

10.3969/j.issn.1671-489X.2019.08.011

学习管理系统中学生行为对学习成绩影响研究^{*}

——以浙江大学国际联合学院为例

◆邵韶昭

摘 要 为进一步探讨学习管理系统大规模应用于中外合作办学环境中对于教学活动的影响,对学习管理系统上用户数据进行更加深度的分析,探索不同专业学生在学习管理系统中访问行为及其对应的专业成绩关系,研究教师在学习管理系统中活跃程度对于学生活动影响,分析学生不同课程中表现特征来判断学生对于专业满意度。研究证明,学习管理系统大规模应用于中外合作办学环境后,对教学产生积极作用。

关键词 学习分析;数据挖掘;学习管理系统;高等教育;中外合作办学;Blackboard

中图分类号: G645 **文献标识码**: B

文章编号: 1671-489X(2019)08-0011-07

Impact of Learning Management System to Students' Academic Achievement at Sino-Foreign Joint Institutions and Joint Programs in China//SHAO Zhaozhao

Abstract Further research is conducted to review the impact of learning management system(LMS) to Sino-Foreign joint institutions and joint programs. The research will cover areas such as the connection between user activity and user academic scores, the impact of instructor's activity of LMS to student, the shared activity profile between different classes. The research proves a positive impact of learning environment at LMS toward Sino-Foreign joint institutions and joint programs.

Key words learning analytics; data mining; learning management

system; higher education; Sino-Foreign joint institutions and joint programs; Blackboard

1 引言

随着数据挖掘在教育行业的广泛应用,学习分析已发展成为一个单独的研究方向,备受广大教师、教育研究者关注。学习管理系统作为高等教育信息化建设的热门方向,其师生访问行为数据被学者广泛研究,相关研究成果已经被用于改善教学过程,建设学业预警机制等方面。

如今,信息技术已经成为高等教育的一个重要组成部分,高校中很多教学过程的管理以及教学活动的快速发展都离不开信息技术的支持,如学生信息管理系统(SIS)、学习管理系统(以下简称LMS)、多媒体互动教学环境等。在国外高校,学习管理系统从单纯的远程学习系统已经逐步转变成课堂学习的重要补充。国外高等教育机构教师利用该系统和学生进行互动,包括共享课件、在线测试、交流学习心得、查看平时成绩等。通过利用学习管理系统,打破教学过程中对于时间和空间的限制^[1]。

国际知名高校的学习管理系统在过去十几年的运行过程中积累了大量的用户行为数据,相关学者已经利用这些数据进行分析和研究,其结果被用来辅助教育质量评价工作的开展;预测学生成绩走向,对学习自觉能力较差的学生进行一定程度的学习干涉和预警^[2-3]。

^{*} 项目来源:浙江大学国际联合学院(海宁国际校区)课题“中外合作办学中信息化项目建设机制探索”(基金编号:1705)。

作者:邵韶昭,浙江大学国际联合学院(海宁国际校区),信息化工程师、工程师,研究方向为教育技术与数据挖掘(314400)。

community college[J].Computers&Education,2014(72):59-67.

[25]Park Y, Yu J H, Jo I H. Clustering blended learning courses by online behavior data: A case study in a Korean higher education institute[J].Internet & Higher Education,2015(29):1-11.

[26]Shimada A, Okubo F, Yin C, et al. Informal learning behavior analysis using action logs and slide features in e-textbooks[C]//The IEEE International Conference on Advanced Learning Technologies.2015:

116-117.

[27] 范逸洲,汪琮.学业成就与学业风险的预测:基于学习分析领域中预测指标的文献综述[J].中国远程教育,2018(1):5-15,44,79.

[28] 刘艳华,徐鹏.大数据教育应用研究综述及其典型案例解析:以美国普渡大学课程信号项目为例[J].软件导刊:教育技术,2014(12):47-51.

[29] 杨旻,胡珊珊,邢宏根,等.借鉴与反思:学习分析在国外高等教育领域的应用[J].现代教育技术,2018(2):107-112.

在这个过程中形成教育数据挖掘(Educational Data Mining, 简称 EDM)——一个更加深层次的研究方向。

在用户行为数据分析和学业成绩预测中,大量的数据挖掘算法(如关联算法、回归算法、聚类分析算法)被用来分析和构建学生的行为特征,通过这些特征,优化教学过程的开展,是教育数据挖掘对于教育领域的重大贡献。

目前,国内高校开始重视学习管理系统的使用,并且对系统运行过程中生成的数据进行了大量分析。相关学者利用分析结果,设计改善教学过程的方法^[4]。例如:根据学生在课程中的内容访问顺序,优化课程的内容设计;针对学生在课程中对相关模块的参与程度,对学习有风险的学生进行预警和提醒。

中外合作办学中,国外高校教学过程非常依赖学习管理系统,如教学材料共享、师生互动等^[5]。因此,不少中外合作办学机构在建校之初就搭建自己的学习管理系统,从而在系统运行过程中形成大量的用户行为数据。如何利用这些数据来评估学习管理系统对于教学的影响,发现教学过程中存在的问题,进而引领教学模式改革,缩短中外高等教育中的教学差距,是中外合作办学中中方办学机构的重要机遇。

2 研究现状

目前国际上教育数据挖掘开展较早,已经取得显著成效,著名的案例包括 George Siemens^[5]提出的学习分析模式, Leah P. Macfadyen 等^[3]学者开发的早期预警系统。其核心思想为:对学习管理系统中的数据进行收集、分析和预测,对过往学业困难的学生在系统中的行为进行分析,抽取关键特征,并用该特征对比当前系统用户行为,从而预测可能存在学业困难的学生,达到降低学生的辍学率,提高学生学业成绩。

国外研究过程大致分为三个阶段。

1) 数据收集。其中数据的收集主要包含学生的 ID、课程名称、系统中该课程每日每学生访问次数,以及该课程中每个模块的访问次数等。

2) 数据分析和建模。在数据分析中,相关学者对学生访问课程资源的行为特征进行分析,特征包括学生访问课程的次数、访问模块的次数、在系统中停留的时间等,通过这些数据,将相关学生进行聚类分组。

3) 学生成绩预测和干涉。在数据预测中,通过分类分析、关联分析,对历史数据中不同小组中学生的学习管理系统访问行为特征和学业成绩进行关联。在预测过程中,将现有学生的学期前中期的行为和历史数据进行对比,从而推出每个学生的学业困难指数。通过该指数来对学业困难的学生进行学习干涉,包括采用谈话、心理辅导、课外辅导等手段。

借助于学习管理系统中的数据,通过以上教学过程的数据分析手段,对目标学生有针对性地进行学业辅导和学习预警干涉。通过这些举措,国外很多高等教育机构提升

了教育质量。

基于国外教学研究取得的显著成就,相关国内学者也在国内教学机构重复该方法,并且也取得一定成效。通过分析学生在课程中的行为特征,对比历史数据,从而预测学生的学业困难指数,对存在学业困难的学生在学期中进行干涉,取得一定的成效^[6-7]。

目前在中外合作办学环境中关于分析学生在学习管理系统中成效的报道较少。在中外合作办学中,课堂授课模式、教学环境和国内传统教育模式存在差距,如在课堂授课模式中以小班讨论为主,在教学环境上基本采用外文授课。在这些外部因素的影响下,相关学者提出的方案能否在中外合作办学环境下取得成效有待考证。其次,目前主流的研究数据的取样都存在一定的局限性,特别是研究中数据取样仅限制在系统特定课程,其研究结论能否被扩展到其他系统课程存在一定质疑。

鉴于此,本研究以浙江大学国际联合学院使用学习管理系统 2017 年秋季到 2018 年春季课程的学生行为数据为例,从学年的角度来分析所在高校中不同年级、不同专业学生的活动特征和学业成绩的关联。基于分析的结果,探讨学习管理系统中学生的活动特征对中外合作办学中教学活动的影响。

3 研究方向和样本

研究问题 本文从不同角度分析学生在学习管理系统中师生行为和学业成绩的关系,包括:

- 1) 不同平均成绩类型的学生在系统中的行为轨迹;
- 2) 教师在系统参与度对学生行为轨迹的影响;
- 3) 学生在不同课程间行为轨迹共性。

研究环境介绍 本次研究截取浙江大学国际联合学院(海宁国际校区)本科生在 2017—2018 学年所有专业学生使用学习管理系统(Blackboard 平台)数据。

海宁国际校区现有浙江大学爱丁堡大学联合学院和浙江大学伊利诺伊大学厄巴纳香槟校区联合学院。目前校区内开设的专业总计六个,分别为生物医学专业、电气工程及其自动化专业、电子与计算机工程专业、机械工程专业、土木工程专业以及针对国际生的 Philosophy-Mathematics-Economics 本科专业。校区内国内外本科学生总计 279 人。

研究样本数据处理 用户的访问行为通过读取 Blackboard 数据库中的访问日志表获取。访问日志表中记录访问者的账号 ID、访问课程的 ID 以及名称、访问课程中的模块的链接、访问课程某个模块的时间戳。目前校区 2017—2018 学年本科生总计开设课程为 60 门,其中使用 Blackboard 平台作为学习管理系统的课程数量为 38 门,有效学生为 279 人。用户在学习管理系统中的访问行为可以分为访问次数、点击次数。

1) 访问次数。访问次数为学生在一天内访问学习管理系统中的相关课程的数量。一个学生在系统有多门课程,每门课程包含多个课程内容,但是无论学生一天内访问多

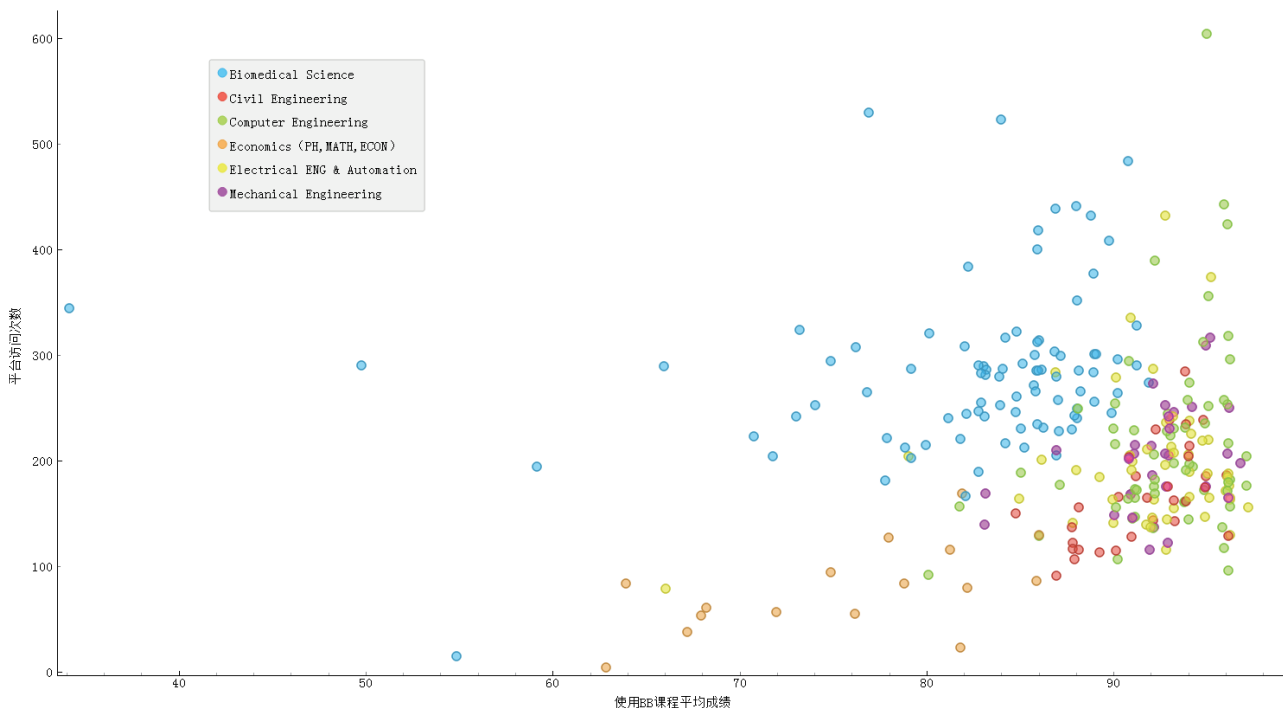


图1 学生访问次数和学生平均成绩（使用 Blackboard 平台课程平均成绩 VS 所有课程平均成绩）

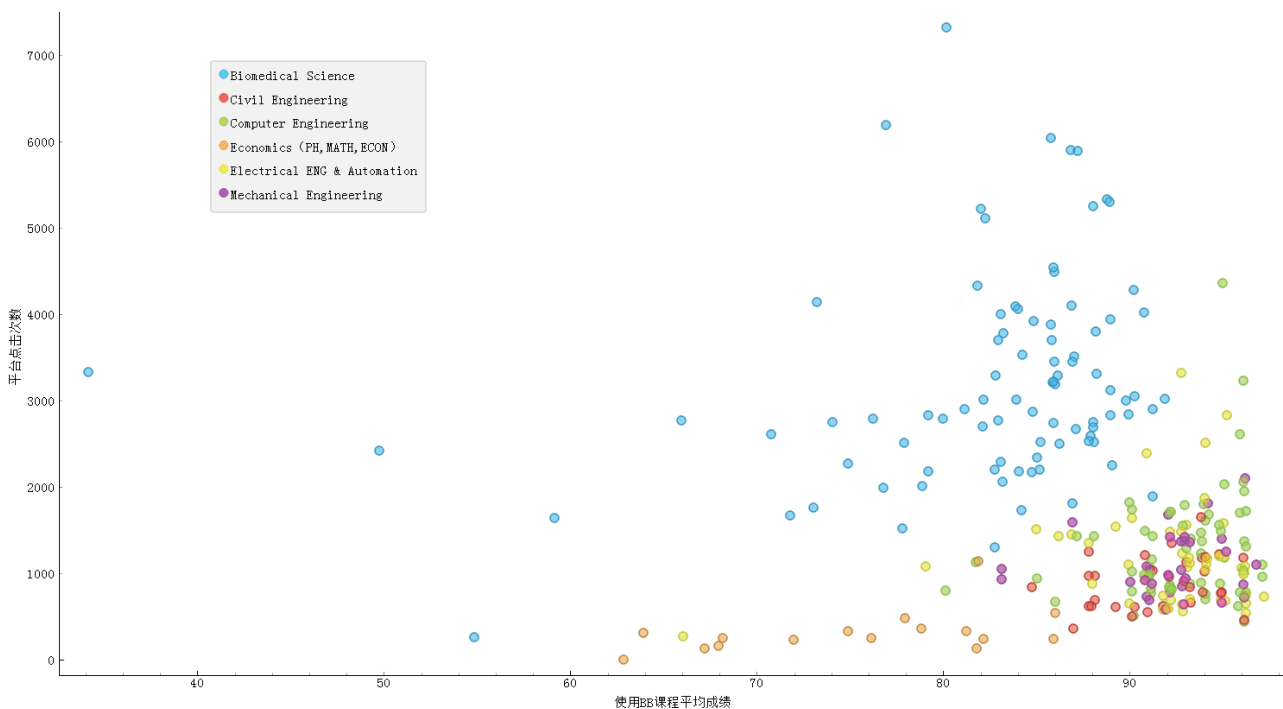


图2 学生点击次数和学生平均成绩（使用 Blackboard 平台课程平均成绩 VS 所有课程平均成绩）

少次课程内容，其在对应课程的访问次数统计仍然为1。访问课程次数意味着学生对于课程的关注度。

2) 点击次数。点击次数为学生在点击课程内容的次数，任何行为都被记录为一次点击次数。这些行为包括查看课程内容，和同学或者教师互动，下载作业题目、上传作业，参加在线测试等。

本次研究使用的统计软件为 R 语言，版本号为 3.4.4，

操作系统为 Windows 7 专业版。

4 研究结果

学生系统行为轨迹和学生平均成绩关联度分析

1) 在不区分专业和年级的情况下。在不区分专业和年级的情况下，获取校区内每名学生在 Blackboard 平台的访问行为（平均访问次数以及平均点击次数），对比其 2017—2018 学年平均成绩。学生访问行为分布和学生平均

成绩分布图分别如图 1、图 2 所示。

图 1 代表学生在学习管理系统中课程访问总次数和学生平均成绩的关系。其中, x 轴代表学生使用系统课程的平均成绩; y 轴代表学生在学习管理系统中课程访问总次数。

图 2 代表学生在学习管理系统中课程点击总次数和学生平均成绩的关系。其中, x 轴代表学生使用 Blackboard 平台课程的平均成绩; y 轴代表学生在学习管理系统中课程点击总次数。

在视觉上, 可以看到用户的平均访问次数和平均点击次数和其最终学业平均成绩关联度不大。通过 R 软件计算各变量之间 Pearson 相关系数。研究结果表明, 在不区分学生专业的情况下, 学生的平均成绩和访问次数、点击次数、停留时间未存在显著关联, 具体结果如表 1 所示。

表 1 访问行为和平均成绩矩阵

	使用 Blackboard 平台 课程平均成绩	所有课程 平均成绩
学生平均访问次数	0.06	0.13
学生平均点击次数	-0.18	-0.13

2) 在区分专业和年级的情况下。2016 级开始三个专业, 分别为生物医学专业、电气工程及其自动化专业、电子与计算机工程专业; 2017 新增三个专业, 分别为机械工程专业、土木工程专业以及针对国际留学生的 Philosophy-Mathematics-Economics 专业。在区分专业和年级后, 分别计算平均访问次数和平均成绩关系 Pearson 系数, 以及平均访问点击次数和平均成绩关系 Pearson 系数, 结果如表 2 所示。

表 2 学习管理系统中使用专业以及年级

入学 年级	专业名称	平均访问次数和 平均成绩关系	平均访问点击次数 和平均成绩关系
2016	生物医学	0.47	0.25
2016	电气工程及其自动化	0.56	0.55
2016	电子与计算机工程	0.51	0.48
2017	生物医学	0.39	0.29
2017	电气工程及其自动化	-0.2	-0.22
2017	电子与计算机工程	0.3	0.21
2017	机械工程	0.38	0.13
2017	土木工程	0.6	0.3
2017	Philosophy- Mathematics- Economics	0.56	0.46

研究结果表明, 学生访问次数和学生的成绩在区分专业和年级后, 存在中等程度相关, 说明在系统中访问频繁的学生, 平均成绩较高的可能性越大。结果同时指出, 学生点击次数也和学生的成绩存在一定的关联度, 但总体上来说, 关联程度平均低于访问次数。在结果中唯一出现访问次数和学生平时成绩存在负关联的是 2017 年电气工程及其自动化专业, 特别是其系统停留时间和学生平时成绩存

在中度程度负关联, 说明成绩越差的学生, 在系统停留的时间越长。该结果可能暗示另外一种学习情况。在个别专业课程中, 成绩相对较差的学生需要反复地在系统中进行学习。

基于用户活动的时序分析 研究显示, 访问次数和成绩的关联度总体上大于点击次数对于成绩的影响。Nespereira 等人在研究中指出, 不同成绩类型的学生在学期前、中、后期三个阶段, 在系统中表现出不同的访问行为轨迹^[8]。为更加清晰地显示平时访问系统的积极程度和学生平均成绩的关系, 对不同专业、学年中的学生的访问次数进行时间序列分解 (Time-series Decomposition)。分解公式如下:

$$Y_t = f(T_t, S_t, C_t)$$

其中, T_t 为长期趋势 (Trend), S_t 为季节变动 (Seasonal), C_t 为不规则变动 (Noisy)。研究首先采用 hierarchy 分类算法对平均成绩进行分类, 将学生的平均成绩分为三组, 求出每个小组每天在系统的平均访问次数 (小组中成员访问总数 / 小组成员数量)。因篇幅所限, 仅列出 2016 级生物医学专业分组结果, 每小组信息如表 3 所示。不同小组平均访问如图 3 所示。

表 3 2016 级生物医学专业学生成绩分类

分组名称	分组学生数量	平均成绩分布范围
优秀小组	33	85 ~ 100
中等小组	25	84 ~ 75
合格小组	9	60 ~ 74

为更加清晰地体现不同成绩分组中人均访问特征, 对其进行时间序列分解。按照样本数据分布类型, 在分解中使用加法分解^[8]。三个小组分解如图 4、图 5、图 6 所示。

在时间序列分解中, 分解出来的 T_t 为长期趋势 (Trend), 其数据更能表明该组学生的行为特征。因此, 对比三个不同小组的趋势分解数据, 对比结果如图 7 所示。

校区内教学时间段划分: 2017 年 9 月之前为英语强化课程, 10 月第一周为国庆节, 10 月第二周至 12 月 31 日为秋季教学阶段; 2018 年 1 月前两周进行秋季学期考试, 1 月后两周至 2 月前一一周为春季教学第一阶段, 3 月至 5 月 31 日为春季教学第二阶段。可以看出, 2016 级生物医学专业中, 平均分 85 分以上的学生在整个教学周期的前期、中期、后期, 相对于其他两组, 整体在系统中表现得更加活跃和积极。

其他专业研究室数据也表明, 平均成绩最好的一组在学年的开始以及期中阶段比其他两组学生更加积极。平均成绩最好的一组和平成绩较差的一组在活动积极程度上差距最为明显。平均成绩居中的一组在平时的活动积极程度稍低于平均成绩最好的一组, 但高于成绩最差的一组。该数据说明成绩好的一组学生在平时表现出更加积极的学习态度, 证明影响学生最终成绩的还是在平时的积极程度。



图3 不同成绩分组日均访问量分布图

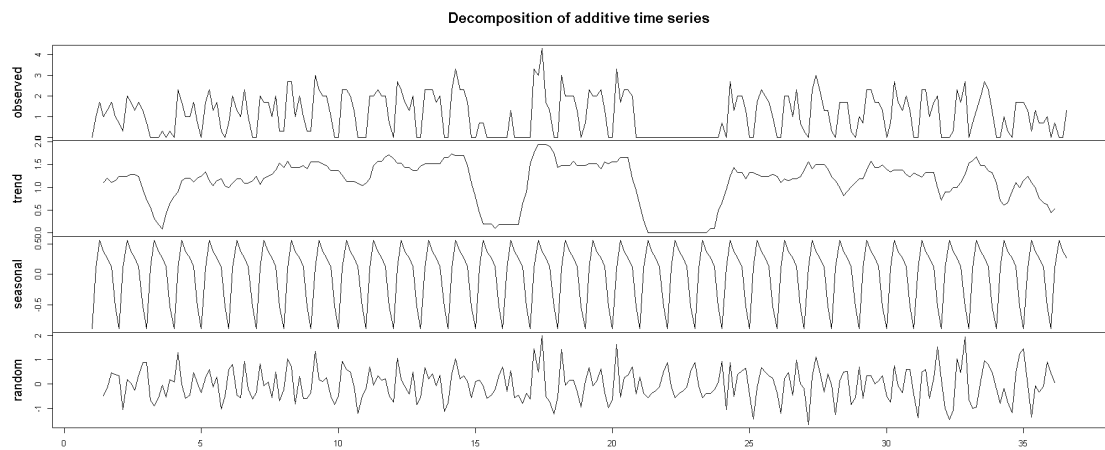


图4 合格小组平均访问次数时序分解

表4 教师在课程中访问次数和学生在系统中访问次数影响

	学生访问总数	学生点击总数	学生平均访问次数	学生平均点击次数	教师访问总数	教师点击总数
学生访问总数	1	0.84	0.62	0.59	0.77	0.73
学生点击总数	0.84	1	0.76	0.86	0.84	0.82
学生平均访问次数	0.62	0.76	1	0.92	0.83	0.74
学生平均点击次数	0.59	0.86	0.92	1	0.81	0.74
教师访问总数	0.77	0.84	0.83	0.81	1	0.95
教师点击总数	0.73	0.82	0.74	0.74	0.95	1

教师在课程中的活跃程度对于学生行为影响分析 分析之前的结果显示，访问次数对于学生的成绩影响最大。研究进一步分析系统中教师参与度对于学生在系统中访问次数的影响。首先统计每门课程中学生的访问总数与点击总数、学生数量、学生平均访问数量、学生平均点击数量、教师访问次数、教师点击次数，然后生产 Pearson 相关系

数矩阵，结果如表4所示。

数据显示，课程中学生的访问次数和教师的访问次数密切相关，即教师在系统中的活动越积极，学生在系统中的活动就会更加积极。

分析不同课程下学生行为的积极度 考虑到访问次数对于学生的成绩影响最大，进一步探索学生不同课程

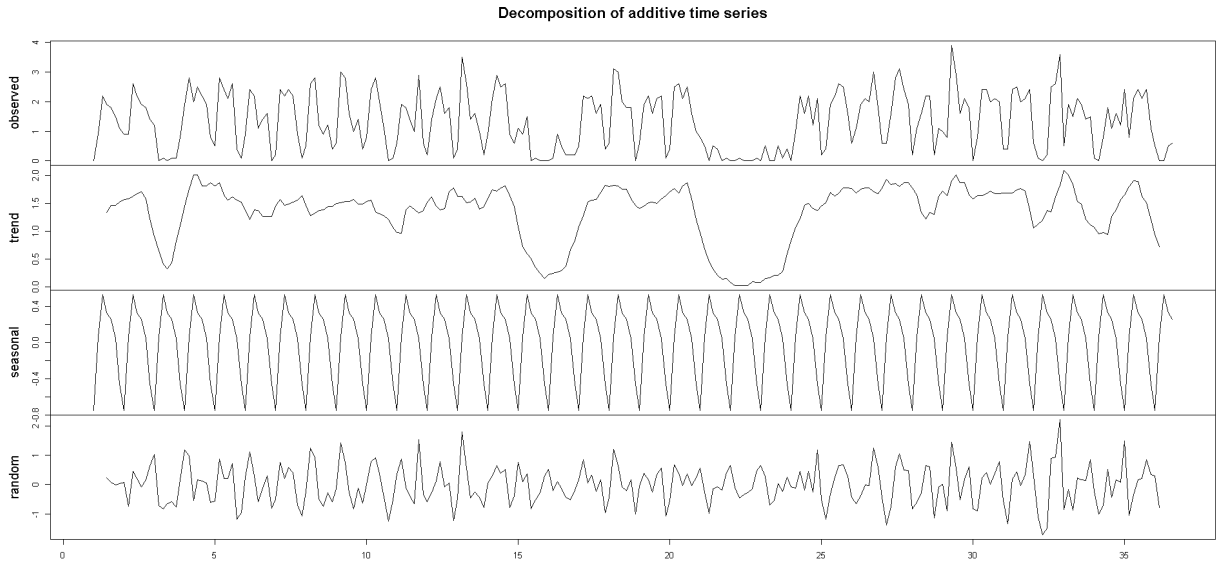


图5 中等小组平均访问次数时序分解

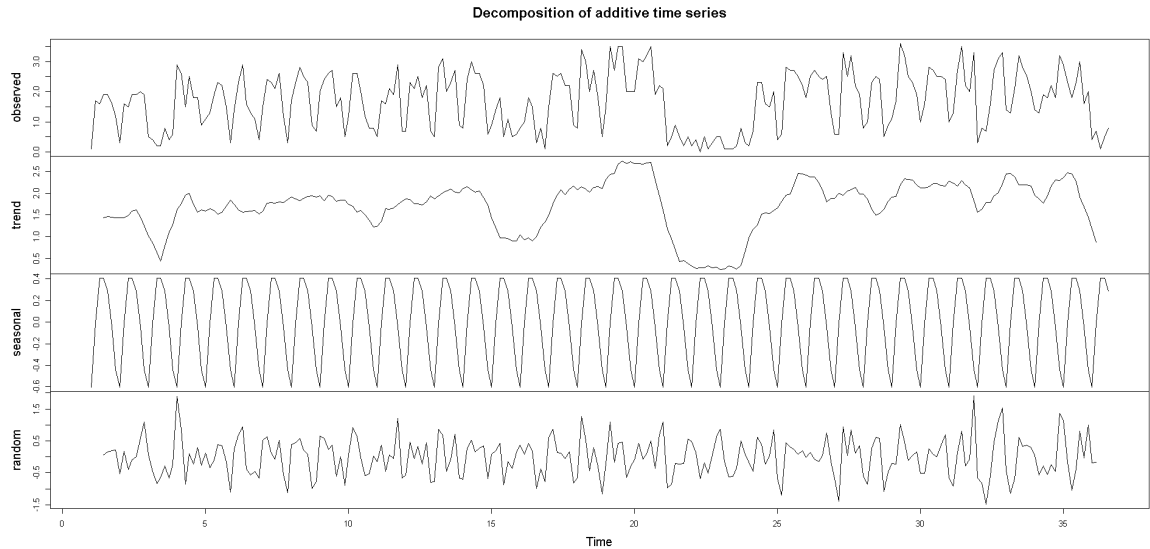


图6 优秀小组平均访问次数时序分解

下的访问次数的关联度。将专业中每个课程的访问次数分为三组，分组采用 hierarchy 分类算法，分组结果分别用“高”“中”“低”来表示，三个分组结果分别代表该小组学生访问次数排序，即小组标记为“高”的成员，在该课程的访问次数处于最高的一组，其他小组依次类推。因为篇幅所限，仅挑选 2016 级生物医学专业学生作为样本分析。其学生在课程中访问分组结果部分数据如表 5 所示。然后使用 Apriori 关联算法对学生在不同课程中的行为进行关联分析，因篇幅所限，仅列举具有代表性的分析结果，如图 8 所示。

以课程 1 为例，12 位学生表现出中等活跃度，在课程 2、课程 4、课程 5 中也表现出中等活跃程度，其中 11 位学生在课程 3 中表现出中等活跃程度。在课程 1 中，九位学生表现出高度积极性，在课程 3 中也表现出高度积极性；其中七位学生在课程 4 中表现出高度活跃性，五位学生在

表 5 学生在课程中访问分组结果

	课程 1	课程 2	课程 3	课程 4	课程 5
学生 1	低	低	低	低	低
学生 2	中	中	中	中	中
学生 3	高	高	高	高	中
...					

课程 5 中表现出高度积极性。研究结果显示，学生在单个课程中的表现和其他课程中的表现存在密切相关性，即：学生在某门课中表现积极，很有可能在其他课程中也表现积极；反之，学生在某门课程中的活跃程度较低，在其他课程中的活跃度也较低。其中活跃程度较低的数据可能暗示学生对于所学专业不喜欢，存在厌学倾向。

5 结语

在中外合作办学环境下，学生在系统的投入程度和成

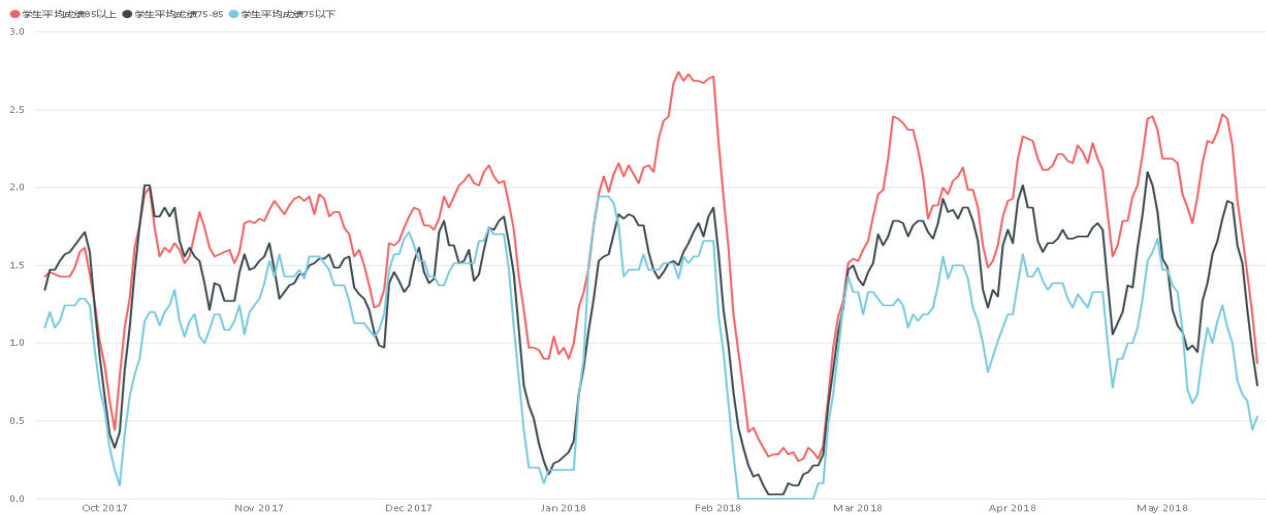


图7 不同分组中趋势分解 (Trend) 对比

课程1=中	12	54.55
课程2=中	12	54.55
课程3=中	11	50
课程4=中	12	54.55
课程5=中	12	54.55
课程1=高	9	40.91
课程2=高	6	27.27
课程3=高	6	27.27
课程4=高	6	27.27
课程5=高	6	27.27
课程1=中	9	40.91
课程2=中	7	31.82
课程3=中	5	22.73
课程4=中	7	31.82
课程5=中	5	22.73
课程2=中	15	68.18
课程3=中	11	50
课程4=中	14	63.64
课程5=中	14	63.64
课程2=高	6	27.27
课程3=高	6	27.27
课程4=高	6	27.27
课程5=高	6	27.27
课程3=中	11	50
课程3=高	10	45.45
课程4=中	14	63.64
课程4=高	7	31.82
课程5=中	16	72.73
课程5=高	5	22.73

图8 系统活动高的学生不同课程中活跃程度关联分析

绩整体上存在相当程度的正相关,这种正相关在一定程度上反映出国外高等教育机构重视学习管理系统建设的原因。本研究也证明学生的投入程度表现出时序性,即不同成绩类型的学生在学年的不同阶段表现出不一样的行为特性,该现象可以在学期中期预测学生成绩走势和实施必要的学业辅导。研究结果也显示,教师对于系统的使用程度对学生的访问程度有很大的促进作用,在系统培训过程中可以利用该结论来鼓励教师在系统中和学生积极互动。同时,针对之前主流的在线学习研究比较关注学生在某个课程中的行为,忽略了学生不同课程间的表现,本研究对这一现象进行补充,分析结果显示,学生不同课程中的投入程度表现出相似性,该相似性可以用来作为分析学生对于自己所学专业满意度参考。

综上,在中外合作办学环境下,学习管理系统对整个教学环境存在积极的促进作用。随着国内中外合作办学机

构的快速增加,相关办学机构可考虑借助学习管理系统系统,学习外方合作院校如何在系统中进行课程设计和开展教学活动,利用系统用户数据优化教学过程和课程设计,从而缩短和国外先进高校教学上的差距。■

参考文献

[1]Looi C K, Seow P, ZHANG B, et al. Leveraging mobile technology for sustainable seamless learning: A research agenda[J].British Journal of Educational Technology,2010(2):154-169.

[2]Sinatra G M, Heddy B C, Lombardi D. The challenges of defining and measuring student engagement in science[J].Educational Psychologist,2015(1):1-13.

[3]Macfadven L P, Dawson S. Mining LMS data to develop an “early warning system” for educators a proof of concept[J].Computer & Education,2010,54(2):588-599.

[4]李爽,钟瑶,喻忱,等.基于行为序列分析对在线学习参与模式的探索[J].中国电化教育,2017(3):88-95.

[5]Siemens G. What are learning analytics[EB/OL].<http://www.elearnspace.org/blog/2010/08/25/what-are-learning-analytics>.

[6]姜强,赵蔚,王朋娇,等.基于大数据的个性化自适应在线学习分析模型及实现[J].中国电化教育,2015(1):85-92.

[7]李爽,王增贤,等.在线学习行为投入分析框架与测量指标研究:基于LMS数据的学习分析[J].开放教育研究,2016(2):77-88.

[8]Nespereira C G, Dai K, Redondo R P, et al. Is the LMS access frequency a sign of students’ success in face-to-face higher education?[M]//Proceedings of the Second International Conference on Technological Ecosystems for Enhancing Multiculturality. Salamanca, Spain. 2014.