

# 基于 LSA 的在线教育行为模式可视化分析

魏莉<sup>1\*</sup> 韩艳辉<sup>2</sup>

(1. 山西广播电视大学 山西省太原市 030027 2. 国家开放大学 北京市 100039)

**摘要:** 本文研究针对基于国家开放大学学习平台, 选择《媒体辅助英语教学》课程的教师和学生的在线行为数据作为研究对象, 将学习过程划分为不同阶段, 对比各阶段教师和学生的在线行为模式的变化, 分析教师和学生的在线行为相关性, 用滞后序列分析法对教师和学生的在线行为模式进行实证探索, 通过 Gephi 工具将分析结果可视化。研究发现教师依据学生的行为模式特点来提供个性化学习支持服务, 有效提高了学生的积极性, 提升了学习效果。

**关键词:** LSA; 滞后序列分析法; 相关性分析

在线教育盛行, 信息数据量正呈几何级指数增长, 通过挖掘和分析在线教育数据, 可以获得其背后隐藏的巨大价值。在线教育是在“互联网+”的大背景下产生的, 在线教育数据成为新的教学要素, 在线学习行为成为分析学生效果的核心资产。国家开放大学基于 Moodle 的学习平台能够记录每个学生的学习行为, 为分析学生的学习状态提供数据基础, 结合数据分析技术, 有助于教师发现学习者的学习行为模式, 预测学习者的表现和风险, 及时有效的干预调整, 使得教师的个性化教学成为可能。

LSA (Lag Sequential Analysis, 滞后序列分析法) 是学习行为模式分析的重要方法之一, 是由 Sackett 于 1978 年提出的, 可以用于检验某一行为发生之后出现另外一种行为的概率在统计意义上的显著性<sup>[1]</sup>。运用 LSA 可以帮助教师把握学生潜在的行为模式, 从行为视角增强学习效果, 有效指导后续教与学活动的设计与实施。GSEQ (General Sequential Querier) 软件可以用于分析在线教育中教师和学生在在线行为序列模式, 提高数据分析处理的效率。

近年来, LSA 被广泛应用于知识建构行为模式分析、学生行为模式分析、师生交互的话语分析、学习社区信息交互分析、学习过程分析与学习效果预测, 线下课堂互动效果分析等<sup>[2-9]</sup>。可是, 整体上结合教师和学生的在线行为模式分析的实证研究较少。学生和教师是在线教育的重要角色, “教”与“学”的行为数据为分析师生互动行为模式提供了有价值的信息。

本研究选取国家开放大学开设的《媒体辅助英语教学》在线课程作为研究对象, 该课程在教学平台上组织实施教学活动, 引导学习者自主学习; 不设终结性考试, 采用百分之百形成性考核的方式, 教师和学生的在线行为过程完整的记录下来, 使得行为日志数据能够较客观、准确地反映学习者的学习状态。

## 1 数据收集与分析

本研究采用 LSA 和相关性分析法对教师和学生的行为频次和行为转换频次进行统计分析, 对教师和学生的在线行为模式进行可视化分析。

### 1.1 研究对象

本研究对象是《媒体辅助英语教学》2018 年秋季学期在线课程, 课程团队将整个在线学习活动划分为七个模块, 分别是一、学习指南和破冰行动; 二、媒体辅助中学英语教学的基本原理介绍; 三、常用办公软件的英语教学; 四、应用语法和词汇训练; 五、阅读训练; 六、听说训练; 七、写作训练<sup>[10]</sup>。核心教师 13 名, 学习者共计 80 名, 课程总时长约十周, 全体师生的在线行为记录 106176 条, 其中教师行为记录 50502 条, 学习者行为记录 55674 条。

表 1: 教师高频在线行为编码

序号	教师高频行为类别	行为频次	编码
1	浏览	22707	TL
2	讨论区	创建	4636
3		上传	3548
4		更新	452
5		浏览	5196
6	互动评价	评价	3337
7		重新评价	223
8	系统	浏览	14607
9		成绩	2162
10		更新	1448

表 2: 学生高频在线行为编码

序号	学生高频行为类别	行为频次	编码
1	浏览	5481	TL
2	讨论区	创建	1278
3		上传	917
4		浏览	3921
5	互动评价	上传	479
6		创建	306
7	文件、文件夹	浏览	3924
8		系统	2762
9	网页	881	WL
10	成绩	371	SL

分析学习者每个模块的在线行为的频次、类型和路径等, 能够帮助教师清晰准确的了解学生的个性化特征, 及时调整教学策略。课程团队设置了“光荣榜”, 激发学生的参与热情, 师生形成良好的在线互动。

### 1.2 数据筛选、编码

在线行为场所主要包括课程页面、讨论区、成绩、作业、学习文件、互动评价和调查问卷等 10 处。每种场所又有多种行为方式, 包括浏览、上传、更新、创建和提交等。将在线行为场所和行为方式组合起来分析, 共有 32 类行为。统计各类行为数据, 筛选频次前 10 的教师和学生的行为并对其进行编码, 表 1 为教师高频在线行为编码、表 2 为学生高频在线行为编码。

### 1.3 数据分析

● 基金项目: 1. 国家开放大学 2016 年度青年课题“Moodle 平台上教学和学习行为的大数据分析——个性化教育实现探索”(G16A0403Q); 2. 山西广播电视大学 2019 年度校级课题“人工智能在现代远程教育中的应用探索”(SXDDKT 201913)。

表 3: 残差表 (部分)

Given:	TL	TS	TC	TG	XL
TL	31.51	14.79	-36.84	7.25	-11.98
TS	-28.42	-9.01	94.23	1.22	-3.99
TC	16.85	-9.64	17.39	-3.01	-4.38
TG	0.44	6.77	-2.99	-0.83	-1.22

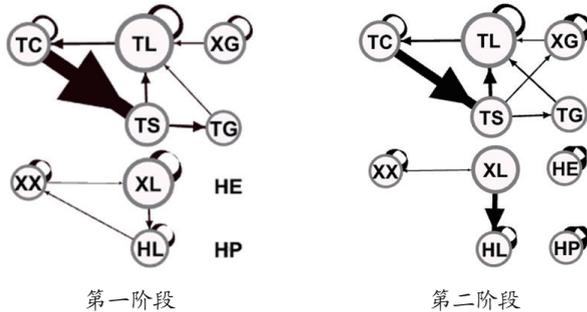


图 1: 不同阶段教师的在线行为转换图

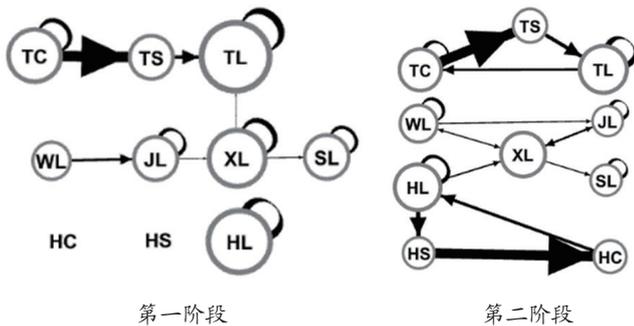


图 2: 不同阶段学生的在线行为转换图

分析表 1、表 2 的高频行为场所主要分布在讨论区、互动评价、系统、文件等。对比表 1 和表 2 的高频行为可以看出, 行为场所高度重合, 教师的行为频次普遍高于学生, 分析各类行为中教师的行为频次和学生的行为频次的相关性, 皮尔逊相关性系数达到 0.835, 可见一致性很高。教师在讨论区积极引导学生参与形成良性互动, 互动评价方面, 课程中设置学生之间互评, 参与度高, 互动评价任务极大激发了学生的学习动力, 增强了学生的评价能力和互相学习的能力, 促进了他们的个性发展和全面提高。

## 2 核心行为序列分析

本研究采用 LSA 和信息数据可视化开源工具 Gephi 对教师和学生行为模式进行可视化分析。研究过程可分为两个部分。第一, 分析整个学习过程中“教”和“学”行为在不同学习阶段的行为模式; 第二, 分析教师和学生行为模式的相关性。

### 2.1 基于LSA的行为模式分析

将整个课程的七个教学模块分为两个阶段, 阶段一包括该课程的模块一、二、三和四, 阶段二包括模块五、六和七。

行为分析流程如下:

(1) 收集在线行为日志数据, 选取时间、姓名、行为类型等特征信息, 按照行为发生的时间生成教师和学生核心行为编码序列;

(2) 使用 GSEQ5 软件, 按照格式要求输入三个阶段的行为编码序列。下文为学生行为转换编码序列的部分序列:

Event

(\$ Behavior=SL HL HS HC TL TS TC WL JL XL);

%LC#lcid 01

HL	HL	HL	HL	HC	HS	HL
HL	HL	HL	HC	HS	HL	TL
TS	TC	TC;				

%LC#lcid 02

XL	JL	JL	JL	XL	JL	TL
TL	XL	XL	JL	JL	JL	XL
XL	XL	JL	JL	JL	WL	XL
JL	XL	XL	JL	JL	HL;	

.....

%LC#lcid 465

HC HS HL TL TS/

(3) 分析行为序列后得到频率转换表及残差分析表, 频率转换次数表体现了某一个行为之后发生另外一种行为的频率, 残差表用于体现两个行为之间的前后转换关系的显著性。

$$z - score = \frac{\sigma - \bar{\sigma}}{\sqrt{\frac{1}{N} \sum_{i=1}^N (\sigma_i - \bar{\sigma})^2}}$$

当残差标准值 z-score 大于 1.96 时, 表示行为转换频次达到统计上的显著性水平, 见表 3, 以第一阶段教师行为转换的残差表(部分)为例, 表 3 中 TL 行为之后发生 TL 行为的调整后的残差值为 31.51。

(4) 筛选意义显著的行为序列, 根据调整后的残差值表, 筛选残差表中 z-score>1.96 的数据, 表 3 中“TL → TL”、“TS → TL”、“TG → TL”、“TC → TS”、“TL → TC”、“TC → TC”、“TS → TG”都是显著的行为转换关系。连续发生同一行为表示该行为的稳定性高。对筛选后的数据运用 Gephi0.9.2 软件生成行为序列有向路径转换图, 可更直观地呈现各种行为之间的转换, 如图 1、图 2 所示。设定行为编码为图的节点标签, 边的信息包括源节点 (source)、目标节点 (target), 权重 (weight), 箭头连线表示了行为发生的顺序。节点直径越大表示该行为发生的频率越高, 边越粗表示权重越大, 连接两个行为之间的关系越强, 即发生上一行为后发生下一行为的概率越高。

### 2.2 结果分析

如图 1 所示, 对比教师在不同阶段的在线行为转换图, 第一阶段中有两个孤立节点 HE 和 HP, 边数为 16; 第二阶段中无孤立节点, 边数为 22。第一阶段的平均度数 1.6, 第二阶段平均度数为 2.2, 高于理想顺序状态下的平均度数 2, 表示教师的行为种类较为丰富, 路径选择较为丰富。图中有很多节点的自环现象, 自环的边越粗自环现象越明显, 表明该行为具有较高的稳定性, 如 TL、TC、XG、XX、XL、HL、HP 等, 可见讨论区和互动评价等场所的转换行为频繁而稳定。

如图 2 所示, 对比学生不同阶段的在线行为转换图, 第一阶段中有两个孤立节点 HC 和 HS, 边数为 13; 第二阶段中无孤立节点, 边数为 22。第一阶段的平均度数 1.3, 第二阶段平均度数为 2.2, 超过高于理想顺序状态下的平均度数 2, 表示学生的行为种类和路径选择较为丰富, 提升明显。第一阶段中, “TC → TS → TL”行为转换显著, 第二阶段中“TC → TS → TL”形成闭合循环路径, 可见学生在讨论区的行为频繁且稳定性高。第二阶段中“HC → TL → TS”行为转换显著且形成闭合循环路径, 可见学生参与互动评价的积极性高。

分析教师和学生两个阶段的行为转换残差值, 教师在各阶段的行为转换中概率前 5 依次为“HP → HP”、“HE → HE”、“XG → XG”、

# 一种码速率自适应的复用型编码调制器

周少骞 罗小成 田运通 林闽佳  
(上海航天电子技术研究所 上海市 201109)

**摘要:** 本文介绍了一种码速率自适应的复用型编码调制器,该编码调制器可以适应 1Kbps-650Mbps 范围的信息传输速率,可同时并行处理发射多个通道(上限取决于单机能容纳的载波信道数量)的信源数据,可同时进行星间、星地数传链路的信号传输发射。

**关键词:** 数传发射机; 编码调制器; FPGA; LVDS 接口; (4, 3, 7) 卷积编码; 8/7LDPC 编码; 射频调制

## 1 引言

星载数传发射机是一种应用于空间飞行器,将信源获取的基带数据调制在载波信号上进行卫星与卫星之间或卫星与地面之间数据传送的设备。

数传发射机原理框图如图 1 所示,数传发射机主要由编码调制器、功率放大器、数传天线组成。

作为数传发射机的关键组成部分,编码调制器的功能是对输入的信源数据进行信道编码处理,提高信道传输增益,增强信号传输抗干扰能力。在对数传信道线性度要求较高的应用场景,编码调制器中还需要进行信号预失真处理,以弥补信号功率放大电路带来的非线性影响。

根据应用场景的不同,编码调制器接收从有效载荷、固态存储

器、信号处理机或加密机等产品输出的基带数据(统称信源数据);在可编程器件中对信源数据进行信道编码处理;编码调制器中的载波生成电路产生调制所需的 X 波段、Ka 波段或其它波段的单载波信号;载波信号和编码数据在直接射频调制器中进行调制,调制信号输出至功率放大器。

功率放大器是数传发射机的重要组成部分,主要包含固态功率放大器和行波管放大器两种类型的放大器,功能是对编码调制器输出的调制信号进行功率放大和带通滤波,可根据总体需求增加三次谐波抑制滤波器。

天线是数传发射机的重要组成部分,一般由天线主体、固定机构、展开机构或转动机构组成,功能是将功率放大器输出的信号按一定的波束指向和增益对外发射。

“TC → TS”和“HP → HP”。学生在各阶段的行为转换中概率前 5 依次为“TC → TS”、“HS → HC”、“TL → TL”、“HL → HS”和“TC → TC”,可见教师和学生的在线行为转换具有较强的相似性。将教师和学生的“TC → TL”和“TC → TC”两种行为转换分别做双变量相关性分析,结果显示相关性在 0.01 水平上显著正相关,皮尔逊相关性系数分别为 0.891 和 0.873。可见教师和学生在讨论区频繁交流,可谓同频共振。

## 3 结论

本研究根据国家开放大学 Moodle 学习平台上的教师和学生的在线学习行为数据进行滞后序列分析,并通过 Gephi 将行为模式可视化,使得教师能够更直观的了解学习过程中的行为模式变化。

研究发现,学生在讨论区和互动评价区的行为转换频率较高,而在学习文件、系统等场所的行为转换频率相对较低。在该课程的在线教学过程中,教师针对学生的行为特征在讨论区和互动评价区与学生充分交流,极大调动了学生的积极性,此外教师还需要发挥线上学习资源在知识呈现方面的优势,优化知识构建,基于学生的需求提供更合适的学习资源提升教学的质量。

在线教育行为模式分析表明,学生的在线学习需要教师精心安排教学活动并不断引导。滞后序列分析方法可以帮助教师开展学习路径的设计与优化,可视化学生的学习路径,还可结合聚类等方法用于分析学生的行为模式与学习成绩的相关性分析,以帮助教师依据在线行为序列的分析结果,及时调整教学策略,做好个性化学习支持服务,促进学生的参与,提高学习的效果。

## 参考文献

[1] 杨现民,王怀波,李冀红. 滞后序列分析法在学习行为分析中的应用[J]. 中国电化教育, 2016(2): 17-23.

[2] 魏顺平,韩艳辉,王丽娜. 基于学习过程数据挖掘与分析的在线教学反思研究[J]. 现代教育技术, 2015(6): 89-95.

[3] Yanhui Han, Shunping Wei and Shaogang Zhang. Analysis of Online Learning Behaviour from a Tutor Perspective: Reflections on Interactive Teaching and Learning in the Big Data Era[J]. AAOU JOURNAL, 2015(2): 29-48.

[4] 李爽,钟瑶喻忱,程罡,魏顺平. 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): 17-23.

[5] 江波,高明等. 基于行为序列的学习过程分析与学习效果预测[J]. 技术应用, 2018(02): 103-112.

[6] 刘聪聪,戴心来. 基于 LSA 与 SNA 的虚拟学习社区信息交互实证研究[J]. 软件导刊, 2018(11): 137-140.

[7] 孙萌,唐雪萍,郑勤华. 基于日行为模式的学生行为序列分析[J]. 开放学习研究, 2019(2): 39-45.

[8] 冷静,黄旦. 基于 LSA 的大学生批判性思维在线话语分析[J]. 中国电化教育, 2019(4): 70-78.

[9] 江毅,王炜,康苗苗. 基于行为序列分析的师生互动效果研究[J]. 现代远程教育, 2019(6): 55-63.

[10] 魏莉,韩艳辉等. 基于 Moodle 平台“教”和“学”行为的个性化分析[J]. 山西广播电视大学学报, 2019(4): 21-25.

## 作者简介

魏莉(1984-) (通讯作者),女,河北省邢台市人。山西广播电视大学,讲师。研究方向为学习分析技术、中文信息处理。  
韩艳辉(1972-),男,河北省沧州市人。国家开放大学,副教授。研究方向为网络化计算机辅助语言学习、数字技术与在线教育的整合。