

# 基于云计算的智能电网信息异常数据挖掘方法

刘晓霞, 周明哲

(国网内蒙古东部电力有限公司信息通信分公司, 内蒙古 呼和浩特 010010)

**摘要** 为优化和完善智能电网内部信息异常数据挖掘, 本文在数据挖掘中应用云计算技术。基于智能电网信息流特性, 将异常数据输入空间映射至特征空间, 从而实现异常数据线性判定函数的构造。实现对偶问题的分布式求解后, 对智能电网信息异常情况进行全面监测, 借助云计算技术, 对智能电网内部的信息异常数据做挖掘工作。研究发现, 云计算技术下的智能电网信息异常数据挖掘, 能够实现异常数据挖掘质量与精度的全面提升。

**关键词** 智能电网; 云计算; 信息异常; 数据挖掘

**中图分类号**: TP3

**文献标识码**: A

**文章编号**: 2097-3365(2024)02-0010-03

智能电网, 其实就是应用信息与通信技术智能化、全系统管理电力系统的输电、发电、变电及用电情况。智能电网为我国电力行业产品创新与有效管理提供了技术保障, 但在电力系统建设期间, 智能电网信息采集是否完整和真实, 通常会对其运行效果产生重大影响<sup>[1]</sup>。所以, 有必要深度挖掘智能电网信息异常数据。陈杰等学者结合孤立森林算法与 Wasserstein 生成式对抗网络, 并提出异常数据辨识方法。对距离生成对抗网络(英文简称 WGAN)交互训练生成器与判别器进行有效融合, 将电力统计数据分布特征提取出来, 借助隔离树算法识别均衡数据, 再对模型识别查全率、准确率和查准率等展开比较, 研究结果发现, 这一异常数据强化算法能够大大提升异常数据识别能力。学者李婧等在大规模电网异常负荷聚类中运用自适应 k-means++, 并检验了不同聚类状态下的聚类结果, 有效防止人工设置产生的聚类结果偏移, 实现异常负荷分类精度的有效提升。本文在数据挖掘中应用云计算技术, 基于智能电网信息流特性, 将异常数据输入空间映射至特征空间, 从而实现异常数据线性判定函数的构造。

## 1 智能电网信息数据挖掘算法和云计算知识架构

### 1.1 信息数据挖掘算法的知识架构

信息数据挖掘, 指的是数据仓库或数据库内, 通过各种分析技术, 分析、归纳和整理以往积累的繁杂数据, 提取有用信息数据, 为决策者提供有用的资讯资料, 其中包括离群分析、分类算法、关联规则等。当前, 针对电力系统的特点, 提出了一种基于概率论的电力系统建模方法。评价适合于只对少数网络信息数据进

行处理的电力系统模拟模式等, 面对海量数据处理, 应改进传统算法, 可采用预测算法与聚类算法, 改进传统信息数据挖掘算法, 最后, 将该方法用于电力系统的负载预报和电力消费的研究中。

### 1.2 基于云计算的智能电网信息数据挖掘

本项目拟从云计算的基本特征出发, 针对智慧电网应用的实际应用, 提出了一种新的解决方案: 云基础设施层、平台层、应用软件集合层、业务接入层。由于智能电网的网络规模很大, 任何一个节点或者一个计算机终端的运算需求都必须通过云端来进行配置; 这样一来, 中央的压力就更大了, 而且权限的判断也很繁琐, 极易导致网络阻塞。因为电力系统有着清晰的分层空中、分级管理层次, 所以, 可根据级别创建主云与子云, 以限定权限与资源分配, 减少了对系统中多余的资源分配和权限的管理, 实现云计算资源的合理化, 同时该系统还可以实现对电力系统的实时处理。

智能电网信息数据驱动体系中, 通过云计算技术搭建智能电网信息异常数据处理平台, 借助该云计算软件, 对数据计算单元进行设置, 以展开数据处理, 同时在输出端与区块链系统相连接, 区块链选择分布式数据存储, 表现为加密算法、对点传输以及共识机制等特性, 将其应用在智能电网中, 既有助于永久性存储数据, 又具备数据保密特性, 便于用户数据共享<sup>[2]</sup>。

## 2 基于云计算的智能电网信息异常数据挖掘方法

### 2.1 监测智能电网信息异常数据

在构建智能电网的过程中, 节点  $j$  是一个数据点,  $j$  表示的是信息流任意数据点,  $a_j$  表示的是智能电网信息流,  $E^n$  表示的是信息流输入空间,  $n$  为空间维度。

通过  $\phi(\cdot):E^n \rightarrow G$  映射电网信息流输入空间至特征空间, 因此, 在该特征空间构建了一种用于电力系统的信息流动的线性决策方程, 具体公式为:  $f(a)=\varepsilon^T \cdot \phi(a)-p$ , 该公式中,  $\phi(a)$  表示智能电网的信息流映射向量, 线性判定函数  $f(a)$  偏移量采用  $P$  表示,  $T$  时刻下  $f(a)$  归一化矢量采用  $\varepsilon^T$  表示。转化  $p$ 、 $\varepsilon^T$  的二次规划问题为对偶问题, 通过精准判断其极值条件, 对智能电网信息异常数据进行有效挖掘, 具体公式为:

$$\varepsilon = \sum_{j=1} \lambda_j \phi(a_j)$$

$$f(a) = \varepsilon - p = \sum_{j=1} \lambda_j K(a_j, a) - p$$

在该公式中,  $K(a_j, a)$  为智能电网信息异常数据监测函数, 信息异常数据监测响应系数由  $\lambda_j$  表示<sup>[3]</sup>。

## 2.2 智能电网信息异常数据预处理

明确电网信息异常数据隶属度, 对信息异常数据的多维度属性进行了高效的抽取, 然后对其进行了清晰的划分, 并用公式对信息异常数据进行了计算, 从而对其求解, 公式是:

$$Q = \sum_{i=1}^m \sum_{a=1}^m (\delta_{pq})^x d_{ab}^2$$

该公式中,  $x$  代表的是智能电网信息异常数据的目标函数值,  $\delta_{pq}$  为异常信息数据特征向量隶属度,  $d_{ab}^2$  代表的是第  $a$ 、 $b$  个异常信息数据之间的距离。以信息异常数据为基础, 以此为基础, 有效地提取智能电网中的异常信息特征。公式表示是:

$$\partial(X, Y) = \phi \sum_{i=1}^M (x_i - x_q) + \sum_{i=M+1}^M \delta_{pq}$$

该公式中,  $M$  是指网络状态下的异常值与簇群中心的间距,  $\phi$  代表了网络异常状态下的权重。根据信息异常数据  $x$  的资讯熵, 将智能电网信息的异常数据输入云端, 并对其参数关联程度的计算, 具体为:

$$U_\beta = \frac{1}{\beta} \times \sum_{x=1}^{\beta} R(x) \times \frac{1}{\partial(X, Y)}$$

该公式中,  $\beta$  代表的是信息异常数据特征相关参数关联系数<sup>[4]</sup>。

## 2.3 智能电网异常信息数据挖掘算法的设计

利用云计算技术, 预处理智能电网内部异常数据信息, 并对其进行高效挖掘。在对智能电网内部异常信息数据进行挖掘期间, 将信息异常数据的复杂性定义为  $f(pq)$ , 那么, 智能电网数据库包括:  $p$  个异常数据库和  $q$  个正常数据库, 经过反复筛选之后, 得到了

如下的公式:  $f(pq)=f(qo), p \neq q$ , 该公式中, 异常信息数据特征数由  $o$  表示。

在筛选出的异常信息资料的基础上, 准确地计算出智能电网的异常信息资料在空间上的分布, 求解电力系统中的异常信息的权重, 具体公式是:

$$\begin{cases} \prod_{i=1}^7 \omega_i = 1 \\ \frac{\xi_i}{\omega_i} = const \end{cases}$$

该公式中,  $\omega_i$  表示的是第  $i$  类别中的智能电网信息异常数据对应权值, 而  $\xi_i$  代表的是智能电网信息异常数据的波动系数<sup>[5]</sup>。

## 3 实验比较

### 3.1 实验环境与数据搭建

为对云计算下的智能电网信息异常数据挖掘性能进行有效验证, 通过实验平台的搭建, 展开测试, 采用 Windows 11 系统, 同时搭载频率为 200GHz、8GB 内存的 CORE i8 系统处理器, 使用 MySQL 2020 作为电力网络中的异常数据源, 并使用 Java 软件来建立试验数据源。

通过对一个实际的电力系统的实际运行资料进行分析, 选取了 200 条异常资料, 其中, 接地故障 49 条, 电压故障 34 条, 共计 26 次故障和 91 次发电设备故障。

### 3.2 实验指标设置

为确保电网信息异常数据挖掘质量, 采用准确率指数来度量电力系统中异常信息的挖掘准确度, 如果存在较高的查准率, 说明电网具有越高的信息异常数据挖掘精度。查准率计算公式为:

$$P_{recision} = \frac{h_b}{T_r} + G_p$$

其中,  $G_p$  为电网信息异常数据特征关联度,  $T_r$  为电网信息异常数据采集周期, 而  $h_b$  代表的是智能电网的异常资料特征。

对电力系统中的异常信息资料进行了挖掘, 利用冗余指数对电力系统中的异常信息资料进行分析, 计算公式为:

$$O = \frac{x(s \times P_{recision})}{h_b} \times \frac{T_r}{\psi_j}$$

该公式中,  $\psi_j$  为电网异常信息数据挖掘均衡性,  $S$

则为电网异常信息数据观测向量。

### 3.3 结果分析

该实验为突出基于云计算的电网信息异常数据挖掘优越性,比较了两种基于适应性的K-means++和基于改进的支持型对抗网络两种不同的数据挖掘方法。在模拟实验平台中,对三种方法的异常数据挖掘冗余度、查准率展开挖掘,最后发现,在检验智能电网信息异常挖掘精度的同时,由于数据规模不断增大,三种算法的准确率也在不断提高,其中,基于云计算的电力系统异常挖掘准确率最高。针对200个电力系统中的异常数据,其准确率达到98.9%;当有200个电力

网络资料不正常时,该算法的召回率仅为72.3%,适用于进行异常数据的分析。在图1中显示了电力系统信息的异常数据挖掘的冗余度。

通过分析图1可知,针对当前电力系统中的异常数据日益增多,现有三种挖掘方法的冗余程度都小于50%,采用基于K均值和基于改进的产生式对抗网络的电力系统异常挖掘技术,可使电力系统的异常挖掘冗余超过20%,应用本文所提出的方法对信息异常数据进行挖掘,虽然有200个信息异常,但冗余还是不到10%,可见基于云计算的智能电网信息异常数据挖掘,有助于数据挖掘质量的提升。

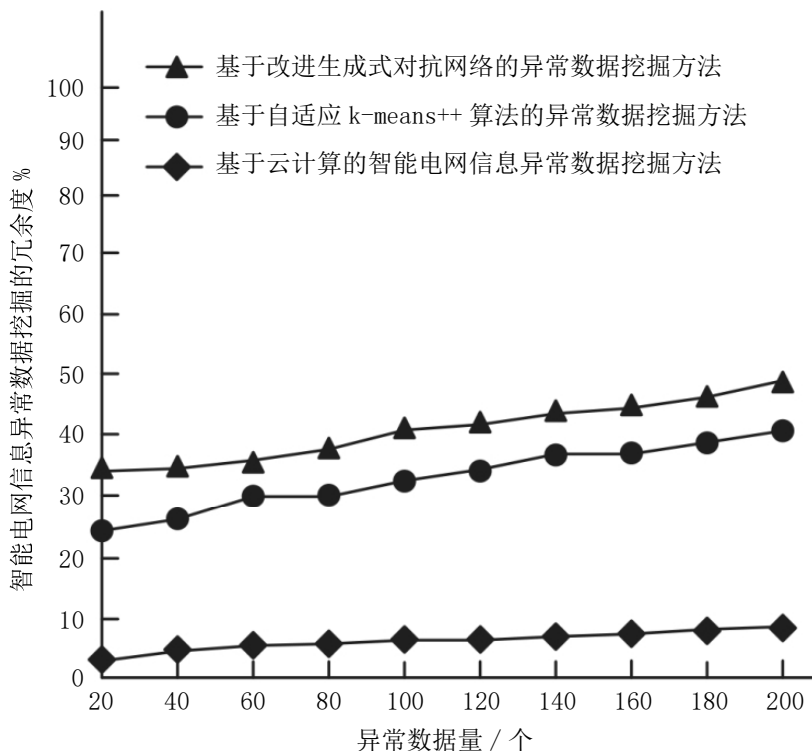


图1 电网信息异常数据挖掘的冗余度

## 4 结语

基于云计算下的智能电网信息异常数据挖掘中,深入分析智能电网中的异常数据,能够实现对信息异常数据准确性、挖掘效率的全面提升。但该研究仍有一系列缺陷,主要是智能电网信息异常数据通常会由于时间的变迁而发生变化,所以挖掘实时状态的电网信息异常数据是今后需重点研究的课题之一。

### 参考文献:

[1] 张亮,武秋阳,方圆,等.物联网、云计算在智能

电网信息通信调度中的运用初探[J].通讯世界,2018(11):114-115.

[2] 于继超.物联网、云计算在智能电网信息通信调度中的应用分析[J].信息通信,2018(07):153-154.

[3] 肖胜.基于云计算的智能电网调度系统设计[J].电源技术,2018,42(02):288-290.

[4] 黄海丽.智能电网信息和通信技术关键问题探讨[J].中国战略新兴产业,2017(24):98.

[5] 陈静,丁学文,荣秀婷,等.基于云计算和光通信网络下智能电网信息模型的建立[J].自动化与仪器仪表,2017(04):181-184.