

DOI:10.19322/j.cnki.issn.1006-4710.2025.03.008

<https://xuebao.xaut.edu.cn>

引文格式:张科峰,马文华,郑庆明,李云红,李丽敏,苏雪平,冯准若.水电站红外监控系统的高效数据汇聚传输聚合算法[J].西安理工大学学报,2025,41(3):381-389.

ZHANG Kefeng, MA Wenhua, ZHENG Qingming, LI Yunhong, LI Limin, SU Xueping, FENG Zhunruo. Efficient data aggregation transmission aggregation algorithm for infrared monitoring system of hydropower station[J]. Journal of Xi'an University of Technology, 2025, 41(3): 381-389.

# 水电站红外监控系统的高效数据汇聚传输聚合算法

张科峰<sup>1</sup>, 马文华<sup>1</sup>, 郑庆明<sup>1</sup>, 李云红<sup>2</sup>, 李丽敏<sup>2</sup>, 苏雪平<sup>2</sup>, 冯准若<sup>2</sup>

(1. 国家能源集团新疆吉林台水电开发有限公司, 新疆 伊犁 835000;

2. 西安工程大学 电子信息学院, 陕西 西安 710048)

**摘要:** 针对水电站监控画面数据量大、传输链路受限、内部环境复杂和传输效率低下等问题,提出了一种基于 MapReduce 差分隐私的 K-means 聚类算法。首先在红外系统所提取的数据中引入 Hadoop 下的 MapReduce 并行框架,以便将数据集转换生成键值对,从而提高容错率和传输效率。其次在迭代聚簇中加入 Laplace 噪声计算聚类中心,以实现数据的隐私保护。然后在传输过程中根据传输链路带宽情况选择无损压或有损压分段式传输,以保障数据传输的实时性和完整性。最后,我们在吉林台流域梯级水电站红外监控系统中进行算法测试和性能评估,基于 MapReduce 差分隐私的聚类汇聚传输算法相较于传统 K-means 聚类算法在不同隐私预算下进行统计量指标对比。结果表明,改进后的聚类算法聚类可用性可达到 94.5%,聚类程度更好,数据汇聚传输的效果更佳。

**关键词:** 水电站; 红外监控系统; 差分隐私; 数据传输; 聚类算法

**中图分类号:** TP391

**文献标志码:** A

**文章编号:** 1006-4710(2025)03-0381-09

## Efficient data aggregation transmission aggregation algorithm for infrared monitoring system of hydropower station

ZHANG Kefeng<sup>1</sup>, MA Wenhua<sup>1</sup>, ZHENG Qingming<sup>1</sup>, LI Yunhong<sup>2</sup>,  
LI Limin<sup>2</sup>, SU Xueping<sup>2</sup>, FENG Zhunruo<sup>2</sup>

(1. CHN ENERGY Xinjiang Jilintai Hydropower Development Co., Ltd., Ili 835000, China;

2. School of Electronics and Information, Xi'an Polytechnic University, Xi'an 710048, China)

**Abstract:** Aiming at the problems from large amount of data, limited transmission links, complex internal environment and low transmission efficiency of hydropower station monitoring images, a K-means clustering algorithm based on the MapReduce differential privacy is proposed. Firstly, the MapReduce parallel framework under Hadoop is introduced into the data extracted by the infrared system, to convert the data set into key-value pairs, in order to improve the fault tolerance rate and transmission efficiency. Secondly, Laplace noise is added to the iterative clustering for calculating the clustering center, so as to realize the privacy protection of data. Then, according to the broadband situation of the transmission link, the lossless voltage or lossy voltage segmen-

收稿日期: 2023-12-20; 网络首发日期: 2024-04-23

网络首发地址: <https://link.cnki.net/urlid/61.1294.n.20240422.1806.006>

基金项目: 国家自然科学基金资助项目(62203344); 陕西省自然科学基金基础研究重点资助项目(2022JZ-35); 陕西高校青年创新团队项目

第一作者: 张科峰,男,高级工程师,研究方向为控制及检测。E-mail:1505261515@qq.com

通信作者: 李云红,女,博士,教授,研究方向为红外热像测温技术、人工智能、深度学习等。E-mail: hitliyunhong@163.com

ted transmission is selected to ensure the real-time and integrity of data transmission. Finally, we test the algorithm and evaluate its performance in the infrared monitoring system of cascade hydropower stations in Jilin Tai Basin. The clustering aggregation transmission algorithm based on MapReduce differential privacy is compared with the traditional K-means clustering algorithm under different privacy budgets. The results show that the clustering availability by the improved clustering algorithm can reach 94.5%, that the degree of clustering is better, and that the effect of data aggregation transmission is better.

**Key words:** hydropower station; infrared monitoring system; differential privacy; data transmission; clustering algorithm

红外监控系统(infrared monitoring system)通常由红外摄像机、红外光源、热成像传感器、处理器、图像分析软件和监控显示设备组成<sup>[1]</sup>。目前,红外监控系统主要应用在水电站监测、煤矿监控、林业防火监控等方面<sup>[2-3]</sup>。在红外监控系统中,数据汇聚传输时,水电站监控画面数据量大、传输链受限、内部环境复杂等因素,往往导致系统中的传感器节点产生大量冗余数据,从而导致数据传输汇聚效率低下<sup>[4-6]</sup>。

目前,国内外的一些学者对数据的高效传输进行了研究。Yang 等<sup>[7]</sup>提出一种基于本地距离隐私的聚类方法,通过一维距离空间引入噪声,在保证隐私的同时提升聚类效果,王辉等<sup>[8]</sup>利用 Hadoop 云计算平台实现了 K-means 聚类算法。还有一些学者将聚类算法转接到 MapReduce 框架下进行设计<sup>[9]</sup>,提出一种 mrk-means 的通道算法,该算法只需读取一次数据,因此聚类速度比现有的编程模块快很多。高榕等<sup>[10]</sup>提出了一种基于 Hadoop 分布式改进的 K-means 算法的聚类中心,克服了传统 K-means 算法因初始中心随机,而易陷入局部最优解的问题。但目前国内外并没有一种改进的 K-means 聚类算法在实现数据隐私保护的前提下能够实现高

效传输汇聚。

针对以上存在问题,利用吉林台流域梯级水电站这一平台,对 MapReduce 框架下的 K-means 聚类算法进行改进,提出了一种基于 MapReduce 差分隐私的 K-means 聚类高效传输算法。在保证水电站集控中心高级应用的同时,满足水电站红外监控系统大量数据高效传输、汇聚的要求。改进后的 MapReduce 差分隐私 K-means 聚类算法实现了隐私保护并且提高了聚类的准确率和传输效率。

## 1 红外监控系统

红外监控系统是一种使用红外技术来实现监控与安全的系统。大型水电站常用一些高级红外监控系统实现对水电站动态检测、人员识别和事件报警等<sup>[11-12]</sup>。

在水电站运行过程中,温度、液位、阀位、流量等参数需要进行实时监控,以便出现问题及时处理,从而保障水电站的正常运行,提高效率。然而红外监控系统可以建立起水电站设备的远程监控和数据采集,将产生的图像和数据进行实时的传输和备份,通过远程监控系统,实现对水电站的远程监控。图 1 为红外监控系统架构图。

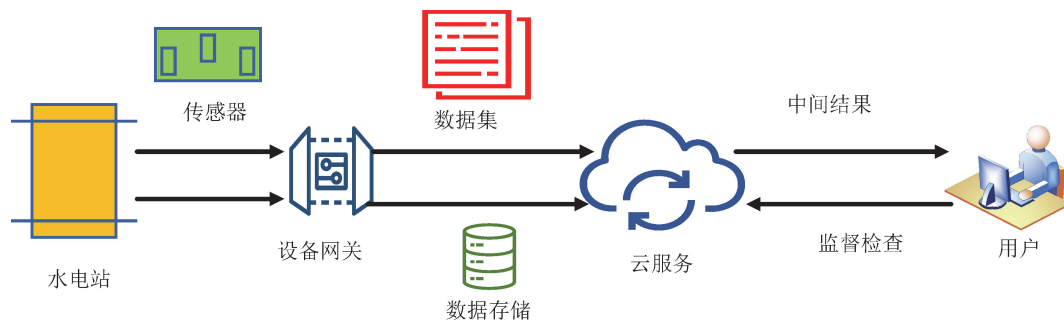


图 1 红外监控系统架构

Fig. 1 Infrared monitoring system architecture

我们以吉林台水电站为例来进行说明,首先将红外传感器和物联网网关安装在水电站的各个关键部位上,实现对温度、液位、阀位、流量诸多数据的采集。数据通过汇聚传输装置进行聚合,并将数据存

储到云端服务器,数据经分析处理后通过云端服务器提供给下位机。工作人员可以通过下位机在监控室通过云端数据平台实时监控水电站运行状态和数据,亦可通过手机等移动终端进行实时监控和数据

查询。对于异常情况,监控系统能及时发出预警,维护人员便可及时通过平台和网关对有关设备进行远程调试并维修,大大缩短了维护周期,提高了运行效率,延长了设备寿命,为水电站的安全稳定运行提供了有效支撑<sup>[13]</sup>。

## 2 隐私保护方法

2006年Dwork在针对统计数据库的隐私泄露时提出了差分隐私这一全新的定义<sup>[14]</sup>。对吉林台流域梯级水电站进行初步数据获取后,在数据传输处理过程中,隐私保护便成了关键。差分隐私是一种保护个人隐私的技术框架,可以使得数据分析时不会泄露个人的私人信息。水电站大量数据进行K-means聚类传输时,迭代过程中会计算各数据点与聚类中心集合的距离大小,这个操作可能导致隐私信息泄露。另外,K-means算法执行完成后得到并且公布的聚类中心点集合若被攻击者联合自身拥有的背景知识可能造成隐私泄露的情况。因此,大型水电站数据汇聚传输过程中加入差分隐私尤为重要。

噪声机制在差分隐私中起着重要的作用,其作用是为原始数据添加一定的随机噪声,从而保护个人隐私,常用的噪声机制包括Laplace机制和指数机制<sup>[15]</sup>。基于MapReduce的差分隐私聚合算法是在K-means数据聚类的基础上引入了Laplace机制,该机制通过向已经聚类的数据中添加随机的Laplace噪声来实现干扰,确保干扰后的结果符合差分隐私的保护要求。

随机噪声服从的Laplace分布,概率密度函

数为:

$$p = \frac{1}{2a} \exp\left(-\frac{|x-\mu|}{a}\right) \quad (1)$$

式中: $p$ 为概率密度函数; $a$ 为尺度参数( $a > 0$ ); $\mu$ 为位置参数; $x \sim \text{Lap}(\mu, a)$ 。而噪声在K-means数据处理之后加入量的多少主要取决于全局敏感度 $\Delta h$ :

$$\Delta h = \max_{D_1, D_2} \|h(D_1) - h(D_2)\|_1 \quad (2)$$

式中: $D$ 为数据集; $D_1, D_2$ 为相邻的两个数据集。而个体隐私泄露的多少,需要通过隐私预算参数 $\epsilon$ 来衡量,其值越小保护程度越高,隐私预算 $\epsilon$ 所在噪声满足条件为:

$$e^{-\epsilon} \leq \frac{P_r[E_{s1}]}{P_r[E_{s2}]} \leq e^{\epsilon} \quad (3)$$

式中: $P_r$ 为随机事件披露的风险; $E_{s1} = t - r$ ,  $E_{s2} = t - r + 1$ ;  $t$ 是对某一操作正常的查询结果; $r$ 为在满足差分隐私保护下的查询结果。

## 3 数据分段传输

监控数据分段传输是将大量监控数据分成小段进行传输,以便高效传递数据。分段传输是在K-means聚合传递时,Laplace加密处理数据过程当中,将数据划分成适当大小的块,并为每个块分配唯一标识符。随后,添加元数据以实现传输控制和管理。然后按照一定顺序和时间间隔进行数据传输,在传输过程中传输链可根据传输的环境对数据进行有损或无损压缩,以保障数据的实时性和完整性,具体分段传输流程见图2。

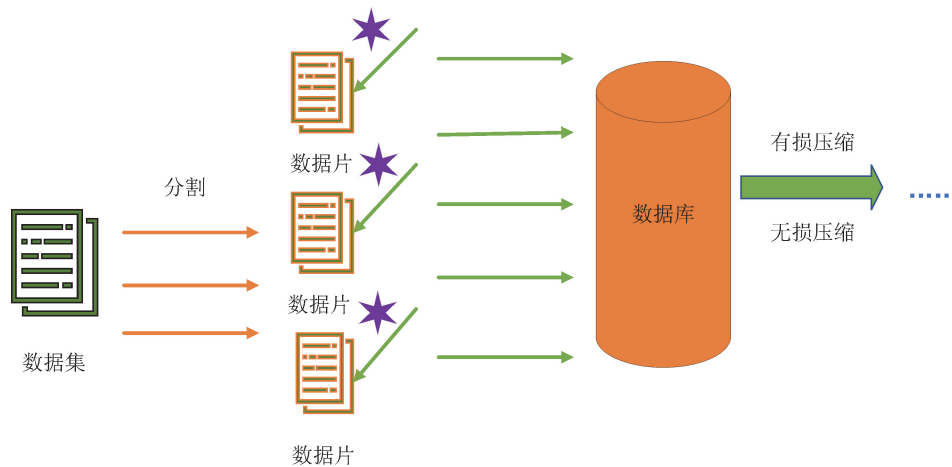


图2 分段传输

Fig. 2 Segmented transmission

水电站红外监控系统高效数据传输时,数据的传输采用分段式,数据被分为多个数据片进行隐私标记,分段传送,减少了数据泄露的风险。数据的分

段式传输结合隐私保护可以减小水电站数据传输的开销和风险。

## 4 聚类算法

### 4.1 基于 Hadoop 的 MapReduce 框架

MapReduce 是源自 Google 在大规模集群上的简化数据处理(Simplified Data Processing on Large Clusters)中提出的方法<sup>[16]</sup>,研究开源后实现,MapReduce 存在于大数据 Hadoop 上的一个编程框

架,MapReduce 工作流程包括 Map 阶段和 Reduce 阶段。在 MapReduce 中,数据集被分割成独立的块,并通过第一阶段(Map)以完全并行的方式进行处理。在 Map 阶段完成后,数据会经过映照和排序,然后输入到第二阶段(Reduce)进行集合计算。通过 Map 和 Reduce 两个阶段的处理,一个大数据集可以被转化为另一个数据集,处理过程见图 3。

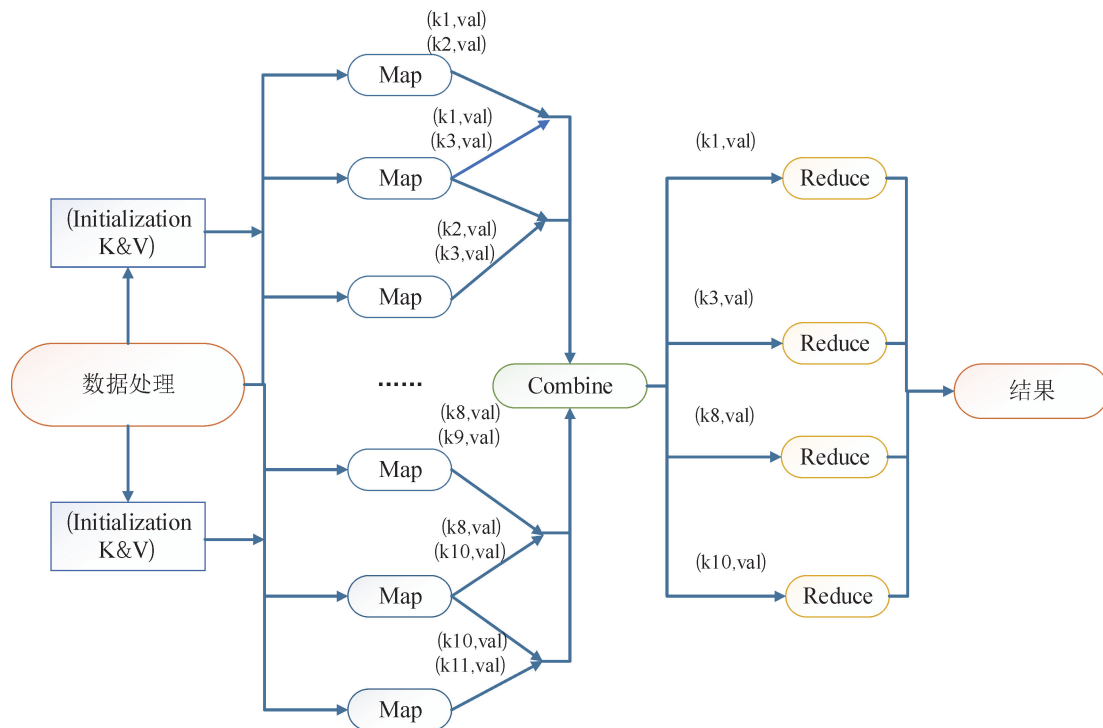


图 3 MapReduce 框架

Fig. 3 MapReduce framework

实际上,MapReduce 框架只对<key, value>形成的键值对处理,此框架将输入的任务进行组建,最后生成一组键值对。MapReduce 能够自动处理数据分布,具有较强容错率,且可以在水电站监控计算机上同时运行,以实现较高的数据处理效率。

Combine 放置 Reduce 之前,可以对数据传输进行优化,提高传输效率,可应用于大规模水电站数据集聚类处理。

### 4.2 数据并行方式

MapReduce 是一种用于处理大规模数据并行计算的编程模型,在差分隐私聚类算法中,重点在于选择聚类的中心点以及计算数据之间的欧式距离,通过这些步骤来获得较为精确的聚类中心点。在处理大量水电站数据集时,差分隐私通常采用数据并行方式,这种方式将监测到的数据集分为多块,分配给同等数量的下游节点。下游节点在更新聚类中心后进行数据块的计算和分类,直到聚类中心的变化小于预设的阈值或者达到预设的迭代上限为止。数据处理时着

重注意的是,数据点到聚类中心点距离的局部归类和对新聚类中心的全局更新的计算,其算法过程见图 4。水电站进行高效数据汇聚时,通过并行计算环境下利用 MapReduce 的差分 K-means 聚类算法解决数据可靠传输问题,降低数据传输的开销,在节点之间进行局部计算和增量更新,以减小通信和计算量。

### 4.3 基于 MapReduce 差分隐私的 K-means 聚类

K-means 聚类算法最早由 Mac Queen 提出,后由 Forgy、Hartigan、Lloyd 对算法做了相关的修正改进<sup>[17]</sup>。K-means 聚类算法是将数据集划分为  $K$  个簇,使得每个数据点都属于最近的簇,每个簇的中心被定义为该簇内所有数据点的平均值,并通过不断地迭代优化来使得每次迭代结果向簇的中心点聚拢,直到满足收敛条件为止。对于采集到的水电站数据总数为  $m$ ,我们将总数据集分成  $K$  个簇,并随机选择  $K$  个数据点作为初始的中心点。然后,将每个数据点分配到最近的簇中心点所属的簇中。通过

迭代计算每个簇内数据点的平均值,并更新簇中心点,不断重复此过程,直到达到预定的迭代次数。最

终,我们可以得到  $K$  个簇以及它们对应的簇中心点,图 5 为二簇聚类过程。

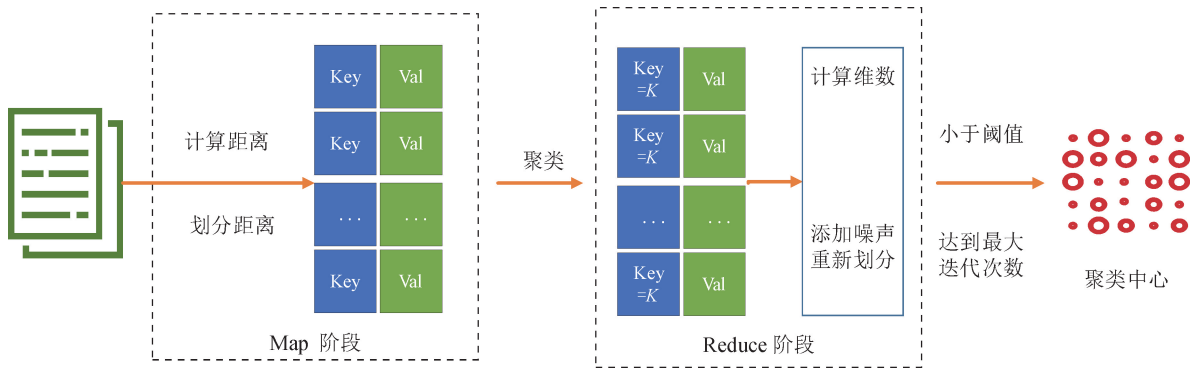


图 4 算法流程

Fig. 4 Process of the algorithm

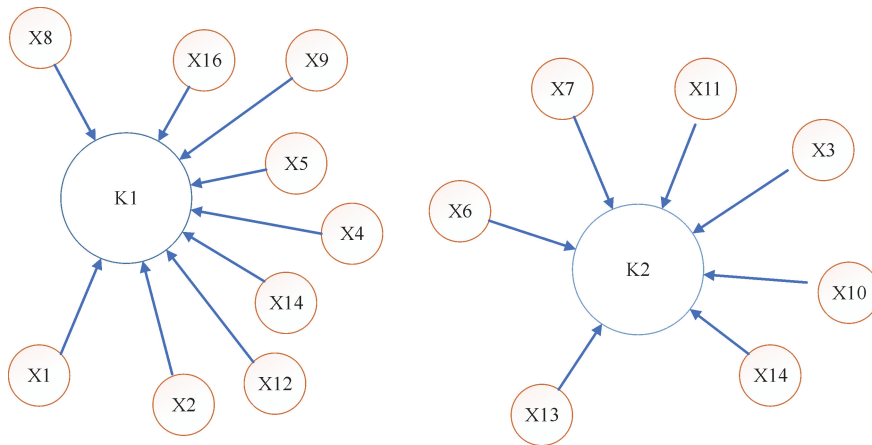


图 5 二簇聚类

Fig. 5 Two-cluster clustering

簇的分配过程中簇的最小聚类距离为  $\delta$ , 样本点至聚类中心欧氏距离分别为  $d$ , 分别用式(4)、(5)表示:

$$\delta = \arg \min_{j=1,2,3,\dots,k} \|x_i - c_j\|^2 \quad (4)$$

$$d = \sqrt{\sum_{k=1}^n (x_{ik} - c_{jk})^2} \quad (5)$$

式中:  $x_{ik}$  为数据样本;  $c_{jk}$  为聚类中心。将数据点分配给聚类之后,需要重新计算质点,作为中心点。重新迭代,需要重新计算新的聚类中心,表示如下:

$$c_j = \frac{1}{|S|} \sum_{i \in S} x_i \quad (6)$$

式中:  $S = \{x_1, x_2, x_3, \dots, x_{|S|}\}$ ;  $S$  是所有的数据点集合;  $c_j$  为新的聚类中心;  $|S|$  为聚类大小表示数据点数量;  $x_i$  代表  $S$  中的样本点,检测中心点是否发生变化。在这个过程中,对于  $K$  的选取十分关键,通常实用手肘法和轮廓系数法,本文主要通过手肘法对  $K$  值选取,用下式表示:

$$S_{se} = \sum_{i=1}^k \sum_{x \in S} \|x_i - c_j\|^2 \quad (7)$$

式中:  $S_{se}$  为误差平方和,该误差平方和可体现聚簇之后的密集程度,  $S_{se}$  与密集程度成反比。基于 MapReduce 差分隐私的 K-means 聚类过程均在 MapReduce 框架中实施。差分隐私 K-means 聚类算法流程图见图 6,图 6 中涉及的各项参数具体含义见表 1。

表 1 参数及含义

Tab. 1 Parameters and meaning

| 符号             | 含义                     |
|----------------|------------------------|
| $\Delta h$     | 噪声敏感度                  |
| $C_k$          | 初始聚类中心点                |
| $S'$           | 归一化数据集                 |
| $\text{Sum}_k$ | 数据记录各维属性总和             |
| $\text{Num}_k$ | 数据数目总和                 |
| $A_k$          | 归一化数据集均分 $K$ 份的第 $k$ 份 |
| $C'_k$         | 添加噪声后新的聚类中心点           |

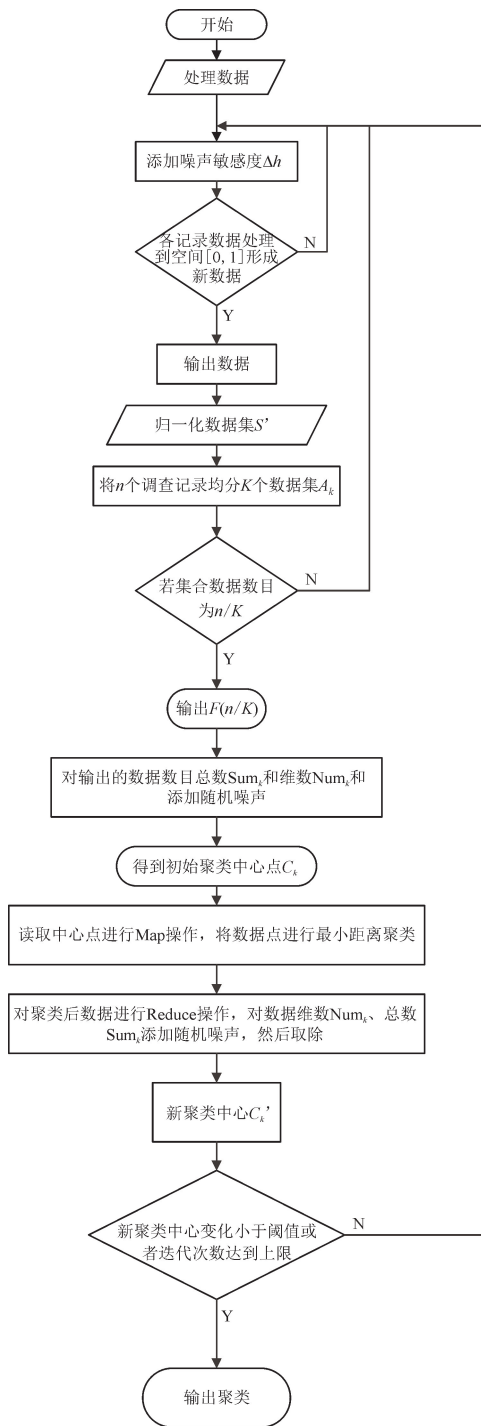


图 6 差分隐私 K-means 聚类算法流程图

Fig. 6 Flowchart of differential privacy K-means clustering algorithms

## 5 仿真与性能分析

在这项工作中,我们介绍了一种基于 MapReduce 差分隐私的 K-means 聚类算法。从传统 K-means 聚类算法的实验结果上得出,算法迭代过程中需要不断调整以计算新的聚类中心点。但是,对大规模数据集处理时效果便不理想,聚类复杂时间

周期较长。本文从传统 K-means 和改进后差分隐私 K-means 的聚类结果对比分析出发,分析改进后的实验仿真性能。在 MATLAB 上执行,采用的数据集来自吉林台流域梯级水电站,包含 5 座大型水电站的温度、液位、阀位、流量等诸多影响因素信息。

### 5.1 K-means 聚类算法

K-means 聚类算法在处理温度、液位和流量数据时,会随机生成 3 个中心以及标准差,之后利用 K-Means 均值聚类算法产生高斯分布数据,利用欧氏距离选择初始聚类质心位置,重复聚类次数。每次均有一个新的初始质心集,计算量、工作量巨大,处理效率较低。我们以水电站某年一季度的数据集进行处理,得到温度、液位、流量三种因素的聚类质心和温泉站的图像分割结果为例进行分析,从  $K=3$  至  $K=8$  进行分割处理,发现  $K=3$  之后的分割效果愈来愈好,但过程繁琐复杂。具体见图 7 和图 8。

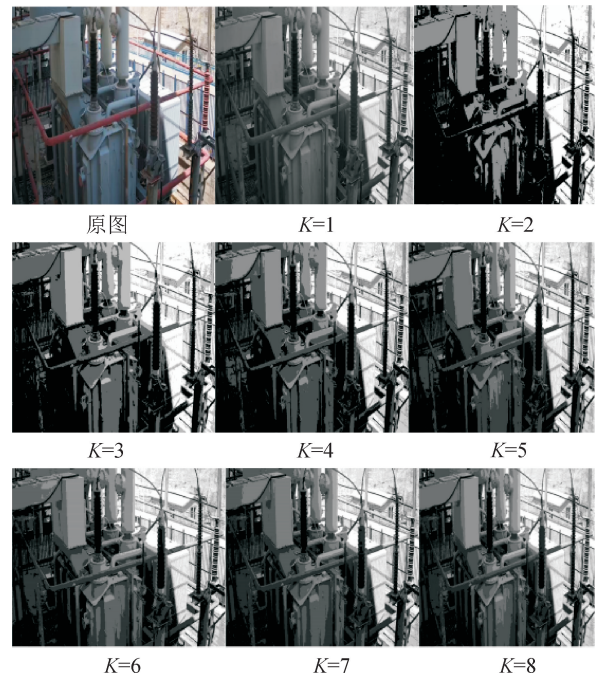


图 7 聚类分割处理图

Fig. 7 Processing graph of clustering segmentation

在实例中,测试到了  $K=i (i=1,2,3,\dots,8)$  的不同值时所对应的聚合程度。随着  $K$  的增大,数据以及图像间聚合程度也不断的增大,  $S_{se}$  大幅度下降。每个样本中心点更逼近于数据中心点,聚合程度提高。但是,在初始设定聚类中心值时很难设定最佳聚类数,而噪声和孤点也存在较大的影响。因此,传统 K-means 聚类算法在确定聚类数、受噪声孤点、处理非凸型数据集、对初始值敏感度以及计算复杂度方面存在较多缺点<sup>[18-19]</sup>。

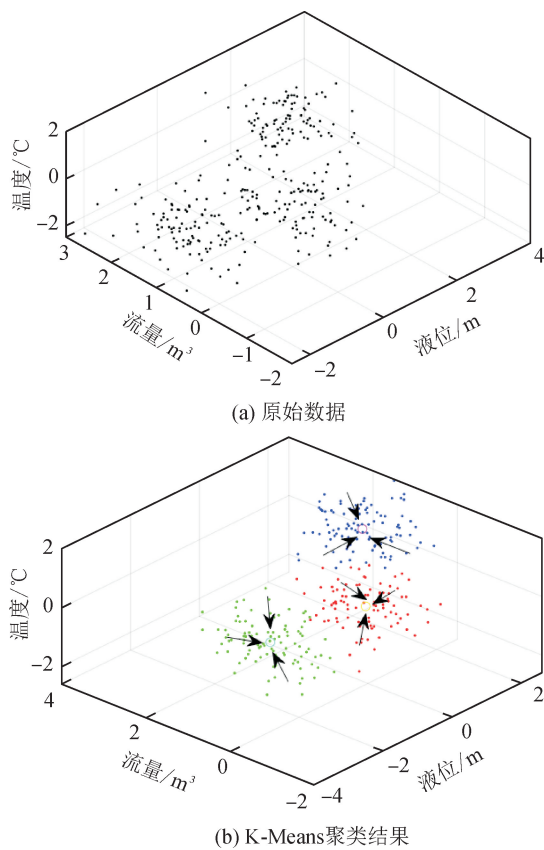


图 8 K-Means 聚类

Fig. 8 K-Means clustering

## 5.2 改进的差分隐私 K-means 聚类

在验证基于 MapReduce 差分隐私 K-means 聚类算法时,利用 F-measure 指标对聚类效果进行评价,而

表 2 改进后算法和传统算法的聚类效果对比

Tab. 2 Comparison of clustering effect by improved algorithm and traditional algorithm

单位: %

| 隐私预算      | 1     | 2     | 3     | 4     | 5     | 6     | 7     | 8     | 9     | 10    |
|-----------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| F-Measure | 0.132 | 0.526 | 0.596 | 0.678 | 0.742 | 0.795 | 0.866 | 0.872 | 0.912 | 0.945 |
| K-Means   | 0.124 | 0.466 | 0.556 | 0.643 | 0.713 | 0.778 | 0.843 | 0.86  | 0.887 | 0.889 |

表 3 同一数据集下相同隐私预算不同 K 值的准确率测算

Tab. 3 Accuracy measurement for different K values of the same privacy budget in the same dataset

| 数据集      | 隐私预算 | K 值/次 | 准确率/% |
|----------|------|-------|-------|
| 梯级水电站数据集 | 5    | 2     | 83.3% |
|          |      | 4     | 81.4% |
|          |      | 8     | 80.1% |
|          | 3    | 2     | 84.3% |
|          |      | 4     | 82.3% |
|          |      | 8     | 81.4% |

图 9 为论文提出方法与传统聚类方法聚类效果的对比,横轴代表隐私预算,评价算法提供隐私保护水平的高低情况,隐私预算根据数据集和算法情况进行设置,数据集规模越大,隐私预算需要越小,以保护

F-measure 与准确率  $P_i$  和召回率  $R_i$  有关。(  $F \leq 1$  )  $P_i$  和  $R_i$  与聚类程度成正比,随着准确率与召回率的增大,聚类程度越高,相似性越高,计算公式为:

$$P_i = \frac{c_i}{|D_i|} \quad (8)$$

$$R_i = \frac{c}{|c_i|} \quad (9)$$

式中:  $c$  为数据集中相同记录的总数;  $n$  为数据集总记录数据;  $C$  和  $D$  分别为数据集,其中  $i$  为数据当中第几个(  $1 \leq i \leq K$  ),随后对准确率  $P_i$  与召回率  $R_i$  加权平均得到  $F_i$ ,即为 F-measure 大小,对各聚类集合  $F_i$  进行加权平均,得到聚类结果指标分别用 (10) 和 (11) 表示:

$$F_i = \frac{2P_iR_i}{R_i + P_i} \quad (10)$$

$$F = \sum_{i=1}^K \frac{|c_i|}{n} F_i \quad (11)$$

本方案是用指标大小评价聚类程度,所得结果值越大其相关度越高,随机噪声值影响也越小。本文根据设置不同的聚类个数  $K$  以及不同隐私预算  $\epsilon$  来衡量差分隐私的 K-means 聚类算法,表 2 是 F-measure 和 K-means 的聚类效果对比结果。很显然, F-measure 在隐私预算相同情况下聚类效果更佳。表 3 所示为同一数据集下,两组隐私设定不同  $K$  值的准确率,根据实验结果分析,改进后的聚类效果更优于传统聚类效果。

个人隐私安全,本实验取值  $\epsilon \in [1, 10]$ , 在保护隐私安全的前提下对吉林台流域梯级水电站进行评价。

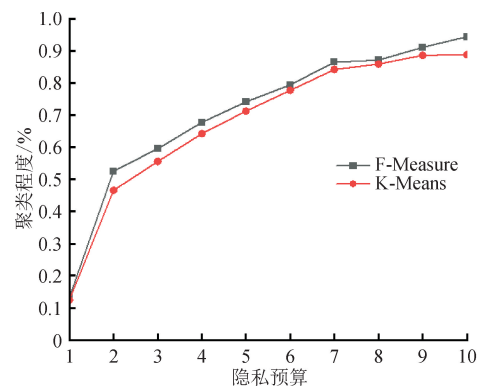


图 9 聚类程度对比

Fig. 9 Comparison of clustering degrees

根据设定不同的隐私预算和  $K$  值评价最终聚类的准确性。为了评价不同隐私预算对实验结果影响,设置两组隐私预算,分别在  $K=2, K=4, K=8$  下评判最终聚类准确性。结果表明, $K$  聚类个数一定时,隐私预算越大,聚类准确率越高,改进算法明显比未改进算法的聚类准确率高。而聚类个数  $K$  不同时,对最终的聚类准确率影响较小。与传统 K-Means 的结果相比,基于 MapReduce 差分隐私的 K-means 高效聚类算法有较高的准确率,提高了水电站的传输效率。

## 6 结 论

在传统 K-Means 聚类算法的基础上,提出了一种新的基于 MapReduce 差分隐私的 K-Means 聚类算法。算法在数据均分后加入了随机噪声,重新得到初始聚类中心,添加随机噪声时对隐私预算和全局敏感度做了相关设置提高了大数据环境下的隐私保护能力,聚类结果可用性也得到了保障。经过理论分析和仿真实验,基于 MapReduce 的差分隐私 K-Means 聚类算法优于传统 K-means 聚类算法且隐私性能也得到了保障,数据的传输性能也有所改善,实现了水电站红外监控系统的高效数据汇聚传输。在以后的工作中,我们将研究如何在隐私预算迭代次数过多的情况下,减小聚类结果和隐私保护级别。

### 参考文献:

- [1] 楼向东,杨兴,孙继康,等.无人值守通信机房远程监控系统的设计与实现[J].电子设计工程,2023,31(12): 119-123.  
LOU Xiangdong, YANG Xing, SUN Jikang, et al. Design and implementation of remote monitoring system for unattended communication machine room[J]. Electronic Design Engineering, 2023, 31(12): 119-123.
- [2] 焦合军,周万春,施进发,等.基于差分隐私保护的机器人视觉图像发布算法[J].应用科学学报,2023,41(4):682-691.  
JIAO Hejun, ZHOU Wanchun, SHI Jinfa, et al. Image publication algorithm of service robot based on differential privacy protection[J]. Journal of Applied Sciences, 2023, 41(4): 682-691.
- [3] 李姣,王秋萍,戴芳.基于改进 HHO 与 K-Medoids 的混合聚类算法[J].西安理工大学学报,2022,38(3): 410-420.  
LI Jiao, WANG Qiuping, DAI Fang. Hybrid clustering algorithm based on improved HHO and K-Medoids[J]. Journal of Xi'an University of Technology, 2022, 38(3): 410-420.
- [4] 张勇,王奎,张松浩,等.某大型水电站智能监控及辅助决策系统应用[J].云南水力发电,2023,39(1): 269-271.  
ZHANG Yong, WANG Kui, ZHANG Songhao, et al. Application of intelligent monitoring and decision support system for a large hydropower station[J]. Yunnan Water Power, 2023, 39(1): 269-271.
- [5] LI Peng, XU Chao, XU He. Data privacy protection algorithm based on redundant slice technology in wireless sensor networks[J]. International Journal of Information Security and Privacy, 2021, 15(1): 190-212.
- [6] 刘永强,杨菊梅,薛邦贵,等.自动气象站工作状态远程监控系统设计与实现[J].气象水文海洋仪器,2023,40(3):131-133,137.  
LIU Yongqiang, YANG Jumei, XUE Banggui, et al. Design and implementation of remote monitoring system for working state of automatic weather station[J]. Meteorological, Hydrological and Marine Instruments, 2023, 40(3): 131-133, 137.
- [7] YANG Mengmeng, HUANG Longxia, TANG Chenghua. K-Means clustering with local distance privacy[J]. Big Data Mining and Analytics, 2023, 6(4): 433-442.
- [8] 王辉,潘俊辉,PETRESCU M,等. Hadoop 下并行化实现文本聚类的优化算法[J].计算机与数字工程,2022,50(12): 2611-2615,2664.  
WANG Hui, PAN Junhui, PETRESCU M, et al. An optimization algorithm for implementing text clustering in parallel under Hadoop[J]. Computer & Digital Engineering, 2022, 50(12): 2611-2615, 2664.
- [9] SHAHRIVARI S, JALILI S. Single-pass and linear-time k-means clustering based on MapReduce[J]. Information Systems, 2016, 60(1): 1-12.
- [10] 高榕,李晶,肖雅夫,等.基于云环境 K-means 聚类的并行算法[J].武汉大学学报(理学版),2015,61(4): 368-374.  
GAO Rong, LI Jing, XIAO Yafu, et al. Parallel algorithm based on K-means clustering in cloud environment[J]. 2015,61(4): 368-374.
- [11] BANAIIE F, MISIC J, MISIC V B, et al. Performance analysis of multithreaded IoT gateway[J]. IEEE Internet of Things Journal, 2019, 6(2): 3143-3155.
- [12] 李云红.红外辐射测温技术及应用研究[M].北京:北京大学出版社,2022.
- [13] FU Wenlong, LIU Xing, CHU Fanwu, et al. A disassembly sequence planning method with improved dis-

- crete grey wolf optimizer for equipment maintenance in hydropower station[J]. The Journal of Supercomputing, 2023, 79(4): 4351-4382.
- [14] YANG Ben, ZHANG Xuetao, LI Zhongheng, et al. Efficient multi-view K-means clustering with multiple anchor graphs[J]. IEEE Transactions on Knowledge and Data Engineering, 2023, 35(7): 6887-6900.
- [15] CROFT W, SACK J-R, SHI WEI. Differential privacy via a truncated and normalized Laplace mechanism[J]. Journal of Computer Science and Technology, 2022, 37(2): 369-388.
- [16] MAO Yimin, GAN Dejin, MWAKAPESA D S, et al. A MapReduce-based K-means clustering algorithm[J]. The Journal of Supercomputing, 2022, 78: 5181-5202.
- [17] LUND Brady, MA Jinxuan. A review of cluster analysis techniques and their uses in library and information science research:  $k$ -means and  $k$ -medoids clustering[J]. Performance Measurement and Metrics, 2021, 22(3): 161-173.
- [18] 赵铮. 大数据差分隐私保护方案设计与实现[D]. 哈尔滨: 哈尔滨工程大学, 2017.
- ZHAO Zheng. Design and implementation of differential privacy scheme for big data[D]. Harbin: Harbin Engineering University, 2017.
- [19] 安建昌, 江俊峰, 徐中原, 等. 光纤传感与红外视频的复合入侵监控系统设计[J]. 红外与激光工程, 2020, 49(5): 170-176.
- AN Jianchang, JIANG Junfeng, XU Zhongyuan, et al. Design of composite intrusion detection system based on optical fiber sensor and infrared video[J]. Infrared and Laser Engineering, 2020, 49(5): 170-176.
- (责任编辑 王绪迪)

(上接第 323 页)

- [21] 严慈, 侯兰功. 基于灰色理论的陕西省土地利用变化及其碳排放研究[J]. 西安理工大学学报, 2021, 37(1): 25-31.
- YAN Ci, HOU Langong. Study on land use change and carbon emission in Shaanxi Province based on grey theory[J]. Journal of Xi'an University of Technology, 2021, 37(1): 25-31.
- [22] 曹周亮, 张欣蓉, 员学锋, 等. 基于土地利用变化的陕西省县域碳排放时空变化及影响因素研究[J]. 水土保持通报, 2022, 42(5): 376-385.
- CAO Zhouliang, ZHANG Xinrong, YUAN Xuefeng, et al. Spatio-temporal variation and influencing factors of CO<sub>2</sub> emission at county scale in Shaanxi Province based on land use change[J]. Bulletin of Soil and Water Conservation, 2022, 42(5): 376-385.
- [23] 卢奕帆, 梁颖然, 卢思言, 等. 结合“珞珈一号”夜间灯光与城市功能分区的广州市碳排放空间分布模拟及其影响因素分析[J]. 地球信息科学学报, 2022, 24(6): 1176-1188.
- LU Yifan, LIANG Yingran, LU Siyan, et al. Spatialization of carbon emissions in Guangzhou City by combining LuoJia1-01 nighttime light and urban functional zoning data[J]. Journal of Geo-information Science, 2022, 24(6): 1176-1188.
- [24] 李秋妍, 曹翠. 固定资产投资对二氧化碳排放影响的实证研究——基于省际面板数据模型[J]. 科技管理研究, 2014, 34(14): 221-225.
- LI Qiuyan, CAO Cui. An empirical study on the impact of fixed asset investment on carbon dioxide emission-based on inter-provincial panel data model[J]. Science and Technology Management Research, 2014, 34(14): 221-225.
- (责任编辑 周 蓓)