

文章编号:1001-9081(2016)51-0228-04

## 基于二次校正的教育考试大数据分析评价方法

杨有科<sup>1</sup>, 罗国辉<sup>2\*</sup>, 李 权<sup>1</sup>, 高西刚<sup>1</sup>, 姜文超<sup>2</sup>, 林 穗<sup>2</sup>

(1. 广东讯飞启明科技发展有限公司, 广州 510530; 2. 广东工业大学 计算机学院, 广州 510006)

(\* 通信作者电子邮箱 luoguo@mailb2014@163.com)

**摘 要:**针对教育考试质量评价过程中面临的单一、片面、复杂特性以及由此而导致的考试评价不科学、不客观等问题,提出了基于二次校正的考试数据分析与评价方法。首先定义考试数据属性,包括正向属性与反向属性,建立梯度式评价规则;其次在定义数据属性与基本评价规则基础上对考试数据进行基于谱聚类的分层过滤,标识出基本的群体簇;然后基于梯度式评价规则对群体簇进行二次校正,识别出因特殊原因而导致分数变动较大的特殊群体,为教师发现问题和有针对性辅导提供准确判断依据;最后基于云技术研发了面向跨区域、多用户的教育考试大数据分析评价服务平台。应用实践证明,该平台可有效提高考试数据获取、存储效率,同时为教育主管和实施部门提供有效分析工具。目前该平台服务对象已遍布全国13个省、100多个城市,考试数据获取、存储与评价分析业务占全国学历考试市场的40%。

**关键词:**教育考试;大数据;分层过滤;二次校正;评价分析

**中图分类号:** TP301.6 **文献标志码:** A

### Second revision based analysis and evaluation method for education and test data

YANG Youke<sup>1</sup>, LUO Guohui<sup>2\*</sup>, LI Quan<sup>1</sup>, GAO Xigang<sup>1</sup>, JIANG Wenchao<sup>2</sup>, LIN Sui<sup>2</sup>

(1. Guangdong Morning Star Technology Company Limited, Guangzhou Guangdong 510530, China;

2. School of Computer, Guangdong University of Technology, Guangzhou Guangdong 510006, China)

**Abstract:** To increase the efficiency and accuracy of test data analysis and evaluation, a second revision based analysis and evaluation method for education and test data was proposed. The positive and negative attributes of the test were defined, and a gradient evaluation rule was constructed. Based on the defined attributes and evaluation rules, a novel spectral clustering algorithm was presented to specify the data into different clusters. Then, the second revision based analysis and evaluation method was implemented on the different clusters to revise the scores of the testers. This second revision method could find the special groups whose scores were low while their real ability were excellent. A real analysis and evaluation platform was developed based on cloud technology. The practical applications from more than 100 cities in 13 provinces show that this method can increase the efficiency of obtaining, storing, analyzing, and evaluating the test data for the education management department. This platform has owned more than 40% market of the education and test data services in China.

**Key words:** education and test; big data; layered filtering; second revision; analysis and evaluation

## 0 引言

教育考试大数据分析评价是对学生原始考试数据进行挖掘,反馈数据关系中隐藏的大量信息和规律,辅助教育部门、教师、学生分析考试结果形成的原因、考试结果与学生实际能力之间的拟合度等,从而在此基础上改进教育与考核的过程与方法。科学、准确、全面的考试数据分析与评价对学生自我诊断、教师反馈总结、教育考试政策制定等具有非常重要的意义,具体体现在:1)理性的自我诊断有助于学生认识自身学习的优势与不足、提高自我效能感、健全自我心理、放平自我心态、准确自我定位,以免在缺乏相关数据支持的情况下对自己的学业成就形成误判。2)教师可以通过对考试结果的分析与总结,联系教学、反馈教学以及反思教学,寻求考试结果形成的原因,客观理性地评价自身教学的不足,及时改进

修正教学计划,以提高教师教学质量。3)教育主管部门可以通过考试评价与分析结果调整教育考试政策、完善考试评价的形式和方法,避免对考试结果只进行统计学生总成绩、平均成绩与及格率等简单的处理方式。4)可以通过对学生考试结果的多维度分析,加强家长与学校之间的信任与合作,为共同制定精确到学生个体的学习与生活计划提供参考和借鉴。

欧美等西方发达国家对教育分析、评价以及依此而进行的按学生特长进行分流非常重视。美国教育进展评价(National Assessment of Educational Progress, NAEP)是美国全国性的评价学生学业成就的评价体系,这项评价由美国国会授权、由教育部所属的全国教育统计资料中心管理、由教育考试服务中心(Educational Testing Service, ETS)实施。英国《1988年教育改革法》规定国家课程评价不仅有统一的测试,还要求教师把自己平时对学生的评价与之相结合。澳大利亚

收稿日期:2015-10-12;修回日期:2015-12-05。

基金项目:国家科技支撑计划项目(2015BAK19B03);广州市科技计划项目(2014XYD-007)。

作者简介:杨有科(1977-),男,湖南涟源人,主要研究方向:大数据与云计算、教育考试系统; 罗国辉(1991-),男,湖南永州人,硕士研究生,主要研究方向:云计算、大数据; 李权(1978-),男,江苏泰州人,主要研究方向:教育信息化、教育大数据分析; 高西刚(1975-),男,山东临沂人,主要研究方向:教育信息系统、大规模考试系统; 姜文超(1977-),男,山东潍坊人,讲师,博士,主要研究方向:分布式计算; 林穗(1972-),女,广东茂名,副教授,硕士,主要研究方向:云计算、云存储、操作系统。

的教育进展评价(National Assessment Program, NAP)对全国各地的教育水平和发展作出定期地、系统的评价。我国教育评价理论研究也有几十年的历史,从国家层面,明确提出要开展由政府、学校、社会各方面共同参与的教育质量评价活动<sup>[1]</sup>。2010年教育部考试中心和北京师范大学合作成立了国家教育考试评价研究院,开展教育评价方面的研究和尝试。

但是,当前考试数据分析与评价主要存在以下不足:

1)对考试数据处理简单化。虽然学校对考试测评非常重视,但对考试结果的处理缺乏应有的关注,存在对考试结果仅以分数为评价处理标准的简单化倾向。2)考试结果分析模糊化。考试结果处理最重要、最有价值的环节是教师及学生共同对考试结果进行分析与反思,寻求考试结果的形成原因,从而有针对性地矫正学生的不良学习方式及转变教师的教学策略。然而教师常常忽视对试卷的效度、信度、难度及区分度等特征量的分析。3)考试结果反馈单一化。考试反馈包括学生、教师、家长、命题组、教育主管部门等多方互动与反馈,然而目前考试结果的单一化极易导致试题命制的僵硬化,丧失试题的生命活力,不利于教育主管部门把握整体教育教学状况,制定适切性的教育教学政策。总之,目前就考试结果处理现状而论,教师及学校管理者并没有意识到整体把握考试结果对学生学业的掌握及教师教学方法改进的重要性。

本文针对教育考试质量评价过程中面临的单一、片面、复杂等特点以及由此而导致的考试评价不科学、不客观等问题,提出了基于二次校正的考试数据分析与评价方法<sup>[2]</sup>。1)定义考试数据属性,包括正向属性与反向属性,建立梯度式评价规则;2)在定义数据属性与基本评价规则基础上对考试数据进行基于谱聚类的分层过滤,标识出基本的群体簇;3)基于梯度式评价规则对群体簇进行二次校正,避免单一分数评价机制的不科学性;4)基于云技术研发了面向跨区域、多用户的教育考试大数据分析评价服务平台。应用实践证明该平台可有效提高考试数据获取、存储效率,同时为教育主管和实施部门提供有效分析工具<sup>[3]</sup>。目前该平台服务对象已遍布全国13个省、100多个城市,考试数据获取、存储与评价分析业务占全国学历考试市场的40%。

## 1 数据分层过滤与二次校正算法

### 1.1 数据属性

结合考试数据挖掘主题、被试学生特点、梯度评价规则以及二次校正目标,定义考试数据的连续性得分属性。以客观题型为例:一个选择题有四个答案,其中一个正确答案,但从命题者的角度看,另外三个选项并非毫无根据地随便加上去的,有一些错误选项离正确答案“距离”很近,学生选择错误选项的原因并非对此知识点一窍不通,而仅仅是因为考虑不周或大意,被具有迷惑性的选项所欺骗,在传统考试评价过程中,这一特点没有得到准确反映。相反,其中有个别选项确实与该知识点没有任何关联,而选择此选项基本可以判断被试学生对此知识点一窍不通。显然这两类被试者是不同的,考试分析评价应该对其有区别对待,并进行针对性辅导和处理。另外,针对不同考试目标,考察学生的能力不应该单纯唯考试分数论,有些考核目标希望得到学生的综合能力考评,包括实践能力、特长、兴趣等。因此,在考试命题时针对不同的题型、可能的不同作答,以及不同考试目标进行考试数据正反多维属性定义,建立考试数据的多维属性模型<sup>[4]</sup>,数据模型如图1所示,除得分属性外,同时设置事实属性,反映被试者万方数据

对考试知识点的事实掌握情况,为二次校正提供数据基础。

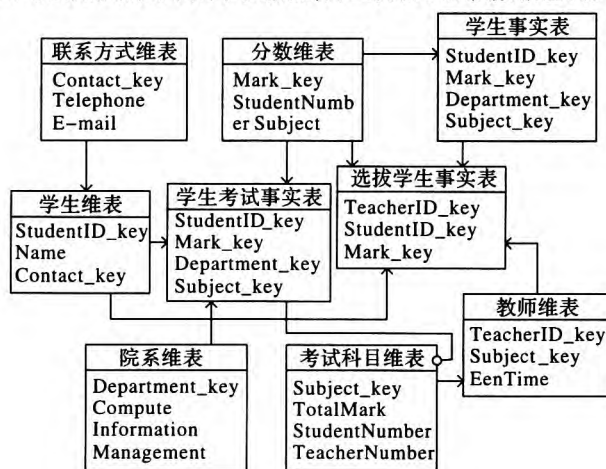


图1 数据模型

根据数据模型定义,考试评价也分成两类:常规评价规则与校正规则。常规评价规则仅包括某次考试的考试成绩。校正规则包括对考生错误作答的梯度评价、平时成绩、考试成绩、特长爱好,以及家庭背景、父母职业等情况。对错误答案的梯度评价可以为发现因特殊原因导致答案出错的特殊学生群体,平时成绩用于评价考生平时学习状态,考试成绩是对某阶段学习的直接评价,特长爱好是对学生综合评价的重要依据,家庭背景和父母职业是考察学生特长和偏科现象的一个重要因素,同时可以分析不同家庭背景和父母职业与学生学习成绩之间的关系。以客观题为例,在四个选项中,唯一正确的答案得分100,而为考察某知识点而设置的混淆答案可以根据其与正确答案之间的关联程度定义为得分50~70,而与改题目考察知识点完全没有关联的选项则得分为0。另外,设置一个统计规则,如果某学生在某科考试中,选错选项的题目中与所考察知识点关联程度高的选项较多,则可以提高其错误选项的得分权值。相反,如果错误选项关联度小的选题增多,则适当减小其得分权值。校正规则更能客观地反映被试者对所考知识点了解的程度。事实评价规则可以发现“大意出错”“彻底不会”“随机乱选”等特殊被试群体,并对其进行针对性辅导和处理。为实现该目标,本文首先对被试数据进行基于谱聚类的分层过滤<sup>[5]</sup>,之后对分层过滤的各数据簇进行二次校正。

### 1.2 基于谱聚类的分层过滤

聚类是按照一定的规律和要求对事物进行区分和分类的过程,在这一过程中没有任何关于分类的先验知识,仅靠事物间的相似性作为类属划分的准则。与传统的聚类算法相比,谱聚类具有能够在任意形状的样本空间上聚类,且收敛于全局最优解的优点。谱聚类算法可简要描述为:给定一个数据集 $X = \{x_1, x_2, \dots, x_n\}$ ,  $x_i \in \mathbb{R}^N$ 。根据数据集 $X$ 建立加权图 $G = (V, E)$ 。其中 $V = \{v_i, i = 1, 2, \dots, n\}$ 是顶点的集合, $E = \{e_{ij}\}$ 是连接顶点 $(v_i, v_j)$ 的边。图 $G$ 中每一个节点 $v_i$ 与数据集 $X$ 中的 $x_i$ 相关。采用一个相似度准则构造图 $G$ 的顶点之间的相似度矩阵 $W$ ,  $W \in \mathbb{R}^{n \times n}$ 。对于两个对象 $x_i$ 和 $x_j$ ,其相似度 $s$ 为 $s(x_i, x_j) = \exp(-\|x_i - x_j\|^2 / 2\sigma^2)$ 。算法输入图 $G$ 的相似度矩阵 $W$ 和聚类个数 $k$ ,输出 $x_1, x_2, \dots, x_n$ 的聚类结果。

基于谱聚类的分层过滤算法根据考试答卷数据对象的特点在谱聚类基础上,增加了答卷数据对象的距离矩阵,得到适合答卷数据分析的谱聚类算法<sup>[6]</sup>。假设研究的答卷数据对象 $x$ 有 $p$ 个测试点,表示为 $(x_1, x_2, \dots, x_p)$ ,两个被试答卷数据对



象  $u$  和  $v$  之间的距离  $d(u, v)$  可定义为:  $d(u, v) = \sum_{i=1}^n |u_i - v_i|$ 。其中,  $|u_i - v_i| = 1, u_i \neq v_i$  或  $|u_i - v_i| = 0, u_i = v_i$ ;  $i = 1, 2, \dots, p$ 。根据答卷数据对象之间的距离, 两个被试的答卷数据对象  $u$  和  $v$  之间的相似度  $s(u, v)$  可定义为:

$$S(u, v) = \exp(-d(u, v)/2\sigma^2)$$

通常情况下,  $\sigma = 1$ 。答卷数据的谱聚类算法对答卷数据进行分层过滤分析算法包括以下几步。

1) 读入考试数据库记录。答卷数据对象集合表示为  $X = \{x_1, x_2, \dots, x_n\}$ ,  $x_i \in \mathbf{R}^p$ 。其中,  $x_i$  为第  $i$  个被试的答题信息,  $p$  为试题数。计算被试的答卷数据对象的距离矩阵  $A$ 。

2) 计算被试的答卷数据对象的相似度矩阵  $W$ , 矩阵  $W$  的阶数与矩阵  $A$  的阶数相同。

3) 计算扩展邻接矩阵  $L$ 。根据相似度矩阵  $W$  建立一个对角矩阵, 记为  $D = (d_{ij})_{n \times n}$ ,  $d_{ii} = \sum_{j=1}^n w_{ij}$ , 令矩阵  $L = D^{-1/2} W D^{-1/2}$ 。

4) 计算矩阵  $L$  的特征值和特征向量, 选择前  $k$  个特征向量  $u_1, u_2, \dots, u_k$ , 构建以特征向量  $u_1, u_2, \dots, u_k$  为列的矩阵  $U \in \mathbf{R}^{n \times k}$ , 并对矩阵  $U$  的每一行进行单位化处理。

5) 设  $y_i \in \mathbf{R}^k (i = 1, 2, \dots, n)$  是对应矩阵  $U$  的第  $i$  行的向量。用  $k$  均值算法, 将向量  $y_i (i = 1, 2, \dots, n)$  聚类到  $C_1, C_2, \dots, C_k$ 。

6) 建立映射  $x_i \in \mathbf{R}^p \mapsto y_i \in \mathbf{R}^k, i = 1, 2, \dots, n$ 。根据  $x_i$  和  $y_i$  之间的对应关系, 利用 5) 中矩阵  $U$  的行向量  $y_i$  的聚类结果, 确定点  $x_1, x_2, \dots, x_n$  在聚类中的结果。

### 1.3 二次校正算法

基于谱聚类的分层过滤是对被试者进行的第一次分类过程, 该分类过程中仅仅用到被试者的考分数据信息, 但是在一些特定的选拔过程中, 仅仅依靠分数来衡量被试者的能力显得科学性不足, 不能真实反映被试者综合能力以及应试过程中因特殊情况导致的特殊群体。因此, 在第一次针对考试分数的分层过滤基础上, 对聚类结果进行二次校正<sup>[7]</sup>。以单选题客观选择题为例, 假设某试题有四个答案 A、B、C、D。其中: A 为正确答案, 得分为  $x$ ; B 为强相关答案, 得分为  $k_1 x$ ; C 为弱相关答案, 得分为  $k_2 x$ ; D 为不相关答案, 得分为  $k_3 x$ 。  $k_1 > k_2 > k_3$  且取值区间为  $[0, 1)$ , 假设考试题目数量为  $N$ , 某学生答错的题目数为  $M, M \leq N$ 。答错题目中权重为  $k_1, k_2, k_3$  的错题数目分别为  $M_1, M_2, M_3, M_1 + M_2 + M_3 = M$ , 则该被试者最终校正得分  $X$  记为:

$$X = (M_1 k_1 / M + M_2 k_2 / M + M_3 k_3 / M + (N - M) x) \quad (1)$$

同理, 对于多项选择题, 假设某题目存在 A、B、C、D、E、F 六个选项, 对应答案出错的题目其正确答案的选项个数为  $N^*, 1 \leq N^* \leq 6$ , 被试者选对的答案选项数为  $M^*, 0 \leq M^* \leq N^*$ , 被试者选错的答案选项数为  $K^*, 0 \leq K^* \leq 6 - N^*$ , 则被试者对该多选题的最终得分  $Y$  定义为:

$$Y = \left( \frac{M^*}{N^*} - \frac{K^*}{6 - N^*} \right) * x^* \quad (2)$$

基于以上描述, 可以在被试者基本考试成绩分类的基础上对其实际能力进行一次新的校正和衡量, 二次校正算法描述如下:

1) 输入单选与多选题目总数分别为  $N, N^*$  以及单选多选题目的开始与结束题号, 每个单选与多选题得分分别为  $x$ 、

万方数据

$x^*$ , 并输入多选题答案得分标准表, 表中的每个记录是一个三元组  $(M^*, N^*, K^*)$ ;

2) 将答题卡进行扫描

3) 对于单选题, 根据式(1) 计算单选题校正得分  $X$ ;

4) 对于多选题, 根据式(2) 计算多选题校正得分  $Y$ ;

5) 计算总校正得分  $X + Y$ 。

在分层过滤和二次校正之后, 每个被试者都具有两个成绩: 真实得分与校正得分, 每个被试者又处于各自的类簇中, 这将为教师分析试卷、被试者的能力提供更精准的判断依据。

## 2 系统实现与分析

### 2.1 总体架构

为了验证分层过滤与二次校正策略对考试大数据分析方面的效率和准确性, 研发了基于云计算的教育考试大数据分析与评价服务平台, 平台总体建设包括一个中心, 两级应用和三个系统, 总体架构如图 2 所示。

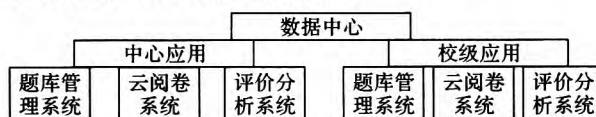


图2 教育考试分析与评价云服务平台总体架构

其中: 一个中心指构建教育评价数据中心, 满足区域内教育评价数据的存储和评价引擎的部署。两级应用主要是中心应用和校级应用。中心应用实现题库、卷库的集中管理, 实现教育评价数据的汇集与分析, 输出个性化的评价分析报告; 校级应用可以实现校本的常态化考试应用及阅卷, 如月考、期中考试、期末考试等。三个系统主要包括题库管理系统、云阅卷系统和评价分析系统<sup>[8]</sup>。

系统开发过程中, 首先基于多年来对教育数据的积累和业务基础, 建立教育分析与评价基础数据库<sup>[9]</sup>, 并研发教育考试数据采集、存储、过滤、挖掘等核心技术; 其次, 基于基础数据库和多维分析方法, 建立教育考试数据分析与评价模型, 设计教育考试数据分析和评价的处理流程; 接着, 基于多维因子及其参数控制设计多维交叉分析算法; 最后, 定义教育考试评价指标体系, 建立学科、学科内容、学生成绩三个评价引擎。

### 2.2 功能模块

平台主要包括三层: 基础设施层、公共支撑基础层以及应用服务层, 功能模块如图 3 所示。

基础设施层主要包括云基础设施和系统软件, 其主要功能是完成教育考试数据的分布式云存储, 主要基于云技术的 HDFS 系统进行文件系统的组织与管理; 公共基础支撑层是本评价分析系统的主要组成部分, 包含数据仓库部分<sup>[10]</sup>和二级数据处理工具, 数据仓库部分主要包括数据采集、数据交换、数据挖掘以及评价引擎的基础功能实现。公共支撑应用部分包括以学校、学科为单位的试题库、数据清洗、试卷生成、历史信息分析、学科分析引擎等。应用服务层是平台面向用户的服务模块, 主要包括基础考试数据的查询、网上阅卷综合服务系统、评价分析及管理子系统和试题库应用系统。该项目建设的教育考试数据分析评价云平台就是通过这些服务直接与用户互动, 因此也是最直接面向用户的部分。

### 2.3 测试分析

该系统既可以针对某个班级某门课程统计分析, 也可以对某个地区某门课程进行统计分析, 另外, 也可以针对某一道

题目对某个范围内的所有学生进行统计分析<sup>[11]</sup>。以某地区初中2年级政治与英语某次测试为例,分别计算被试者原始得分与二次校正后的得分,统计结果如图4、5所示。数据系列1表示原始得分,数据系列2表示校正得分。由图4和图5可知,被试者基本被分成三大类:第一类成绩在35~65,第二类成绩在65~90,第三类成绩在90以上。同时可以看到经过二次校正后,所有同学的成绩较原始得分都有所提高,但第

一类和第三类提高幅度明显小于第二类。第一类提高幅度较小的原因在于这一群体平时学习成绩就较差,因此在考试过程中,做错题目基本属于不懂的类型,因此其校正幅度较小。第三类提高幅度较小的原因在于该群体成绩本来就已经很高,已经严重压缩了可校正的空间。而第二类提高幅度较大,因为这一批学生属于中等,存在较大提升空间,教师应该针对其出错的题目给予重点关注和辅导。



图3 教育评价云平台功能

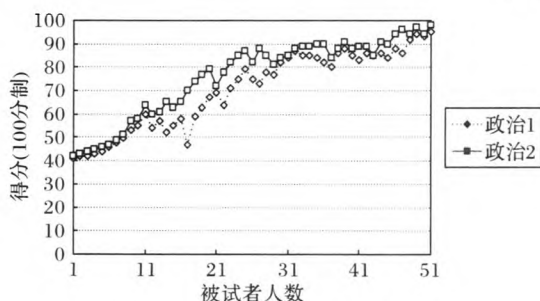


图4 某班政治课校对前后成绩对比

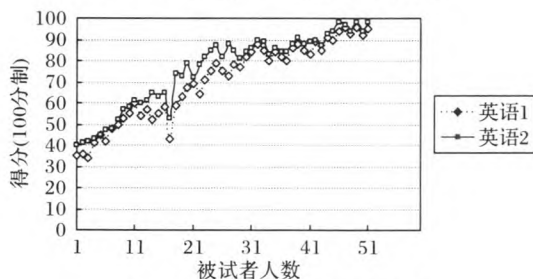


图5 某班英语课校对前后成绩对比

## 2.4 分析报告

基于云计算的教育考试大数据分析评价服务系统目前已经服务于全国13个省、100多个城市,考试数据获取、存储与评价分析业务占全国学历考试与分析评价业务市场的40%。系统支持多种类型的报告格式,通过各类参数,针对不同人群提供相关的分析报告。主要有:学生报告、机构报告和教师报告。

### 1) 学生报告。

学生用户报表主要展示了学生在考试中的考试情况的体现,从学生整体情况和学生单科情况对学生进行了统计分析,指出其强项及薄弱环节,给出针对性改进意见<sup>[12]</sup>。

### 2) 机构报告。

这里的机构,涵盖学校及学校的主管部门。机构用户报表主要展示了机构学生整体在考试中的考试情况的体现,从学生整体情况和学生单科情况对机构进行了统计分析,用各种指标研究机构的强项及薄弱环节。机构报表分为:测验概

况报告、机构科目报告、小题作答报告、机构知识对比、学生知识对比。

### 3) 教师报告。

教师的报告,主要是通过把知识点、知识板块、认知水平、能力领域建立以一对多的多维度映射关系<sup>[13]</sup>,以各维度项目的得分率为主要统计学指标,以校、年级或班级作为不同考生样本群体,指出教学的薄弱环节,对老师后续教学起到指导、启示和反馈的作用,分析报告示例如图6所示。

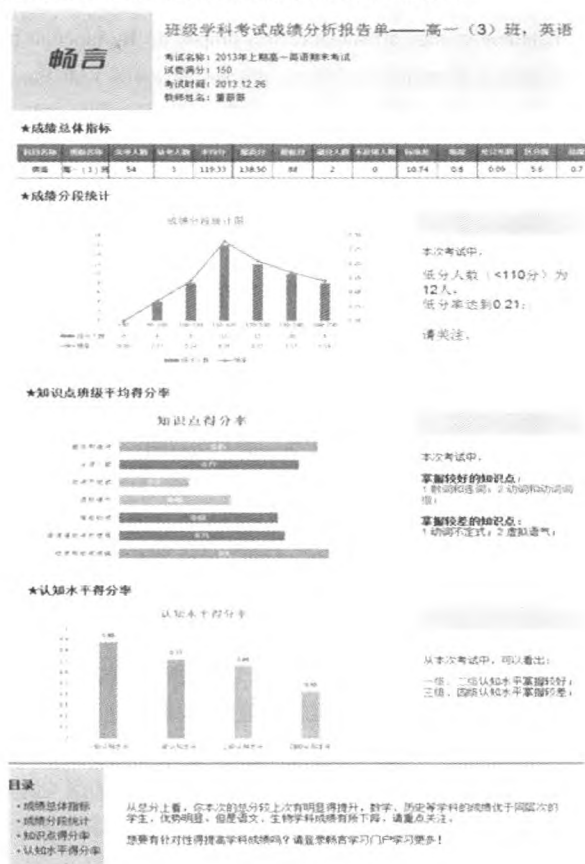


图6 教师报告示例

(下转第236页)

由测试的结果来看,在数据量大或者访问量大的情况下,MongoDB 的具有显著的优势。

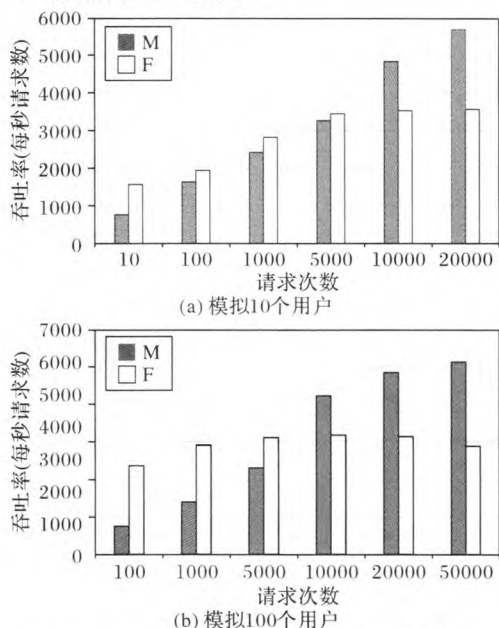


图 15 ApacheBench 模拟文件和数据库的对比测试

## 6 结语

本次研究使用 MongoDB 存储架构对现有的蛋白质组学数据存储架构进行改造,并且重新设计的系统能顺利通过最后的功能测试和性能测试。在科研中,在保证研发速度的同时,我们也需要追求数据处理的更高效率,同时也要控制研发成本。经过本次研究,可以看出 NoSQL 数据库技术已经发展到了很高的水平,使用 MongoDB 数据库的存储架构能很好地

代替原有的数据存储架构。未来的科研过程中,将继续对这种存储方法进行优化,以进一步提高其性能。另外,我们应当尝试使用较新的技术去解决原有的性能瓶颈,充分了解前沿的计算机技术,让蛋白质组学的研究更高效,处理过程更方便。

### 参考文献:

- [1] VENTER J C, ADAMS M D, MYERS E W, et al. The sequence of the human genome[J]. Science, 2001, 291(5507): 1304 - 1351.
- [2] 曾嵘,夏其昌. 蛋白质组学研究进展与趋势[J]. 中国科学院院刊, 2012, 17(3): 166 - 169.
- [3] 张克. 我国全面启动人类蛋白质组计划[N]. 科技日报, 2014 - 06 - 11(10).
- [4] 刘卓. 基于 NoSQL 的空间数据云存储的研究[D]. 郑州: 河南大学, 2014: 22 - 24.
- [5] 王光磊. MongoDB 数据库的应用研究和方案优化[J]. 中国科技信息, 2011(20): 93 - 94.
- [6] 郑静静, 叶焱, 刘太君. 基于 Flex, Red5 和 MongoDB 的视频直播、录制及存储系统设计[J]. 计算机应用, 2014, 34(2): 589 - 592.
- [7] CHODOROW K, DIROLF M. MongoDB 权威指南[M]. 程先锋, 译. 北京: 人民邮电出版社, 2014: 3 - 28.
- [8] 隋春明. 多核环境下 Key-Value 数据库性能分析与优化[J]. 电力信息与通信技术, 2014, 12(2): 44 - 49.
- [9] 朱亚兴, 余爱民, 王夷. 基于 Redis + MySQL + MongoDB 存储架构应用[J]. 微型机与应用, 2014(13): 3 - 5.
- [10] 王彦明, 薛云, 饶洪, 等. 图书馆云计算可行性研究[J]. 现代情报, 2014(7): 3 - 8.
- [11] 蔡平. 基于 Hadoop 的 NoSQL 数据库安全研究[D]. 上海: 上海交通大学, 2013: 23.
- [12] 王培建. 云计算环境下大规模数据存储技术研究[D]. 北京: 北京邮电大学, 2013: 15.

(上接第 231 页)

## 3 结语

本文针对教育考试质量评价过程中面临的单一、片面、复杂等特性以及由此而导致的教育考试评价不科学、不客观等问题,提出了基于二次校正的考试数据分析与评价方法<sup>[14]</sup>。1)定义考试数据属性,包括正向属性与反向属性,建立梯度式评价规则;2)在定义数据属性与基本评价规则基础上对考试数据进行基于谱聚类算法的分层过滤,标识出基本的群体簇;3)基于梯度式评价规则对群体簇数据进行二次校正,避免单一分数评价机制的不科学性,进一步识别出因特殊原因导致分数不高的特殊群体,以及彻底不懂所考察知识点的群体,从而为教师发现问题和有针对性辅导提供准确依据;4)基于云服务技术研发面向跨区域、多用户的教育考试大数据分析评价云平台。应用实践证明该平台可有效提高考试数据获取、存储效率,同时为教育主管和实施部门提供了有效分析工具支持,目前平台服务对象已遍布全国 13 个省、100 多个城市,考试数据获取、存储与评价分析业务占全国学历考试市场的 40%。

### 参考文献:

- [1] 靳希斌. 论教育服务及其价值[J]. 教育研究, 2003(1): 44 - 47.
- [2] 黄志纯, 刘必千. 关于构建高职生创新创业教育评价体系的思考[J]. 教育与职业, 2007(30): 80 - 81.

- [3] 韩圣龙. 网络信息检索工具评价指标[J]. 情报学报, 2001, 20(4): 87 - 93.
- [4] 李林花, 钱越英. 数据仓库多维分析模型的设计[J]. 计算机工程与应用, 2004, 40(11): 185 - 187.
- [5] 郭小芳, 李锋, 叶华. 基于主元的多元时间序列聚类分析方法研究[J]. 测控技术, 2012, 31(12): 104 - 107.
- [6] 奚茂龙, 盛歆漪, 孙俊. 基于多维问题的交叉算子量子粒子群优化算法[J]. 计算机应用, 2015, 35(3): 680 - 684.
- [7] 王雯霞. 一种专家系统中多维交叉层次时序关联分析挖掘技术[D]. 长沙: 国防科技大学, 2011: 1 - 70.
- [8] 朱庆华, 杜佳. 搜索引擎评价指标体系的建立与应用[J]. 情报学报, 2007, 26(5): 684 - 690.
- [9] 伍力, 吴捷, 周乐荣. 面向规划的广东电网数据库的建立及其应用[J]. 电力系统自动化, 1999, 23(13): 50 - 52.
- [10] 刘素梅, 徐礼华, 彭少民. 丹江口水库诱发地震基础数据库建设[J]. 武汉理工大学学报, 2005, 27(10): 89 - 92.
- [11] 董文鹭, 吴娟仙. 图像搜索引擎综合评价指标体系的构建[J]. 嘉兴学院学报, 2008, 20(4): 121 - 125.
- [12] 龚宪成. “三段四分法”在环境空气质量统计预报模型建立“基础数据库”中的应用[J]. 科技创新导报, 2011(4): 148 - 149.
- [13] 杨斌. 战略能力多维分析模型构建[J]. 生产力研究, 2008(16): 52 - 54, 70.
- [14] 邹良影. 温州科技职业学院的创业实践评价体系[J]. 华章, 2011(34): 163.