

基于Spark的输变电路实时故障监测研究

陈建峡¹, 朱季骐¹, 张月¹, 张晓星², 吕俊涛³, 白德盟³

CHEN Jianxia¹, ZHU Jiqi¹, ZHANG Yue¹, ZHANG Xiaoxing², LV Juntao³, BAI Demeng³

1. 湖北工业大学 计算机学院, 武汉 430068

2. 武汉大学 电气工程学院, 武汉 430072

3. 国网山东省电力公司 电力科学研究院, 济南 250002

1. School of Computer Science, Hubei University of Technology, Wuhan 430068, China

2. School of Electrical Engineering, Wuhan University, Wuhan 430072, China

3. State Grid Shandong Electric Power Research Institute, Jinan 250002, China

CHEN Jianxia, ZHU Jiqi, ZHANG Yue, et al. Real-time fault monitoring of transmission lines based on Spark. *Computer Engineering and Applications*, 2018, 54(5): 265-270.

Abstract: Since the monitoring data of transmission lines are the largest part of the amount of data in the smart grid, including not only the online condition monitoring data, but also the basic information of the devices, the experimental data, defect records, it requires a higher performance of the reliability and real-time in the data processing. The paper designs and realizes a novel model to solve the real-time fault monitoring of transmission lines according to the practical application of power transmission line faults' types. In particular, the paper constructs a distributed cluster based on Spark, an efficient real-time data processing system, for the transmission line fault real-time monitoring, develops a distributed ISODATA and fuzzy KNN big data analysis algorithm. Compared with standalone KNN algorithm, it improves 70.75% efficiency of the time performance. Experimental results demonstrate the proposed approach has the obvious advantages of the computational efficiency.

Key words: real-time big data; transmission lines; fault monitoring; distributed Iterative Self Organizing Data Analysis Techniques Algorithm (ISODATA) algorithm; distributed fuzzy k -Nearest Neighbor (KNN) algorithm

摘 要: 输变电路状态监测数据是智能电网中数据量很大的一部分, 不仅包括在线的状态监测数据, 还包括设备的基本信息、实验数据、缺陷记录等, 在数据处理的可靠性和实时性方面的要求都很高。根据实际应用中输变电路的故障类型, 设计并实现了输变电路实时数据故障监测模型。其中, 利用高效处理实时数据的Spark系统, 研发出基于Spark的分布式ISODATA和模糊KNN大数据分析算法, 与单机KNN算法相比, 在时间性能上提高了70.75%效率, 具有明显的计算效率优势。

关键词: 实时大数据; 输变电路; 故障监测; 分布式迭代自组织数据分析算法 (ISODATA); 分布式模糊 k 最近邻分类算法 (KNN)

文献标志码: A 中图分类号: TM715 doi:10.3778/j.issn.1002-8331.1609-0025

1 引言

随着全球能源问题日益严峻, 世界各国都在积极开展智能电网方面的研究工作。其中, 智能电网运行中实

时数据的获取、传输和存储, 以及海量数据的实时分析, 是支撑安全可靠的智能电网的必要基础^[1]。目前, 随着智能技术在电网中的广泛应用, 电网运行过程中会产生

基金项目: 国家高技术研究发展计划(863)(No.2015AA050204)。

作者简介: 陈建峡(1971—), 女, 副教授, 研究领域为数据挖掘、机器学习、搜索引擎, E-mail: cjsxwh@126.com; 朱季骐(1994—), 男, 研究领域为数据挖掘; 张月(1990—), 女, 研究生, 研究领域为数据挖掘; 张晓星(1972—), 男, 教授, 博士生导师, 研究领域为输变电装备在线监测与故障智能诊断技术。

收稿日期: 2016-09-02 **修回日期:** 2017-04-24 **文章编号:** 1002-8331(2018)05-0265-06

CNKI网络优先出版: 2017-07-19, <http://kns.cnki.net/kcms/detail/11.2127.TP.20170719.1127.044.html>

各种海量异构、多态的数据,也即大数据。其中,数据量最大的应属于输变电路状态监测数据。状态监测数据不仅包括在线的状态监测数据(时序数据和视频),还包括设备的基本信息、实验数据、缺陷记录等,数据量极大,在可靠性和实时性方面的要求都很高^[2]。如何利用这些海量数据实施大数据分析,对输变电路状态进行故障的实时监测与预测,成为制约国内外电网管理水平发展的瓶颈,也是当前智能电网中重要的研究课题。

1.1 实时大数据分析的研究现状

根据处理时间的不同,大数据的处理方式主要包括三种形式:对静态数据的批量处理;对在线数据的实时处理;对图数据的综合处理^[2]。2010年,Google率先推出了Dremel实时数据处理系统,开创了大数据实时处理技术的先锋,弥补了以Hadoop为代表的传统大数据批处理方式中,计算效率低下的不足之处。

实时数据处理可以分为交互式数据处理和流式数据处理两种方式。流式数据处理用于服务器日志的实时数据采集,其处理系统的典型代表有Twitter的Storm,流式处理作业被分发至不同类型的组件,每个组件负责一项简单的、特定的处理任务^[3]。交互式数据处理则是将PB级数据的处理时间缩短到秒级,典型代表系统是Berkeley的Spark系统和Google的Dremel系统。传统的Map Reduce在处理大量网络传输时,因为磁盘I/O使得数据处理效率低效,与它不同的是,Spark是一个基于内存计算、可扩展、开源的集群计算系统,可以使用内存进行数据计算以便快速处理查询,实时返回分析结果^[4]。

1.2 输电线路故障监测的研究现状

在整个电力工业系统中,输变电路的状态监测对数据存储与处理平台的性能或实时性具有较高的要求,云计算技术虽然可以有效地处理大数据,但很有必要改善对海量监测数据的存取性能,以满足实时性的要求。以往的大规模停电事故,比如大风导致的线路跳闸等。现有SCADA(Supervisory Control And Data Acquisition,即数据采集与监视控制)系统的监控范围仅限于系统的主参数,对构成系统的各重要设备的具体信息无法及时提供,致使运行人员在事故面前难以做出正确的处理^[5]。未来智能电网要求具有故障自愈功能,其SCADA系统须拥有全网的监测数据,这对平台的实时处理提出了更高的要求^[6]。

在国外,文献[7]已经通过变电站智能电子设备(IED)获取相关记录,结合传统的(历史)设备管理数据,提出了对电力设备进行评估和预测的详细系统架构和具体实施方案。特别是当输电线路出现故障时,能够准确及时地确定电力系统中故障发生的位置,对故障进行抢修。然而,文献[8]指出,随着智能电网的发展,已有的电网调度和控制模型不能处理大量的智能化小型

发电系统产生的波动和不可预知的行为,需要创建一种新型的电网监测系统,能够更加细粒度地跟踪电网的实时状态。因此,国外关于实时大数据分析在智能电网中的应用也处于起步状态。

目前,国内也开展了输变电路故障实时处理的研究。文献[9]认为状态监测数据是一种典型的流式数据,具有持续不断、数据量大、规模及顺序无法预知等特点,并提出了复杂装备监测实时流数据处理框架。文献[10]综述了当前处理流数据的主流技术和发展趋势,数据流式计算系统应具有数据吞吐量高、状态一致性强、负载均衡稳定、系统容错性强和可伸缩性强等特点。文献[11-12]论述了流数据处理技术在电力系统自动化中的应用。然而,具有代表性的Storm或者Spark等高效处理实时数据的系统,在互联网企业应用广泛,但是在电力系统中鲜有应用^[7]。

众所周知,以Hadoop为代表的海量数据离线处理方式,需要频繁的磁盘I/O操作,计算效率低下,很难满足电力系统中在线状态监测、评估的需求。另外,Hadoop技术在处理需求多次迭代算法的问题中存在效率低的问题,而输变电路的状态评估算法大多需要多次迭代。

Spark是UC Berkeley AMP lab所开源的类Hadoop MapReduce的通用的并行计算框架,不仅拥有Hadoop MapReduce所具有的优点,Spark的计算中间结果不需要反复读写磁盘,而是暂时存储在内存当中,所以,基于内存的计算平台Spark在处理此类问题的效率比Hadoop高好几倍甚至更高,具有支持复杂查询和实时的流处理等优点^[13]。因此,用Spark技术分析输变电路的故障监测数据,可以良好地满足实时监测的需要。

2 输变电路故障智能检测技术简介

输变电路不对称故障信息量大而且基本规律稳定,十分便于数据挖掘技术处理^[14]。目前,输变电路故障监测领域常用的人工智能技术包括专家系统ES、决策树DT、神经网络NN、支持向量机SVM、模糊理论等^[15]。其中,虽然基于规则的专家诊断系统得以广泛应用,但由于其规则库还存在搜索、更新和扩充较慢的问题,实际效果良好的专家系统并不多^[16];决策树无法删除输变电路中的带噪声的不相关属性,影响故障监测的准确性^[17-18];神经网络的训练时间很长而且参数很多,主要需要人为的经验去确定^[19-20];支持向量机是针对有限样本的情况,它所求的是现有信息下的最优解,无法反映样本数量很大时的最优解^[21-22]。

K最近邻分类算法(K-Nearest Neighbor, KNN)作为一种经典的统计模式识别方法,也是效果最好的分类方法之一,而且KNN方法不是靠判别类别界限的方法来判定所属类别,而是主要靠周围有限的邻近样本,因此对于类别界限交叉或重叠较多的数据来说,KNN方

法较其他方法更为适合^[23]。但是,由于 KNN 只计算最近邻样本,对于类别界限不明显的输变电路数据特点,其预测结果存在偏差。而模糊理论可以弥补这种不足。

模糊理论主要利用隶属度函数从数据中进行知识发现并揭示其潜在的规律,将数据样本模糊化,然后根据模糊规则推断输入样本的故障类型^[24]。采用的模糊化方法一般根据预先选取的隶属度函数将输入样本转化为相应的隶属度。模糊逻辑类似于人工神经网络,是一种重要的非线性映射模型^[25],其独特之处在于能够更加充分有效地利用多种不同类别的数据。并且,模糊理论解决了输变电路中的模糊现象,从而对系统进行综合评价。

因此,根据上述算法在输变电路故障监测中的问题,及实际应用中存在的不足之处,本文根据 KNN 算法的特点,结合模糊理论,提出了模糊 KNN 分类算法,对输变电路故障进行预测分析,解决了 KNN 在处理类别界限不明显数据时的问题。

3 基于 Spark 的输变电故障检测模型

3.1 输变电路实时故障监测模型

本文利用模糊理论和基于实例的算法,建立了输变电路实时故障监测模型。在训练阶段,本文对已知故障类别标签的样本数据,运用研发的分布式 ISODATA 聚类算法进行数据训练。数据测试阶段,根据模糊理论将训练后的数据和测试数据进行模糊化,然后把新数据与样本数据进行比较。通过研发的分布式模糊 KNN 算法计算测试数据的 K 个最近邻集群,计算测试数据的隶属度来确定测试数据的故障类别,并进行线路故障定位分析。本文研发的基于 Spark 的输变电路故障监测模型如图 1 所示。

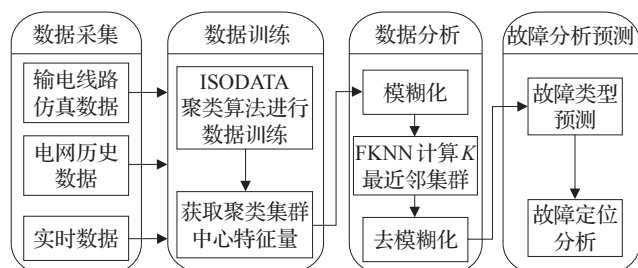


图1 基于 Spark 的输变电路实时故障监测模型

如图 1 所示,输电故障分析及监测模型分为数据采集、数据训练、数据分析及故障预测等四个程序模块。其中,数据采集模块包含了电网历史数据和实时数据,及输变电路仿真结果的三相电压和三相电流信号模型。从三相电压和电流信号采集的输入数据矢量被映射到包含了故障类型的集群。由于实际情况故障案例数不多,所需要的故障和非故障情形下的部分数据需要通过仿真软件获得,与输变电系统数据库记录中获得的

大量数据一起被用于数据训练,然后与来自真实系统获得的数据原型结合,通过模糊 KNN 分类器来进行在线故障分析。

3.2 基于 Spark 的分布式 ISODATA 算法

3.2.1 ISODATA 算法

ISODATA (Iterative Self Organizing Data Analysis Techniques Algorithm) 算法全称为迭代自组织数据分析方法,是一种常用的非监督学习的聚类分析方法。它是在 K -means 算法的基础上,对聚类结果增加“分裂”和“合并”操作,同时对算法中运行控制参数进行设定的聚类算法。根据初始设定的聚类中心,以及设定的最小聚类样本数等参数迭代,最终得到一个比较符合预期的结果。其算法流程图如图 2 所示。

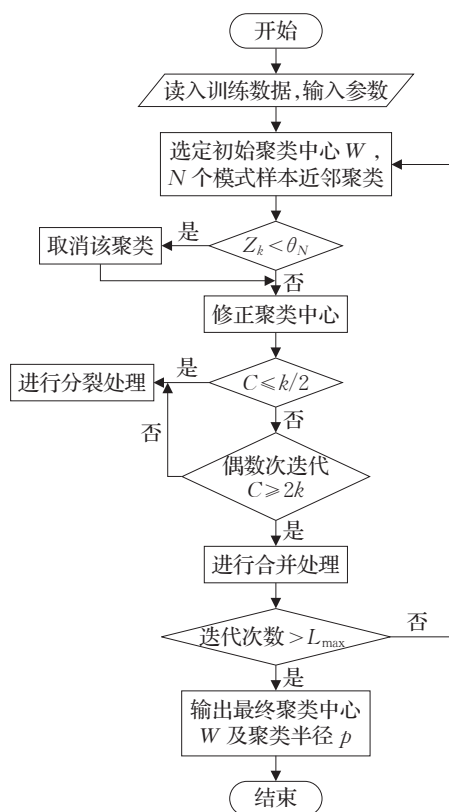


图2 ISODATA 算法流程图

3.2.2 基于 Spark 的分布式 ISODATA 算法

本文结合 Spark 的 RDD 模型和其 API 对算法进行综合设计,考虑到算法中各个步骤的计算量,算法的设计主要分为以下两步:

步骤 1 样本分配,过滤,更新聚类中心,计算样本到中心距离和,计算样本各维方差和。

步骤 2 收集步骤 1 的计算结果,根据条件判断是否执行分裂、合并步骤。

分布式 ISODATA 算法流程图如图 3 所示。

3.3 基于 Spark 的分布式模糊 KNN 算法

3.3.1 模糊 KNN 算法

根据输变电路数据的特点和模糊集合变换原理,

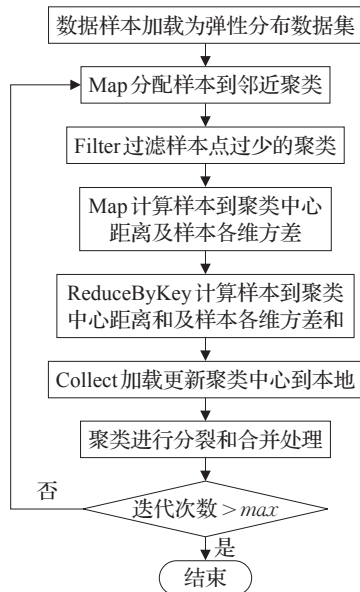


图3 分布式 ISODATA 算法流程图

文中利用隶属度函数来描述各个因素及因子的模糊界线,对输变电路测试数据进行隶属度函数复合运算来预测和分析输变电路故障类型^[26],有效解决了类别界限不明显的问题。

本文研发的模糊 KNN 算法,是根据待分类的样本 X 和训练阶段得到的聚类中心 W 之间的欧式距离,选择与待分类样本距离最小的 k 个样本,并且给模式样本的近邻赋予很大的权值,最后根据模式样本 x_i 的隶属度函数来判定 x_i 的类别。其判定模式样本 x_i 隶属于类别 c 的隶属度函数 $u_c(x_i)$,具体表示如表达式(1)所示:

$$u_c(x_i) = \frac{\sum_{k=1}^K \frac{u_c(w_k)}{\|x_i - w_k\|^2}}{\sum_{k=1}^K \frac{\rho_k}{\|x_i - w_k\|^2}} \quad (1)$$

其中,聚类中心 w_k 的隶属度函数 $u_c(w_k)$ 的取值等于聚类中心的半径 ρ_k ,聚类中心的类别标签是聚类 z_k 中所有样本的类别标签 c 。聚类中心 w_k 的隶属度函数如表达式(2)所示:

$$u_c(w_k) = \begin{cases} \rho_k \\ 0 \end{cases} \quad (2)$$

同一测试模式样本的隶属度函数式(1)中的分母部分,都为该测试模式样本 k 个 $\rho \frac{1}{\|x_i - w_k\|^2}$ 求和,即:最近邻聚类半径 ρ 及聚类中心与测试模式样本之间距离倒数的乘积,求得模式样本 x_i 的最大隶属度 $u_c(w_k)$,则:测试模式样本的类别标签就为 c 。模式样本 x_i 的隶属度函数如表达式(3)所示:

$$u_c(x_i) = \max_c u_c(x_i) \quad (3)$$

模糊 KNN 算法的伪代码如图 4 所示。模糊 KNN 算法与 KNN 算法的最大区别就是,模糊 KNN 对训练后的聚类中心特征数据 W 进行模糊化处理,根据第二部分

给出的聚类中心隶属度函数,即公式(2),计算聚类中心 W 隶属于某一类别 c 的隶属度。然后将该聚类中心的隶属度和测试样本 X_i 之间的距离存入大小为 k 的优先级队列。在遍历完训练集群中心 W ,求得测试数据 X_i 的 k 个最近邻聚类中心后,根据测试数据的隶属度函数 $u_c(x_i)$,即公式(1),计算测试数据 X_i 隶属于类别标签 c 的隶属度,测试数据 X_i 的类别标签为隶属度最大的聚类集群的类别标签。训练数据聚类中心的抽取,不仅提高了模糊 KNN 的分类效率,而且削弱了样本分布不均与对分类的影响,提高了分类精度。

```

1: Input: an unlabeled test set(X,Y)
2: For i=1,2,...,n do
3:   getDistance(test,clusterCenter);
4: end for
5: For j=1,2,...,k do
6:   getTopMinDistance;
7:   saveClusterCenter;
8:   getMaxMembership(test,clusterCenter);
9:   testLabel=ClusterCenterLabel;
10: end for
11: Output: accuracy
  
```

图4 模糊 KNN 算法伪代码

3.3.2 基于 Spark 的模糊分布式 KNN 算法

本文结合 Spark 的 RDD 模型和其 API 对算法进行综合设计。由于此处模糊 KNN 算法使用的训练集合来自于 ISODATA 算法得到的聚类中心点,数据量不大,所以训练数据集采用广播变量的方式发放到各节点。然后利用 RDD 模型的算子进行分类计算,分布式模糊 KNN 算法伪代码如图 5 所示。

```

1: Input: an unlabeled test set RDD<Point>;
2: Broadcast clusterCenter;
3: Map Point to DataNode;
4: For i=1,2,...,n do
5:   getDistance(test,clusterCenter);
6: end for
7: For j=1,2,...,k do
8:   getTopMinDistance;
9:   saveClusterCenter;
10:  getMaxMembership(test,clusterCenter);
11:  testLabel=ClusterCenterLabel;
12: end for
13: Output: Cost Time
  
```

图5 分布式模糊 KNN 算法伪代码

4 实验结果及分析

4.1 实验环境和条件

实验构建了 20 台计算机组成的 Spark 分布式集群,单机都是 2.70 GHz 的 CPU 和 8 GB 内存的 Marcos 系统配置。其中,单机实验用 Java 语言,分布式实验采用 Scala 语言开发程序。

因为在实际工作环境中,线路发生故障的情况并不是很多,文中所需的实验数据是对输电系统进行仿真获取。实验采用了 MATLAB/Simulink 环境下的 PSB 模型

库。PSB(Power System Blocker)是 MATLAB 软件中的电力系统模块,它经常用于电气工程专业的电路、电机系统、电力传输等过程的仿真。本文利用它提供的一种类似电路建模的方式进行电路模型绘制,然后转换成状态方程描述的系统形式,最后在 Simulink 环境下进行仿真分析^[27]。

其中,每条线路的数据样本包括不同距离参数,从线路故障位置的 10%处到 90%处,以 10%递增;不同过渡电阻值 1 Ω、3 Ω、9 Ω、15 Ω;在发生 10 种不同故障类型时三相电压和三相电流的突变量参数;共 252 条训练数据以验证分布式 ISODATA 算法与单机版 ISODATA 算法结果的一致性;使用 3 600 000 条训练数据迭代 100 次测试分布式 ISODATA 算法的运行效率与单机版 ISODATA 算法的运行效率。

文中主要对 10 种输变电路不对称故障类型进行预测分析,包括单相短路接地故障(AG,BG,CG)、两相短路故障(AB,AC,BC)、两相短路接地故障(ABG,ACG,BCG)和三相短路故障(ABC)。参照电力系统故障分析原理,实验中选取故障后与故障前三相电压及三相电流突变量^[28],过渡电阻和故障点距离线路的位置,从线路总长度的 10%到 90%处,以 10%递增。仿真数据序号及含义见表 1 所示。

表 1 仿真数据序号与含义

仿真数据序号	仿真数据含义
1	线路总长度
2	故障距离
3	过渡电阻
4	A 相电压突变量
5	B 相电压突变量
6	C 相电压突变量
7	A 相电流突变量
8	B 相电流突变量
9	C 相电流突变量

实验中数据样本为 9 条不同类型的输变电路的不同位置,及 4 种不同过渡电阻值 1 Ω、3 Ω、9 Ω、15 Ω,共 252×9=2 268 条数据;测试样本是从总样本中随机抽取去标签的 1 800 条数据进行故障预测,并与 KNN 算法预测进行对比。分布式模糊 KNN 则使用 400 万条数据进行效率测试与单机版模糊 KNN 算法进行效率对比。

4.2 分布式 ISODATA 算法实验结果与分析

4.2.1 数据训练结果分析

在少量训练数据情况下,Java 单机实现的 ISODATA 聚类算法和 Scala 并行实现的 ISODATA 算法运行结果一致。为了方便起见,本文通过降维,画出输变电路故障数据训练后的聚类效果如图 6 所示,图中的纵轴代表了不同的故障类型,横轴代表集群的样本个数^[29]。

4.2.2 数据训练时间

在 3 600 000 条训练数据迭代 100 次的条件下,单机

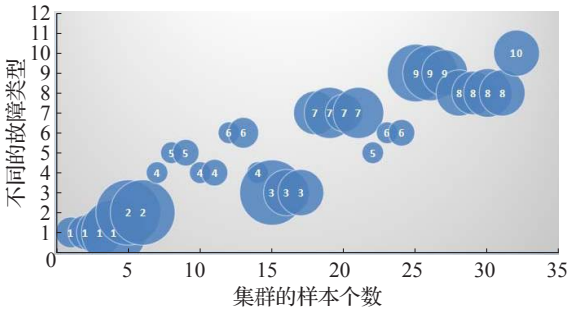


图 6 ISODATA 算法故障聚类效果

版 ISODATA 训练的耗时 211 283 ms,分布式 ISODATA 训练的耗时 110 823 ms。由此可以看出,基于 Spark 的分布式 ISODATA 聚类算法实现,在大数据情况下计算效率要高于单机版的 ISODATA 算法 90%。

4.3 分布式模糊 KNN 算法实验结果与分析

KNN 算法和模糊 KNN 算法故障预测结果如图 7、8 所示,其中横轴代表的是每条线路的故障数据。分布式模糊 KNN 算法的结果与单机版一致。在 4 000 000 样本, K 取值为 6 的条件下分别运行单机版模糊 KNN 算法和分布式模糊 KNN 算法,单机版模糊 KNN 算法运行耗时为 132 839 ms,分布式模糊 KNN 算法的运行耗时 86 512 ms,由此可看出基于 Spark 的模糊 KNN 算法实现效率比单机实现版本的计算效率提升 53.5%。

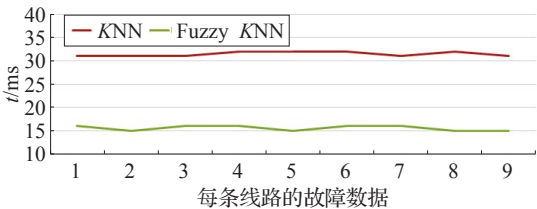


图 7 KNN 与模糊 KNN 故障预测时间

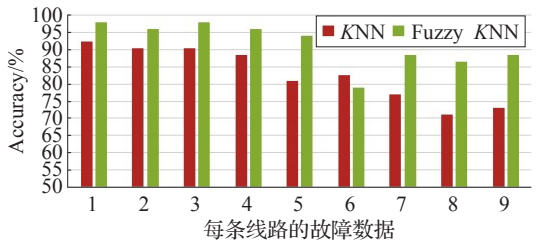


图 8 KNN 与模糊 KNN 故障预测准确率

图 7 和图 8 的横轴是 9 条不同类型的输变电路数据,纵轴分别是测试时间和故障预测准确率。从图 7 和图 8 中可以看出,在输变电路不对称故障预测的应用中,该算法与普通 KNN 算法相较而言,不仅有更高的准确率,而且耗时少。并且,根据故障距离参数,对故障发生的位置有一定的指导性意义。

5 结束语

本文研究了输变电路故障的实时处理技术,基于 Scala 语言和 Spark 框架,实现了 Spark 系统中 MLlib 没

有的 ISODATA 聚类算法和模糊 KNN 分类算法,并在大数据环境下与相应的单机算法进行比较。本文研发的分布式 ISODATA 聚类算法的时间效率提高了 90%,模糊 KNN 分类算法的时间效率提高了 53.5%,在处理大数据时本文研究的分布式算法比单机版的算法的效率优势提高了 70.75%。最终,利用基于 Spark 的 ISODATA 聚类算法和模糊 KNN 分类算法构建了输变电路实时故障监测模型,用于输变电路的实时故障监测工作中。

在今后工作中,将进一步改进分布式 ISODATA 聚类算法和模糊 KNN 分类算法的效率,例如,与 MLlib 模块中已有的机器学习算法的结合、ISODATA 动态 partition 分配等方面,并且用于智能电网中的其他电力设备,例如断路设备和变压器设备等,进行有效的监测分析。同时,本文将对实验中基于 Yarn 的 Spark 集群进行实际的线性扩展,以便直接应用于真实的生产环境中。

参考文献:

- [1] Fang Xi, Misra S, Xue Guoliang, et al. Smart grid, the new and improved power grid: A survey[J]. IEEE Communications Surveys and Tutorials (COMST), 2012, 14(4): 944-980.
- [2] 程学旗, 靳小龙, 王元卓, 等. 大数据系统和分析技术综述[J]. 软件学报, 2014, 25(9): 1889-1908.
- [3] The Apache Software Foundation. Storm: Distributed and fault-tolerant realtime computation[EB/OL]. (2014-07-17) [2016-08-15]. <http://storm.incubator.apache.org/>.
- [4] Zaharia M, Chowdhury M, Franklin M, et al. Spark: Cluster computing with working sets[C]//Proceedings of the 2nd USENIX Conference on Hot Topics in Cloud Computing, 2010.
- [5] Wikipedia. Northeast blackout of 2003[EB/OL]. (2003) [2016-08-15]. http://en.wikipedia.org/wiki/Northeast_Blackout_of_2003.
- [6] 宋亚奇. 智能电网大数据处理技术现状与挑战[J]. 电网技术, 2013, 37(4).
- [7] Kezunovic M. Intelligent design[J]. Power & Energy Magazine IEEE, 2010, 8(6): 37-44.
- [8] Birman K P, Ganesh L, van Renesse R. Running smart grid control software on cloud computing architectures[C]//Workshop on Computational Needs for the Next Generation Electric Grid, 2011: 1-28.
- [9] 庄雪吟, 张力, 翁晓奇, 等. 复杂装备状态监测实时流数据处理框架[J]. 计算机集成制造系统, 2013, 19(12): 2929-2939.
- [10] 孙大为, 张广艳, 郑伟民. 大数据流式计算: 关键技术及系统实例[J]. 软件学报, 2014, 25(4): 839-862.
- [11] 赵加奎, 杨国凤, 沐连顺, 等. 数据流技术在电网自动化中的应用研究[J]. 电网技术, 2011, 35(8): 6-11.
- [12] 王阳光, 尹项根, 游大海. 无线传感器网络应用于智能电网的探讨[J]. 电网技术, 2010, 34(5): 7-11.
- [13] 周国亮, 朱永利, 王桂兰, 等. 实时大数据处理技术在状态监测中的应用[J]. 电工技术学报, 2014(S1): 432-437.
- [14] 叶昊, 王桂增, 方崇智. 小波变换在故障检测中的应用[J]. 自动化学报, 1997, 23(6): 736-741.
- [15] 占才亮. 人工智能技术在电力系统故障监测中的应用[J]. 广东电力, 2011, 24(9): 87-92.
- [16] 马刚. 输变电设备在线状态分析与智能诊断系统的研究[D]. 北京: 华北电力大学, 2013.
- [17] 李再华, 刘明昆. 电力系统故障的智能诊断综述[J]. 电气技术, 2010(8): 21-24.
- [18] 冯少荣. 决策树算法的研究与改进[J]. 厦门大学学报: 自然科学版, 2007, 46(4): 496-500.
- [19] 白建社, 樊波, 黄文华, 等. 基于决策树的变电站故障监测知识表示与获取[J]. 电力系统及其自动化学报, 2004, 16(2): 5-8.
- [20] 尹丽子. 基于动态模型的神经网络稳定性研究[D]. 济南: 山东师范大学, 2011.
- [21] 严凤, 许海梅. 基于神经网络的配电线路综合故障定位方法[J]. 电力系统及其自动化学报, 2015, 27(5): 86-91.
- [22] 丁世飞, 齐丙娟, 谭红艳. 支持向量机理论与算法研究综述[J]. 电子科技大学学报, 2011, 40(1): 2-10.
- [23] 马新明, 韩占忠, 刘平. 基于支持向量机的输电线路故障诊断分类[J]. 电气开关, 2009, 47(5): 61-63.
- [24] 耿丽娟, 李星毅. 用于大数据分类的 KNN 算法研究[J]. 计算机应用研究, 2014, 31(5): 1342-1344.
- [25] 郑圆圆, 陈再良. 模糊理论的应用与研究[J]. 苏州大学学报: 工科版, 2011, 31(1): 52-58.
- [26] 张勇. 基于模糊理论的智能电网运行风险评估系统研究[J]. 电力与能源, 2015, 36(3): 283-286.
- [27] Vasilic S, Kezunovic M. Fuzzy ART neural network algorithm for classifying the power system faults[J]. IEEE Transactions on Power Delivery, 2005, 20(2): 1306-1314.
- [28] 于群. MATLAB/Simulink 电力系统建模与仿真[M]. 北京: 机械工业出版社, 2011.
- [29] Zhang Yue, Chen Jianxia. Fault analysis and prediction of transmission line based on fuzzy K-nearest neighbor algorithm[C]//Proceedings of the 12th International Conference on Natural Computation and Fuzzy Systems and Knowledge Discovery, 2016.