警应当被归并在一起,且不同缺陷单对应的告警不应被归并在一起,选择了 V-measure 方法^[23]进行有效性评估。

$$h = 1 - \frac{H(C|K)}{H(C)} \quad (3)$$

$$c = 1 - \frac{H(K|C)}{H(C)} \quad (4)$$

$$v = \frac{2 \times h \times c}{h + c} \quad (5)$$

其中, $H(C|K)$ 是给定簇划分条件下类别划分的条件熵, $H(C)$ 是类别划分熵, $H(K|C)$ 是给定类别划分条件下的簇划分的条件熵, 即:

$$H(C|K) = - \sum_{c=1}^{|C|} \sum_{k=1}^{|K|} \frac{n_{c,k}}{n} \log\left(\frac{n_{c,k}}{n_k}\right) \quad (6)$$

$$H(C) = - \sum_{c=1}^{|C|} \frac{n_c}{n} \log\left(\frac{n_c}{n}\right) \quad (7)$$

$$H(K|C) = - \sum_{c=1}^{|C|} \sum_{k=1}^{|K|} \frac{n_{c,k}}{n} \log\left(\frac{n_{c,k}}{n_c}\right) \quad (8)$$

式中, N 表示实例总数, n_c 表示类别 c 下的实例数, n_k 表示簇 k 下的实例数, $n_{c,k}$ 表示类 c 中被划分到簇 k 的实例数。

表 2 给出了分别基于规则匹配、K-MEANS、DBSCAN 以及本系统中 DBSCAN 与规则结合的告警归并方法的特性和效果对比。其中 K-MEANS 方法需要业务专家的先验知识推测出可能发生的缺陷个数进行预设簇的数目, 其他方法不需要进行预设簇; 规则匹配方法不具备自学习能力, 只能够在规则中学习, 而 K-MEANS 方法与 DBSCAN 方法可以从数据中进行自学习。

表 2 不同告警归并方法的特性和效果对比

方法	预设簇	自学习	h-score	c-score	v-score
规则匹配	×	×	0.99	0.97	0.98
K-MEANS	√	√	0.96	0.93	0.94
DBSCAN	×	√	0.93	1.00	0.99
本系统方法	×	√	0.99	0.99	0.99

h-score、c-score 和 v-score 分别表明了归并结果的同质性、完整性和同质性与完整性的调和平均值, 取值为 0 到 1, 取值为 1 时为最理想结果。

显然观察表中结果, 可以直观地看出, 几种方法在

信息熵上的表现都能够有效消除不确定性。其中在同质性表现上, 规则匹配和本系统方法表现最佳; 在完整性表现上, DBSCAN 方法表现最佳; 综合考虑同质性与完整性表现, 本系统方法和 DBSCAN 方法表现最佳。性能表现具有可解释性。K-MEANS 方法在缺陷具有突发性的前提中并不适用, 因此性能表现都不太理想。规则匹配方法得到的归并结果基于人工经验因此归并的结果同质性较高, 但是对于规则以外的模式无法进行捕获因此完整性欠缺。DBSCAN 方法基于数据之间的粘连程度进行聚类, 会对所有数据进行归并, 因此完整性较好, 同质性欠缺。本系统方法结合了 DBSCAN 和规则方法, 在完整性和同质性上都能获得较好的性能表现。本文的基于聚类结合规则的告警归并方法在消除不确定性上表现更强, 具有自学习能力, 不需要预先人为预测缺陷数目。

表 3 给出了分别基于规则匹配、K-MEANS、DBSCAN 以及本系统中 DBSCAN 与规则结合的告警归并方法的缺陷一致性对比。

表 3 不同告警归并方法的缺陷一致性对比

方法	簇数目	归并与缺陷/%			
		一对一	一对多	多对一	多对多
规则匹配	163	76.71	7.53	13.70	2.05
K-MEANS	146	58.71	30.32	7.74	3.23
DBSCAN	122	60.27	36.99	1.37	1.36
本系统方法	153	84.93	10.96	3.42	0.68

归并与缺陷一对一: 表明归并集合中仅包含一个缺陷且一个缺陷对应的告警被归并到了同一个集合中。归并与缺陷一对多表明告警被正确归并, 显然本系统方法显著优于其他方法。

归并与缺陷一对多: 表明归并集合中包含多个缺陷但一个缺陷对应的告警被归并到了同一个集合中。归并与缺陷一对多表明部分簇被划分得过大, 可以通过细化簇来降低该比例, 可见本系统方法通过结合规则匹配比起单纯的 DBSCAN 方法降低了更多归并与缺陷一对多比例, 提高了归并与缺陷一对一比例。

归并与缺陷多对一: 表明归并集合中仅包含一个缺陷但一个缺陷对应的告警被归并到了多个集合中。归并与缺陷多对一表明部分簇被划分得过小, 可见本系统方法通过结合 DBSCAN 方法比起单纯的规则匹配降低了更多归并与缺陷多对一比例, 提高了归并与缺陷一对一比例。

归并与缺陷多对多: 表明归并集合中包含多个缺陷且一个缺陷对应的告警被归并到了多个集合中。本系统方法在归并与缺陷多对多上占比最小, 表现最优。

系统评估使用了 146 个缺陷单数据, 其中 12 个缺陷单存在重复派单的现象, 因此归并与缺陷一对多的比例较高。由聚类所生成的簇是一组数据对象的集合, 这些对象与同一个簇中的对象彼此相似, 与其他簇中的对象相异。在本实例中, 每一个簇对应着归并方法得到的

被预测为由同一个缺陷导致的告警的集合。则簇的数目对应着归并告警集合数,也就是预测的缺陷数目。在簇的数目上,K-MEANS 算法需要提前预设簇数目才能运行,预设簇数目为 146,因此生成簇的数目与缺陷总数保持一致,而其他方法生成簇的数目与实际缺陷数目有偏差。簇的数目与实际缺陷数目的一致性部分显示了归并方法的准确性。由表中结果可知,本系统的基于聚类结合规则的告警归并方法在归并与缺陷一致性表现上更强,不需要预设簇数目且生成簇与实际缺陷数目较为一致。进一步经过数据分析可得,未能正确被归并到一起的 5 个缺陷对应的告警时间间隔较大往往在 3 小时以上,不符合预设的时间聚集前提,同样也不符合运维的常规情况,推测可能采集装置存在问题或由检修动作引发^[24]。结果表明,该方法能被部署在边缘侧节点完成数据收集和告警归并,具有较强的鲁棒性和泛化能力,同时实验效果表明将无监督方法应用到告警归并流程中能够显著提高归并率和告警与缺陷的匹配率,给后续缺陷定位和故障诊断打下了坚实的基础。

4 结论

本文提出了一种面向边缘计算的基于密度的无监督聚类方法结合人工规则应用于电力通信网中的告警归并方法,讨论了告警特征空间构建和时间与相关参数指标的权重矫正方法,并且引进了有效性评价指标表明该方法对时间上聚集和性质相似的告警数据有着较好的效果。后续研究可以通过加入空间拓扑关系特征^[25]以及调整特征^[26]权重来进一步提高归并效果。

参考文献

- [1] 刘建戈,张鹏宇,姜蒙娜,等.适应电网运行模式的电力通信管理系统演进分析[J].电力信息与通信技术,2020,18(9):111-117.
- [2] 林志方.试论电力通信网运维的信息化与电子化[J].计算机产品与流通,2020(9):63.
- [3] 韩啸.动环监控系统告警清理解决方案[J].电信技术,2014(6):81-84.
- [4] 黄颖祺,邱玉祥,胡游君,等.日志采集归并技术在电力安全管理中的应用[J].电力信息化,2013(4):46-50.
- [5] MAZDZIARI A. Alarm correlation in mobile telecommunications networks based on k-means cluster analysis method[J]. Journal of Telecommunications and Information Technology, 2018, 2(2): 95-102.
- [6] 田力,向敏.基于密度聚类技术的电力系统用电量异常分析算法[J].电力系统自动化,2017,41(5):64-70.
- [7] 易鸣.江西电信传输网告警自动派单功能研究与实施[D].南京:南京邮电大学,2016.
- [8] 张振伟,何杰,喻鹏,等.电力通信管理系统数据采集功能的研究与设计[J].电力信息与通信技术,2013,11(12):94-99.
- [9] ESTER M, KRIEGEL H P, SANDER J, et al. A density-based algorithm for discovering clusters in large spatial databases with noise[C]. Proceedings of the Second International Conference on Knowledge Discovery & Data Mining, 1996.
- [10] 伍育红.聚类算法综述[J].计算机科学,2015,42(S1):491-499,524.
- [11] YANG Y, WANG Y, WEI Y. Adaptive density peak clustering for determining cluster center[C]. 2019 15th International Conference on Computational Intelligence and Security(CIS), 2019.
- [12] 陈诚伟,曹德发,张伟渊.大电网监控信息管理系统建设与应用[J].通讯世界,2019,26(11):193-194.
- [13] CHEN P C P, PONOCKO J, MILOSEVIC N, et al. Towards application of text mining for enhanced power network data analytics-part i: retrieval and ranking of textual data from the internet[C]. Mediterranean Conference on Power Generation, Transmission, Distribution and Energy Conversion (MedPower 2016). IET, 2016.
- [14] JIANG J, CHENG Y, FENG L, et al. Electricity anomalies diagnose method based on the electricity information acquisition system[C]. 2019 IEEE 4th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), 2019.
- [15] 周建华.基于告知告警数据的电网事故演化决策模型研究[D].成都:电子科技大学,2020.
- [16] JACKSON E, AGRAWAL R. Performance evaluation of different feature encoding schemes on cybersecurity logs[C]. 2019 SoutheastCon, 2019.
- [17] Lin Xiqiao, Wu Wanlu, Zeng Bo, et al. Analysis of large-scale electricity load profile using clustering method[C]. 2018 IEEE 15th International Conference on Networking, Sensing and Control(ICNSC), 2018.
- [18] GIONIS A, MANNILA H, TSAPARAS P. Clustering aggregation[J]. ACM Transactions on Knowledge Discovery from Data, 2007, 1(1): 4-es.
- [19] 黄颖祺,邱玉祥,胡游君,等.日志采集归并技术在电力安全管理中的应用[J].电力信息化,2013,11(4):46-50.
- [20] 刘剑,陈一超,江虹.基于规则的通用专家知识库故障诊断方法[J].计算机与数字工程,2010,38(6):72-76.
- [21] FOX W R. Finding groups in data: an introduction to cluster analysis[J]. Journal of the Royal Statistical Society: Series C(Applied Statistics), 1991, 40: 486-487.
- [22] HOU J, LIU W. Evaluating the density parameter in density peak based clustering[C]. 2016 Seventh International Conference on Intelligent Control and Information Processing (ICICIP), 2016.
- [23] NOWOSAD J, STEPINSKI T F. Spatial association between regionalizations using the information-theoretical V-measure[J]. International Journal of Geographical Information Science, 2018, 32(12): 2386-2401.
- [24] 梁清清,李世译.电力运维安全隐患排查治理措施研究[J].黑龙江科学,2020,11(22):108-109.

- [25] RIVERA J, NASIRIFARD P, LEIMHOFER J, et al. Automatic generation of real power transmission grid models from crowdsourced data[J]. IEEE Transactions on Smart Grid, 2019, 10(5): 5436–5448.
- [26] SHI Z, ZENG Y, SUN L. Operation and maintenance analysis for power communication networks based on big data[C]. 2016 China International Conference on Electricity Distribution(CICED), 2016.

(上接第 8 页)

- [2] ETSI.GS MEC 003. Mobile Edge Computing(MEC); framework and reference architecture[S]. 2016.
- [3] ETSI.GR MEC 004. Mobile Edge Computing(MEC); service scenarios[S]. 2015.
- [4] ETSI.GR MEC 010-2. Mobile Edge Computing(MEC); Part 2: Application lifecycle, rules and requirements management[S]. 2017.
- [5] ETSI.GR MEC 017. Mobile Edge Computing(MEC); Deployment of mobile edge computing in an NFV environment[S]. 2018.
- [6] 沈彬, 李海花, 高腾. 工业互联网技术洞察[J]. 中兴通讯技术, 2020, 26(6): 34–37.
- [7] 李庆, 刘金娣, 李栋. 面向边缘计算的工业互联网工厂内网络架构及关键技术探究[J]. 电信科学, 2019, 35(S2): 160–168.
- [8] 钮丹, 乐兰娟, 李欣, 等. 边缘计算在工业互联网平台建设中的应用探讨[C]. 中国造船工程学会会议论文集, 2019.
- [9] 穆琰博, 柴瑶琳, 宋平, 等. 边缘计算发展现状及标准体系研究[J]. 信息通信技术, 2020, 14(4): 23–30.
- [10] 马吉军, 贾雪琴, 寿颜波, 等. 基于边缘计算的工业数据

(上接第 16 页)

- Service and Applications(BigDataService). IEEE, 2016.
- [9] SHAO Z, WANG L, ZHANG H. A fault line selection method for small current grounding system based on big data[C]. 2016 IEEE PES Asia-Pacific Power and Energy Engineering Conference(APPEEC). IEEE, 2016.
- [10] KNORR E, GRUMAN G. What cloud computing really means[J]. InfoWorld, 2008(7): 20.
- [11] DENG R, LU R, LAI C, et al. Optimal workload allocation in fog-cloud computing toward balanced delay and power consumption[J]. IEEE Internet of Things Journal, 2017, 3(6): 1171–1181.
- [12] ZHAO Z M, LIU F, CAI Z P, et al. Edge computing: platforms, applications and challenges[J]. Computer Research and Development, 2018, 55(2): 327–337.
- [13] HAMEED Z, HONG Y S, CHO Y M, et al. Condition monitoring and fault detection of wind turbines and related

(收稿日期: 2020-12-31)

作者简介:

李霁轩(1991-), 男, 硕士, 工程师, 主要研究方向: 边缘计算、电力通信调度管理、人工智能无监督学习。

吴子辰(1988-), 男, 硕士, 高级工程师, 主要研究方向: 边缘计算, 电力通信。

郭焘(1988-), 男, 工程师, 工程硕士, 主要研究方向: 电力通信。

采集[J]. 信息技术与网络安全, 2018, 37(4): 91–93.

- [11] 路璐. 工业互联网中数据集成和边缘处理技术国际态势分析[J]. 高科技与产业化, 2020(2): 42–48.
- [12] 孙念, 傅为. 对 5G 工业互联网的边缘计算技术架构与应用探讨[J]. 数字通信世界, 2020(3): 211, 213.
- [13] 黄智国, 李宏杰, 钱岭, 等. 一种云边协同的工业应用云化方案[J]. 电信科学, 2019, 35(S2): 84–88.
- [14] 李辉, 李秀华, 熊庆宇, 等. 边缘计算助力工业互联网: 架构、应用与挑战[J]. 计算机科学, 2021, 48(1): 1–10.
- [15] 刘炜, 张为等. 基于边缘智能 2.0 的工业互联网平台架构及应用[J]. 光通信研究, 2020.
- [16] 乔爱锋. 边缘计算多业务融合架构及关键技术[J]. 通信与信息技术, 2020(6): 27–28, 33.

(收稿日期: 2020-11-03)

作者简介:

甘雨莹(1991-), 通信作者, 女, 硕士, 工程师, 主要研究方向: MEC 关键技术的研究, E-mail: ganyy@chinatelecom.cn。

孙慧(1986-), 女, 硕士, 工程师, 主要研究方向: 工业互联网的工业 PON、MEC 和 5G 典型行业应用。

梁真铭(1993-), 男, 本科, 工程师, 主要研究方向: 接入网网络重构。

algorithms: a review[J]. Renewable and Sustainable Energy Reviews, 2009, 13(1): 1–39.

- [14] LIU W Y, TANG B P, HAN J G, et al. The structure healthy condition monitoring and fault diagnosis methods in wind turbines: a review[J]. Renewable and Sustainable Energy Reviews, 2015, 44: 466–472.
- [15] 任玉亭. 振动监测技术在风力发电机组的应用[J]. 内蒙古电力技术, 2010, 28(2): 8–12.
- [16] DC 轴承故障检测数据集[EB/OL]. <http://www.js.dclab.run/v2/cmptDetail.html?id=248>.

(收稿日期: 2021-01-25)

作者简介:

方科(1979-), 男, 硕士, 高级工程师, 主要研究方向: 信号处理及敏捷智能计算。

宁琨(1985-), 男, 硕士, 高级工程师, 主要研究方向: 智慧风电应用。

潘灵(1981-), 男, 硕士, 高级工程师, 主要研究方向: 嵌入式平台软件、分布式计算框架。

版权声明

经作者授权，本论文版权和信息网络传播权归属于《电子技术应用》杂志，凡未经本刊书面同意任何机构、组织和个人不得擅自复印、汇编、翻译和进行信息网络传播。未经本刊书面同意，禁止一切互联网论文资源平台非法上传、收录本论文。

截至目前，本论文已经授权被中国期刊全文数据库（CNKI）、万方数据知识服务平台、中文科技期刊数据库（维普网）、DOAJ、美国《乌利希期刊指南》、JST 日本科技技术振兴机构数据库等数据库全文收录。

对于违反上述禁止行为并违法使用本论文的机构、组织和个人，本刊将采取一切必要法律行动来维护正当权益。

特此声明！

《电子技术应用》编辑部

中国电子信息产业集团有限公司第六研究所