DOI:10.20033/j.1003-7241.(2025)02-0080-05.

基于大数据技术的DRGs绩效评价信息分析处理系统

陈立真
(郑州市妇幼保健院,河南 郑州 450000)

摘 要: 当前的信息处理分析系统在处理数据时缺乏实时性,导致系统平均响应时间较长。对此,设计基于大数据技术的DRGs绩效评价信息处理分析系统。根据绩效评价信息处理要求,建立包含三层结构软件层次总体架构;采用改进烟花算法建立信息集成模型,汇总DRGs绩效评价信息;采用离散化数据流对汇总的信息进行实时、快速聚类分析;最后结合内存存储模式和同步复制技术,实现信息的分布式存储。实验结果表明:所提系统在不同数据量条件下的系统操作指令平均响应时间始终保持在4 s,说明所提系统的实时性较好。

关键词: 大数据技术;绩效评价系统;信息集成;流处理;信息存储
中图分类号:TP311.13 文献标识码:A 文章编号:1003-7241(2025)02-0080-05

Information Analysis and Processing System of DRGs Performance Evaluation Based on Big Data Technology

CHEN Li-zhen
(Women and Infants Hospital of Zhengzhou, Zhengzhou 450000 China)

Abstract: The current system lacks real-time performance when processing data, resulting in a long average response time of the system. The DRGs performance evaluation information processing and analysis system based on big data technology is designed. According to the requirements of performance evaluation information processing, it establishes the overall software hierarchy structure including three layers. The improved fireworks algorithm is used to establish an information integration model and summarize DRGs performance evaluation information. The discrete data stream is used for real-time and fast clustering analysis of the summarized information. Finally, the memory storage mode and synchronous replication technology are combined to realize the distributed storage of information. The experimental results show that the average response time of the system operation instructions under different data volumes is 4 s, indicating that the proposed system has good real-time performance.

Keywords: big data technology; performance evaluation system; information integration; stream processing; information storage

0 引言

评价信息是DRGs绩效评价的基础,想要得到准确的绩效评价结果,需要先提升评价信息质量[1]。通常情况下,绩效评价信息可以通过问卷法、调查法等多种方法收集,多方法结合使用收集了全面且完整的评价信息。但也正是因为信息采集渠道不同,最终汇集的信息量较大且杂乱无章[2]。因此,如何进行有效的信息处理和分析,提取有价值的信息进行DRGs绩效评价,成为一项重要问题。

已有研究采用层次分析法与功效系数,对所有待处理数据进行优先级别的判断。针对急需处理的数据,运用语义识别模型对海量信息进行初步处理,再通过大数据技术,完成信息的存储和分析。但是,该系统具有响应时间较长的问题[3]。也有研究依托于分布式架构设计信息处理策略,将所有数据汇总到一起,进行报文解析处理。通过云计算节点将解析后的数据保存至数据库,进行数据聚类分析,得到信息处理分析结果。但是该方法处理后的信息可用性较低[4]。或针对信息处理分析问题,以可逆信息隐藏算法为核心进行信息处理。从信息隐藏、信息提取和信息恢复三方面入手,得到信息处理分析结果。但是该方法的信息处理系统鲁棒性较差[5]。

对此,本文设计基于大数据技术的DRGs绩效评价信息处理分析系统。根据绩效评价信息处理要求,搭建包含表示层、中间层和数据层三层应用结构的软件层次架构,对评价信息进行处理、分析。采用改进烟花算法IF-WA建立信息集成模型,汇总待处理的DRGs绩效评价信息。依托于大数据技术,设计基于离散化数据流的分布式数据处理模型,以批处理的形式对汇总的信息进行实时、快速的聚类分析。最后结合内存存储模式和同步复制技术,实现信息的分布式存储。至此,完成绩效评价信

*基金项目:河南省医学科技攻关计划项目(LHG20200742)
收稿日期:2023-02-23

息的处理分析。实验结果表明，所提系统在处理DRGs绩效评价信息时，在不同数据量条件下，系统操作指令平均响应时间始终保持在4 s，表明所提系统的实时性较好。

1 基于大数据的DRGs绩效评价信息处理分析系统

1.1 搭建信息处理分析系统总体架构

根据DRGs绩效评价信息处理要求，文中建立了包含表示层、中间层和数据层三层应用结构的软件层次架构，总体架构示意图如图1所示。

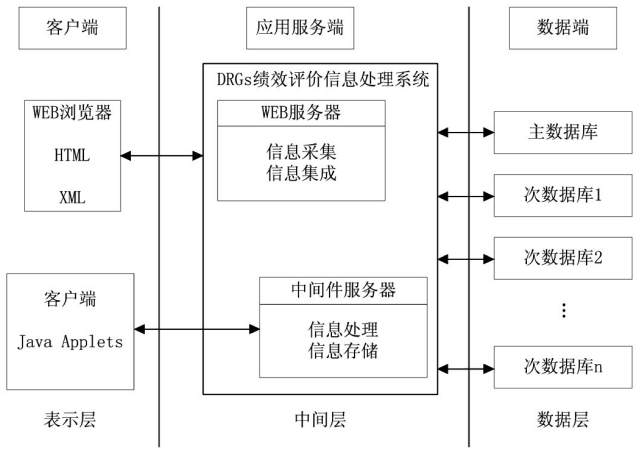


图1 系统软件总体架构示意图

图1中，数据层是三层结构中的最后一层，也是信息处理分析系统的基本层次结构，其中包含一个主要数据库，主要用来存放待处理的DRGs绩效评价信息，其下存在n个次数据库，用来保存处理后的评价信息。

中间层是信息处理分析的核心结构层，与数据层和表现层相连接，该层具有可扩展性，可以运行在服务器端，实现信息采集、信息集成、信息处理和信息存储等多项功能。

最后，表现层与中间层相连接，通过统一的应用界面，负责展示DRGs绩效评价信息处理分析结果，同时用户交互行为也是在该结构层得以实现。三层结构互相配合，完成DRGs绩效评价信息的流畅处理和分析。

1.2 构建DRGs绩效评价信息集成模型

信息的集成是系统进行处理和分析的基础[6]，文中建立的DRGs绩效评价信息集成模型如图2所示。该模型主要由两部分组成，一部分包括4条信息流，分别表示绩效评价主体和主体意见权重，另一部分则包括6条信息流，用以描述绩效评价对象。

应用图2所示的模型进行信息集成，需要经过确定评价主体权重、计算指标直觉模糊绩效综合评价值[7]、确定指标权重、计算融合评价指标绩效值四个步骤。

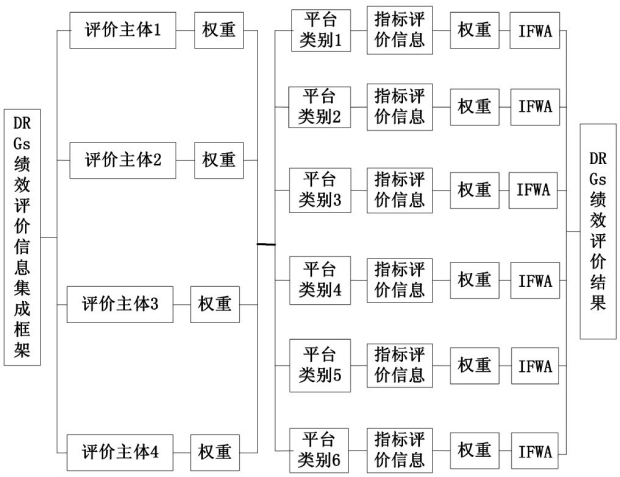


图2 DRGs绩效评价信息集成模型

其中，评价主体对应的权重具有差异化特点，具体评价权重计算公式为：

$$\eta_a^{(p_\varepsilon)}=\frac{(\delta_a^{(p_\varepsilon)}+\frac{\delta_a^{(p_\varepsilon)}}{\delta_a^{(p_\varepsilon)}+\rho_a^{(p_\varepsilon)}})}{\sum_{p_\varepsilon=1}^{A}(\delta_a^{(p_\varepsilon)}+\frac{\delta_a^{(p_\varepsilon)}}{\delta_a^{(p_\varepsilon)}+\rho_a^{(p_\varepsilon)}})},a=1,2,\cdots,n \tag{1}$$

式中，η表示评价主体权重，a表示指标，n表示DRGs绩效评价指标总数量，p_ε表示第ε个评价主体，A表示评价主体总数量，δ表示隶属度，ρ表示非隶属度。

为了得出评价指标对应的综合评价值，先量化分析所有评价结果，结合评价主体的差异性权重，建立如下所示综合评价矩阵：

$$R_{ab}=I_\varpi(r_{ab}^{(p_1)},r_{ab}^{(p_2)},\cdots,r_{ab}^{(p_A)}) \tag{2}$$

式中，R表示综合评价矩阵，b表示评价要素，$r_{ab}^{(p_1)}$表示评价者主体对评价要素关于指标的评价值，I表示改进烟花算法IFWA，ϖ表示指标权重向量。

其中，指标权重的计算需要依托于直觉模糊熵度量方法，针对综合评价矩阵内每个要素指标，分别赋予一个对应的直觉模糊数，再计算其对应的权重，具体计算公式为：

$$\varpi_b^{(a)}=\frac{1-E_{ab}}{\sum_{j=1}^{J}(1-E_{ab})} \tag{3}$$

式中，J表示评价要素总数量，E表示指标权重向量。

根据所有评价指标绩效值计算结果，建立融合后的综合评价矩阵，集结所有要素。

$$R'_a=I_\varpi(r_{a1},r_{a2},\cdots,r_J) \tag{4}$$

式中，R'表示融合后的综合评价矩阵。

通过上述处理，完成DRGs绩效评价信息集成，将其保存至系统数据库内，作为后续处理和分析的基础。

1.3 设计大数据信息处理分析技术

集成后的绩效评价信息，呈现为数据流的形式[8]，为了对海量集成信息进行快速处理，文中依托于大数据环

境，设计大数据信息处理分析技术，选用大数据技术中最活跃的离散化数据流Spark Streaming项目，建立并行计算框架[9]，实现对DRGs绩效评价信息的分布式数据流处理。

Spark Streaming流式数据处理模型可以同时进行批处理和流处理，模型结构如图3所示。应用Giallo模型进行数据处理分析时，先将原始信息队列分成多个信息段，得到一个离散流，再对信息进行微批处理。

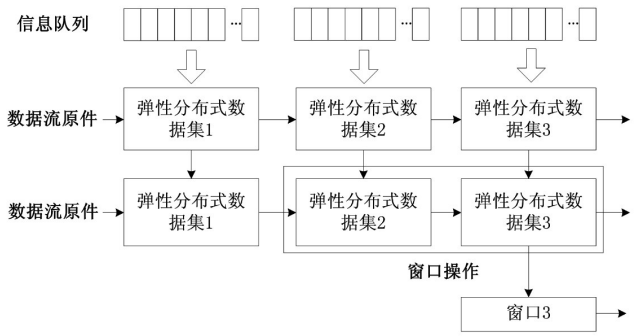


图3 Spark Streaming流式数据处理模型

如图3所示，应用Spark Streaming流式数据处理模型进行数据处理时，会将每一次接收的信息组成弹性分布式数据集，再通过窗口对数据集进行转化操作和动作。这种操作形式下，可以将不同时刻的评价信息结合起来，提升数据处理效果。

上述处理模型运行时，同步应用了大数据技术进行信息聚类，设置欧氏距离为衡量指标[10]，将DRGs绩效评价信息划分为多个簇，每个簇内包含的信息具有属性或特征，便于DRGs绩效评价。聚类过程中，欧氏距离的计算公式为：

$$D(\xi,F)=\sqrt{|s_1-f_1|^2+|s_2-f_2|^2+\cdots+|s_m-f_m|^2} \tag{5}$$

式中，D表示欧氏距离，ξ，F表示两条DRGs绩效评价信息，s表示信息ξ中的元素，f表示信息F中的元素，m表示DRGs绩效评价信息中元素总数量。

文中设计的信息处理分析系统，在信息处理阶段采用了基于密度的聚类算法Dbscan，将所有待处理的DRGs绩效评价信息，映射到一个数据空间内，选取一个数据点作为圆心，并定义一个半径，得到包含多个数据的圆形区域，作为数据邻域。

$$\tau(\xi)=\left\{F\middle|F\in\phi\wedge D(\xi,F)\le\theta\right\} \tag{6}$$

式中，τ表示数据邻域，ξ表示作为圆心的数据点，ϕ表示DRGs绩效评价信息数据集合，θ表示圆形区域半径。

在Spark Streaming流式数据处理模型中，通过上述聚类分析方法，对DRGs绩效评价信息进行处理，需要深入解析所有评价信息。考虑到评价信息的大数据量纲差异较大，在聚类处理之前需要先依托于极差标准化要求，得到归一化处理数据：

$$\overline{X}=\frac{X-X_{\min}}{X_{\max}-X_{\min}} \tag{7}$$

式中，$\overline{X}$表示归一化处理数据，X表示原始数据，$X_{\max}$、$X_{\min}$表示DRGs绩效评价信息中最大变量与最小变量。

根据欧氏距离值计算结果，按照从小到大的顺序，对所有DRGs绩效评价信息进行排序，得到如下所示排序后信息集合：

$$\chi=\left\{D(1),D(2),\cdots,D(k-1),D(k)\cdots,D(L)\right\} \tag{8}$$

式中，χ表示排序后数据距离集合，k表示欧氏距离，L表示最大欧氏距离值，$D(k)$表示k距离节点。

选取所有的k距离节点，并对这些节点进行拟合分析，生成一条变化曲线，根据曲线拐点确定数据邻域的实际取值范围。深入分析领域范围内类簇核心数量，信息聚类处理过程中，针对核心点进行遍历分析，得到DRGs绩效评价信息聚类分析结果。

1.4 实现处理数据的分布式存储

保存处理分析后的数据，是系统设计的最后一个环节。文中为了提升数据存储的实时性，设计基于内存存储的分布式存储模式，设置在该模式下存储的数据具有持久化，且数据存储容量较大。

分布式存储过程中，为了保证DRGs绩效评价信息的完整性，应用了同步复制技术。在随机云计算节点的支持下，复制阶段处理后的信息，将其传播至其他节点，最终转移至数据层的数据库内。同时在数据存储阶段，还可以根据系统负载情况，选择合适的信息复制方式，切换存储数据库[11]。此外，在该阶段提出了“失效接管”运行方式，通过高级复制功能，在主数据库运行存在故障时，可以将评价信息转移至次数据库进行接管，避免数据丢失。高级复制功能的应用，需要以当前大数据可用性为基础[12-13]。通过如下公式计算平均数据丢失时间，以此来判断是否执行复制行为。

$$\lambda=\frac{G^2}{2\Phi} \tag{9}$$

式中，λ表示平均数据丢失时间，G表示节点平均故障间隔时间，Φ表示节点故障的平均恢复时间。

按照上述计算结果，完成处理后信息向分布式数据库的合理转移，完成DRGs绩效评价信息处理分析结果的整体存储。

2 系统测试

文中针对DRGs绩效评价信息，设计了以大数据技术为核心的处理分析系统。考虑到该系统是自主研发的，为了确保系统具有良好的应用效果，需要在正式投入应用之前进行系统测试。

2.1 开发工具

在Jdeveloper 10132开发平台上，依托于Struts+JSP技术进行信息处理分析系统的开发。本系统的前后端操作界面，均属于B/S架构，二者的实现技术为java技术。而系统中数据库和Web应用服务器，分别选用了SQL server 20059i、Tomcat6服务器。

除此之外，为了确保系统正常运行，在系统开发阶段应用了自主研发的数据中心服务结构，可以建立多种主流数据库，并在系统运行过程中，对统计数据库进行随时修改，避免不同数据库内采集的DRGs绩效评价信息编码格式不统一。考虑到绩效评价信息存在一定的隐私性，在系统开发阶段，针对系统权限服务进行深入研究，将整个系统的资源划分为功能资源、操作资源和统计资源，并对不同种类的资源定义不同的权限。建立用户管理模块，根据系统登录用户的信息和组织信息，赋予用户对应的权限。依托于本体开发工具protégé设计一个具有开放性特点的初始本体库，在该库内保存本体数据，作为系统运行的基础。

在本体编辑界面导入DRGs绩效评价信息，并建立一个多类层次结构，每个类内都存在数据类型属性、对象类型属性。

2.2 系统基本功能实现

按照上述开发模式，得到基于大数据技术的DRGs绩效评价信息处理分析系统，软件开发人员和专业测试人员在系统开发结束后，都需要对系统基本功能进行测试，通过白盒测试方法，验证系统各模块功能都可以顺利实现。

系统基本功能实现阶段，由系统管理员向测试人员发布系统登录网址，输入登录网址后进入系统的登录页面，如图4(a)所示。



(a) 系统登录界面 (b) 用户权限设置界面 (c) 信息处理分析结果展示界面

图4 系统主要功能界面

考虑到该系统是应用在医院内部的，系统的登录账号需要由管理人员统一分配。在登录页面进行多次登录测试，验证不符合要求的用户名和口令，是否能够登录系统，系统登录测试结果如表1所示。

表1 系统登录测试结果

用户名	口令	预期结果	说明
User10	Pass10	进入系统	用户名和口令正确
User123	Pass123	进入系统	用户名和口令正确
User00010	Pass00010	进入系统	用户名和口令正确
	Pass	提示输入用户名	用户名为空
空格	Pass	提示无效用户名	用户名为空格
User	UserPass	提示用户名太短	用户名小于6位
User1000001	UserPass	提示用户名太长	用户名大于10位

根据表1可知，系统用户登录功能符合预期要求，可以满足用户的基本使用要求。

之后，进入权限管理页面进行批量授权，权限设置界面如图4(b)所示。最后，以某大型三甲医院为例，针对内分泌科、心内科、甲乳外科、心脏大血管外科四个科室，调查2017年−2019年的DRGs绩效评价信息，作为系统测试数据，导入系统内进行信息处理分析。根据检索要求，得到图4(c)所示的绩效评价信息处理分析结果。

通过上述实现结果可知，本次开发的系统主要功能均得到了良好实现，证明了文中设计系统具有可行性。

2.3 系统响应时间分析

随着DRGs绩效评价信息量的成倍增长，待处理的信息越来越多，而人们对海量数据处理实时性要求却越来越高，要求系统操作指令平均响应时间不大于5 s。因此，在确定系统可以正常运行后，以系统操作指令平均响应时间为衡量指标，验证系统性能。在不同数据量条件下分别应用本系统进行信息处理分析，得到图5所示的响应时间变化结果。

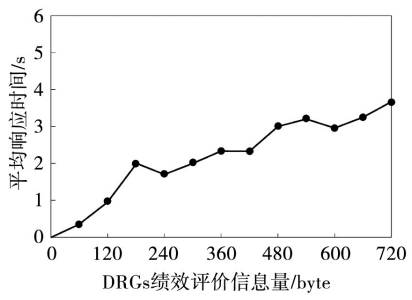


图5 不同数据量条件下系统操作指令平均响应时间

根据图5可知，随着待处理DRGs绩效评价信息量的增长，系统操作指令平均响应时间呈现出增长趋势，当信

息量达到720 byte时,操作指令平均响应时间长为3.72 s。总体来看,该系统的操作指令平均响应时间总是低于4 s的,符合信息处理系统的运行要求,体现了文中设计系统性能的优越性。

3 结束语

医疗数据的增长使得DRGs绩效评价信息处理难度越来越大,为了更好地进行数据处理和分析,文中考虑到系统面对海量数据的响应实时性要求,设计以大数据技术为核心的新型信息处理系统。根据系统测试结果可知,所提系统的操作指令平均响应时间满足预期要求,可以在短时间内有效处理大量绩效评价信息,提升医院内部绩效评价效率。

参考文献:

[1] 李文彬,陈晓绚.大数据背景下提升财政支出绩效评价数据质量的路径研究[J].学习论坛,2021(1):61-68.

[2] 刘晓玲,刘娜,秦富贞.基于数字信号处理器的舰船导航信息处理系统[J].舰船科学技术,2022,44(5):118-121.

[3] 李攀,刘庆杰,周兆军,等.大数据技术的震后救援信息处理平台研制与应用[J].科学技术与工程,2021,21(15):6154-6164.

[4] 周雪芳,刘树龙,周海龙.复杂环境中高可用船舶AIS大数据信息处理方法[J].舰船科学技术,2022,44(6):137-140.

[5] 方昱璋,张海文,郭嘉成.基于BIM与遗传算法的电网工程数据信息处理系统设计[J].电子设计工程,2021,29(6):168-172.

[6] 谭怀英,王丁禾,郭建明,等.大数据挖掘技术在预警探测信息处理领域的应用前景[J].现代雷达,2020,42(11):1-5.

[7] 祝旭,彭湘华,陈义.船舶嵌入式数字导航信息处理系统构建[J].舰船科学技术,2020,42(24):121-123.

[8] 张春霞.基于可逆信息隐藏算法的图像信息处理系统设计[J].现代电子技术,2020,43(19):23-26.

[9] 范一群,何亦斌,张凌,等.基于DRGs的加权TOPSIS法对江西省南昌市某三甲综合医院医疗绩效评价[J].中国卫生事业管理,2020,37(4):265-268.

[10] 李卫星,王峰,李智国,等.面向多源数据的军事信息系统设计[J].中国电子科学研究院学报,2020,15(3):237-243.

[11] 康达,张卓.基于多元评价模型的多主体协同创新绩效评价系统[J].自动化技术与应用,2022,41(6):161-165.

[12] 肖鹰,曾志丹,张艳.基于云计算下现代生态农业物联网监控系统的设计[J].农机化研究,2023,45(11):117-121.

[13] 陈海燕,康峰,苏立伟,等.基于数据挖掘的电力数据资产运营系统设计[J].自动化技术与应用,2024,43(9):159-162.

作者简介:陈立真(1983-),女,本科,高级工程师,研究方向:医疗卫生信息化。

(上接第12页)

果良好,能够在面对非结构障碍物时更快地抵达目标点。

参考文献:

[1] Roberge V,Tarbouchi M,Labonté G.Comparison of parallel genetic algorithm and particle swarm optimization for real-time UAV path planning[J].IEEE Transactions on industrial informatics,2012,9(1):132-141.

[2] Bolourian N,Hammad A.LiDAR-equipped UAV path planning considering potential locations of defects for bridge inspection[J].Automation in construction,2020(10):117-132.

[3] Li F,Sisi Z,Martijn K,et al.Universal path planning for an indoor drone[J].Automation in Construction,2018(95):275-283.

[4] Liu Y,Wang Q,Hu H,et al.A novel real-time moving target tracking and path planning system for a quadrotor UAV in unknown unstructured outdoor scenes[J].IEEE Transactions on Systems,Man,and Cybernetics:Systems,2018,49(11):2362-2372.

[5] He Z F,Zhao L.The Comparison of Four UAV Path Planning Algorithms Based on Geometry Search Algorithm[C].2017 9th International Conference on Intelligent Human-Machine Systems and Cybernetics(IHMSC).2017(2):33-36.

[6] Heng L,Meier L,Tanskanen P,et al.Autonomous obstacle avoidance and maneuvering on a vision-guided MAV using on-board processing[C].IEEE International Conference on Robotics&Automation.IEEE,2011:2472-2477.

[7] Hornung,Armin,et al.OctoMap:An efficient probabilistic 3D mapping framework based on octrees[J].Autonomous robots,2013:189-206.

[8] Xiao K,Tan S,Wang G,et al.XTDrone:A Customizable Multi-rotor UAVs Simulation Platform[C].2020 4th International Conference on Robotics and Automation Sciences(ICRAS).2020:55-61.

[9] Meyer J,Sendobry A,Kohlbrecher S,et al.Comprehensive simulation of quadrotor UAVs using ROS and gazebo[C].Simulation,Modeling,and Programming for Autonomous-Robots.Springer,Berlin,Heidelberg,2012:400-411.

作者简介:周恒旭(1993-),男,硕士,研究方向:无人系统与智能控制。