

远程教育机构在线教学活跃指数研究

魏顺平

(国家开放大学,北京 100039)

摘要:用于科研工作人员绩效评价的h指数同样可以适用于远程教育。远程教育机构在线教学活跃程度可以用w指数来表示。在提出w指数定义及算法后,以国家开放大学学习网2016年秋季1.1亿在线教学日志为样本,计算国开44所分部的w指数,并对达到w指数的人的比例加以约束,使得这个比例不能低于40%,也不能高于60%,此时得到的w指数与师生数、人均行为次数均能达到一个中等相关的水平,表明此时的w指数能同时兼顾数量和质量,这也达到了预期的目标。

关键词:远程教育机构;在线教学;w指数;活跃指数;学习分析;教育大数据

中图分类号:G434 **文献标识码:**A **文章编号:**1008-3006(2017)04-0004-06

一、引言

中国互联网络信息中心(CNNIC)发布的第39次《中国互联网络发展状况统计报告》显示,截至2016年底中国在线教育用户规模达1.38亿^[1]。《国家教育事业“十三五”规划》对“教育质量”非常关注,并提出“制定在线教育和数字教育资源质量标准”“制定在线开放课程教学质量评价标准”,可见国家对在线教育质量的重视^[2]。在我国,在线教育主要由开放大学、广播电视大学以及普通高等教育的网络教育学院等机构供给,远程教育机构每个学期教学工作的成功开展,是在线教育质量保证的关键。我们不仅要远程教育机构教学效果进行评价,更要对其教学过程进行实时监控与评价。

然而,在过去以及当前一段时期内,我们必须面对在线教育低投入率这一现实并逐步改变这个现实。梅耶(Meyer)报告称,在斯坦福大学、麻省理工学院、加州大学伯克利分校提供的慕课课程中,学生退出率高达80%~95%^[3]。在edX开设的1门电子电路慕课课程中,15万余人报名学习,最终7000余人获得证书,通过率为4.6%。在Coursera的1门社会网络分析课程中,也有类似情况:只有2%的学习者获得了基本

的证书,仅有0.17%的学习者获得了高级别的证书。樊超和宗利永基于MOOC学习平台的“学堂在线”(www.xuetangx.com)1300万师生行为日志,分析得出绝大多数学生选课行为的持续天数都非常短,95.1%学习者选课后的上课天数不超过10天,学习者的退出率现象比较严重,大多数用户并没有坚持学习完成全部课程内容^[4]。

面对在线教育迅速发展的大好形势以及学习者低投入、高退出率的现状,我们要对远程教育机构在线教学过程进行监控和评价,并充分发挥评价的导向功能,引导远程教育机构朝着理想目标前进。现阶段的评价导向应该是促使在线教学发生,提升师生在线教学活跃度,研制出一种简单易行的远程教育机构在线教学活跃度评价工具显得很有必要。

二、文献综述

用户活跃度一般是指用户使用一项功能或服务的频繁程度,一般用于衡量一个网站或一个移动应用的用户活跃程度。常用的公式如下:

用户活跃度=日活跃用户/月活跃用户

对于一个网站或一个移动应用来说,用户活跃度

基金项目:2017年度教育部在线教育研究中心在线教育基金(全通教育)一般项目“基于职教大数据的学习分析系统开发与应用研究”(课题编号:2017YB120)。

收稿日期:2017-10-11

作者简介:魏顺平(1981-),男,博士,国家开放大学信息化部(工程中心),副研究员。

的理论值不低于0.2,即用户使用网站或移动应用的天数不少于6天。

对于远程教育机构而言,其师生活跃度的计算要放在一个学期的周期内,并不是以月为单位;并且并不要求学生天天学习,而是在一个学期内选择自己合适的时间学习。因此,采用常规的用户活跃度来衡量远程教育机构的师生活跃度并不恰当。一般而言,师生活跃度要基于师生行为日志计算得到。师生行为一般以一个行为动词为中心,辅之以行为主体、行为客体、发生场所、发生时间以及行为产生的结果等语义关系。一些研究者基于这一“教与学行为”要素框架提出了师生活跃度评价模型。

2017年,贾积有和于悦洋提出了一个描绘学生学习活动的指数LAI(Learning Activity Index)及其在线形式——在线学习活动指数OLAI(Online Learning Activity Index),从速度(speed)、质量(quality)和数量(quantity)三个维度进行综合评判并进行数字化处理,得到3个独立的无量纲指标,而学习活动指数LAI就等于这3个维度指标之和^[5]。郑勤华等人对基于数据的在线教学测评展开了较为系统的研究,以学生综合评价为目标,通过理论演绎和专家访谈构建了以投入度、完成度、调控度、联通度和主动性为核心的五维度综合评价参考理论模型,通过学习行为数据聚合特征变量,构建了相应的计算模型^[6]。魏顺平、程罡采用演绎和归纳相结合的办法构建了“基于数据驱动的教育机构在线教学过程评价指标体系”,主要从“广度”、“深度”、“持续度”(或“及时度”)等维度进行评价^[7]。

已有在线学习评价模型主要适用于个体在线学习评价。当用于机构评价时,累加个体评价得分往往是困难而复杂的。因此,有必要找到一种更简便的方法。一般来说,“参与学与教的师生数”和“人均行为数”常用来评价机构表现。以国家开放大学为例,在每年的网上教学检查中,一般采用求和、求平均等两种聚合函数计算“实际参与教学的师生比例”和“师生人均有效行为数”两个指标来评价国开各分部的网上教学活跃程度,第一个指标注重评价活跃数量,第二个指标注重评价活跃质量。采用两个指标的弊端在于规模较小的分部,由于师生基数小,无论是师生占比还是人均行为数都容易排在前面,这对于规模较大的分部有失公平;活跃数量和活跃质量的分开考察,容易给被评价机构造成的导向是,要求大部分师生登录学习平台,1次即可;同时组织少数师生大量点击,

如此一来,也可以实现两个指标同时得高分的结果。因此,迫切需要提出一个同时兼顾数量和质量的评价指标,本文称之为机构在线教学活跃指数。

我们可以从科研工作者的绩效评价得到一点启发。过去,科研工作者的绩效评价一般采用学者论文发表数量和论文的总被引频次,前者着眼于量的分析,后者着眼于“质”的考量。针对“论文数量”的缺点是不能测度论文的重要性和影响力、“总被引频次”的缺点是作者的影响力会被少数重要的合作论文夸大等缺陷,美国加州大学圣迭哥分校的物理学家J.E. Hirsch教授在2005年提出了一项旨在评价科学家个人绩效的指标——h指数。h指数的定义是:一个科学家的分值为h,当且仅当在他(她)发表的 N_p 篇论文中有h篇论文每篇获得了不少于h次的引文数,科学家剩下的 (N_p-h) 篇论文中每篇论文的引文数都小于h次^[8]。使用h指数可以让被引频次、论文数量两个指标相互制衡:只有当论文数量多,而且多数论文具有高被引频次时,才能获得高h指数^[9]。

我们发现,采用h指数评价科研工作者科研绩效所要达到的导向目标即“论文数量多,而且多数论文具有高被引频次”,非常类似于我们评价教育机构在线教学活跃度所想达到的目标即“参与师生数量较多,而且多数师生产生了较高的行为次数”。因此,我们可以借用h指数的思路来设计一个教育机构在线教学活跃指数。

三、研究过程与讨论

本文选取国家开放大学这一大型在线教育机构作为研究对象,评价该大学44个分部的在线教学情况。目前国家开放大学学习平台依托云服务平台采用集中式部署,全国学生使用统一的学习平台——国开学习网(www.ouchn.cn,改造开源学习管理系统Moodle得来),所以国家开放大学学习网中生产的师生行为数据基本上就是国家开放大学所有师生的在线教与学数据。本文选取国开学习网2016年秋季学期的数据作为研究样本,这一学期国家开放大学44所分部共计开设万门课程,1.4万余名教师和62万余名学生参与在线教与学,产生了1.1亿条行为日志。

(一)机构活跃指数定义与算法

参照h指数的命名方法及定义,作者将机构活跃指数命名为w指数(取作者姓氏拼音的首字母),并定义为:1所远程教育机构的在线教学活跃指数为w,当

且仅当在它的N个用户中有w²个用户的行为次数不少于w次,机构剩下的(N-w²)个用户的行为数都小于w次。

具体算法是:计算一个学期中1所远程教育机构学习平台中所有师生的行为次数,然后将师生用户及对应的行为次数按行为次数逆序排列并从1开始标上序号;然后从第1名用户开始,比较每个用户的序号开方与行为次数的大小,直至某个用户的序号开方大于对应的行为次数,假设此时的序号为n,那么w指数可用以下公式 $w = \sqrt{n-1}$ 算出,其计算过程见图1。

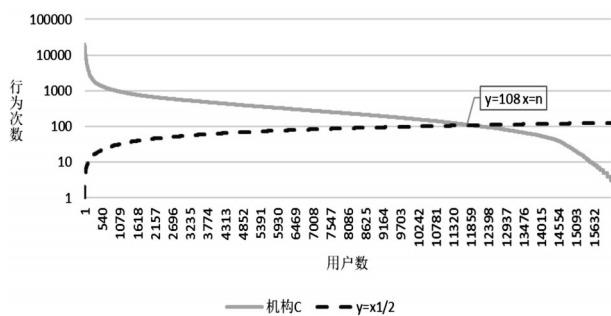


图1 w指数计算过程图示

如图1所示,实线是机构C中用户数及对应的行为次数分布曲线,虚线是幂函数 $y=x^{1/2}$ 的曲线,此时机构C的w指数就是两曲线交点的y值。

正如h指数能够综合反映1位研究人员的发文量和被引次数,w指数则能够综合反映1所远程教育机构在一个学期中在线教学的参与师生数及师生行为次数。参与学与教的师生数量(广度)和师生学与教行为次数(深度)两个指标互相制衡:参与师生很多,但师生行为次数不高的机构,w指数一定不高,如图2

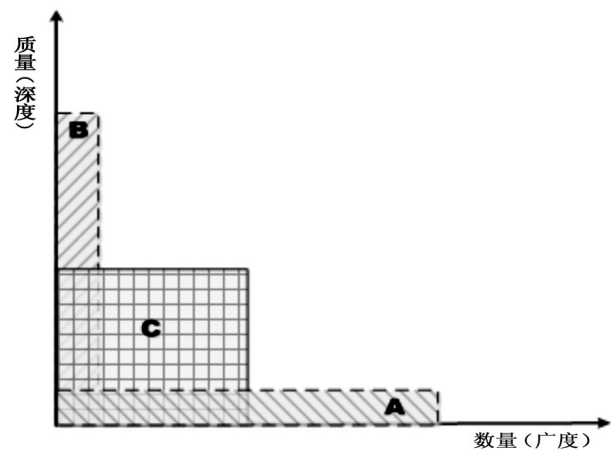


图2 w指数示意图

区域A所示;师生行为次数很高,参与师生数量不多的机构,w指数也会被参与师生数量所限,而不能得到高w值,如图2区域B所示;只有当参与师生数量多,而且多数师生具有高行为次数时,才能获得高w指数,如图2区域C所示。可见,用w指数评价机构教学绩效可以遏制片面追求参与师生数量的不良倾向,同时又能够激发机构探索深层次教学的热情。这就是w指数与其他单项机构教学评价指标相比所独具的高明和绝妙之处。

(二)w指数应用

将w指数用于国家开放大学各分部在线教学绩效评价,计算出国家开放大学44所分部的w指数。

1.初步应用

如表1所示,是按照前文给出的w指数算法得出的初始w指数计算结果,记为w₀,并按w₀逆序排列呈现。同时列出“参与师生数”“人均行为数”,并给出行为次数不低于w₀的师生数占比。

表1 初始w指数计算结果(部分)

分部代码	参与师生数	人均行为数	w ₀	行为次数高于w ₀ 的师生占比
A0U	28773	375.42	129	57.84%
B5I	37043	186.78	128	44.23%
E4U	24831	577.12	128	65.98%
N6N	41306	155.49	125	37.83%
C1N	36636	259.79	117	37.36%
N7N	16167	403.08	108	72.15%
N5G	21577	225.58	100	46.35%
A2G	30253	95.39	94	29.21%
I2G	23278	116.94	94	37.96%
N3N	27715	102.87	89	28.58%
N0U	14738	184.28	87	51.36%
O3G	19199	114.56	87	39.42%
A8U	17427	101.28	86	42.44%
H9I	21062	98.12	86	35.12%

从 w_0 排序来看,基本上排在前面的都是拥有较多师生数的分部,一些师生数较少但是人均行为次数较高的分部如厦门分部,却由于用户少处在垫底的位置。计算 w_0 位序、参与师生数位序、人均行为次数位序3种排序之间的相关性,结果如表2所示。

表2 相关系数计算

	师生数位序	人均行为次数位序
w_0 位序	0.93	0.21

从中可以发现,此时的 w 指数主要反映了参与师生数量,而不能反映参与质量。问题就出在达到 w 指数的人数占全部师生数的比例上。

在这个比例上,最高为是A8N分部,83.41%;最低是E7G分部,27.26%,平均值51.98%,标准差15.86%,变异系数30.51%,可以看出个体间的差异已经达到了一个很高的值。因此,在计算 w 指数时,要对达到 w 指数的人的比例加以约束,使得这个比例不能低于某个最低值,也不能高于某个最高值。如果不设下限,就会让一些师生人数多但是总体表现不佳的分部“占便宜”;如果不设上限,就会让一些师生人数少但是总体表现良好的分部“吃亏”;如果不设区间,而是指定某一位置,如50%的位置,这就与计算中位数、四分位数无异,并且限制了规模不一的机构调整的空间。

2.算法修正

首先还是计算出 w_0 这个值,然后判断“行为次数不低于 w_0 的师生数”即 w_0^2 个师生占全部参与师生数 N 的比例。此时存在3种情况,并对应3种算法,得到修正后的 w 指数,记为 w_1 。

表3 w 指数算法修正

条件	算法
$0.4 \leq w_0^2/N \leq 0.6$	$w_1 = w_0$
$w_0^2/N < 0.4$	$w_1 =$ 处在40%位置的用户的行为次数
$w_0^2/N > 0.6$	$w_1 =$ 处在60%位置的用户的行为次数

其次,按照上述的相关条件和算法,我们得到了40%和60%的用户行为次数,还需要进行修正后得到

以下的结果,如表4所示。

表4 修正后的 w 指数(部分)

分部代码	参与师生数	人均行为数	W_0	W_1	行为次数高于 W_1 的师生比例	W_0 排序	W_1 排序
N7N	16167	403.08	108	176	60.00%	6	1
E4U	24831	577.12	128	171	60.00%	2	2
AOU	28773	375.42	129	129	57.84%	1	3
B5I	37043	186.78	128	128	44.23%	2	4
N6N	41306	155.49	125	120	40.00%	4	5
C1N	36636	259.79	117	110	40.00%	5	6

此时发现,前后两个 w 指数排名已经有了一定的变化,一些师生规模较小但人均行为频次较高的分部排名有了大幅提升;一些师生规模较大但人均行为频次较低的分部排名则大幅下降。接着,我们来检验一下调整后的 w 指数与师生规模、人均行为次数之间的关系,相关系数计算结果如表5所示。

由表5可知,调整后的 w_1 位序与师生数位序、人均行为次数位序均能达到一个中等相关的水平,表明此时的 w 指数能同时兼顾数量和质量,这也达到了我们预期的目标。

表5 相关系数计算

	师生数位序	人均行为次数位序
w_1 位序	0.59	0.47

(三) w 指数拓展

1.机构质量维度的拓展

对于一个机构而言,在计算 w 指数时,除了可以将每个用户的行为次数作为质量维度外,还可以将每个用户的投入时间(天数、小时数)、使用的资源模块和活动模块数等作为质量维度。当然在选用不同的质量维度前需要比较不同质量维度取值的个体差异性。一般而言,个体差异性越大,越适合做质量维度。如表6所示,是三个质量维度的个性差异比较。

表6 三个质量维度的个体差异比较

统计项	样本数	最大值	最小值	平均值	标准差
行为次数	650642	56395	1	172.9792	471.7529
使用模块数	642527	881	1	7.2743	13.98497
投入天数	642527	106	1	3.076	4.148462

从标准差来看,在质量维度选取上,“行为次数”最佳,其次是“使用模块数”,再次是“投入天数”。

2.评价对象的拓展

W指数还可以拓展到其他的评价对象如课程、教师、学生等,从而构建课程实施活跃度指数、教师在线教学活跃度指数、学生在线学习活跃度指数(如表7所示)。当然,不同的评价对象,在计算w指数时所选取的数量维度和质量维度又会有所不同。

表7 更多评价对象的w指数

w指数拓展类型	数量维度	质量维度	定义
课程w指数	参与学生数量	学生的行为次数、投入时间或使用模块数。	以投入天数为例,1门在线课程的实施过程活跃指数为w,当且仅当在它的N个学生中有w个学生的投入天数不少于w天,课程剩下的(N-w)个学生的投入天数都小于w天。
	使用模块数量	各个模块的使用学生数占全部参与学生的比例。	一门在线课程的实施过程活跃指数为w,当且仅当在它的100%的模块中有w%个模块的参与学生比例不低于w%,课程剩下的(100-w)%个模块的参与学生比例都小于w%。
辅导教师w指数	使用模块数量	在活动模块上做出响应(回复学生发帖、批改作业等)后涉及的学生数量;在资源模块上共同浏览同一资源的学生数量。	1位辅导教师在1门在线课程的活跃指数为w,当且仅当在他(她)使用的N个模块中有w个模块回应或共同参与的学生人数w人,剩下的(N-w)个模块的学生人数都小于w人。
学生w指数	使用模块数量	在每个模块上产生的行为次数、投入时间。	以行为次数为例,1个学生在1门在线课程的活跃指数为w,当且仅当在他(她)使用的N个模块中有w个模块的行为次数不少于w次,剩下的(N-w)个模块的行为次数都小于w次。

四、研究结论

本文从加强远程教育机构在线教学过程监测、引导师生在线教学活动更多发生这一问题出发,提出要构建远程教育机构在线教学活跃指数,并且这一指数要能同时兼顾常用的参与师生数量及师生行为次数两项指标。作者从用于科研工作人员绩效评价的h指数获得启发,提出了一项远程教育机构在线教学活跃指数——w指数。w指数的定义是:一所远程教育机构的在线教学活跃指数为w,当且仅当在它的N个用户中有w²个用户的行为次数不少于w次,机构剩下的(N-w²)个用户的行为数都小于w次。

在提出w指数定义及算法后,以国家开放大学2016年秋季学习的在线教学1.1亿日志为样本,计算国开44所分部的w指数。此时的w指数主要反映了

参与师生数量,而不能反映参与质量。问题就出在达到w指数的人数占全部师生数的比例上。在这个比例上,最高为83.41%;最低为27.26%,标准差为15.86%,变异系数为30.51%,已经达到了一个很高的差异值。因此,作者进一步改进算法,在计算w指数时,要对达到w指数的人的比例加以约束,使得这个比例不能低于40%,也不能高于60%。使用修正的算法计算后,w指数与师生数、人均行为次数均能达到一个中等相关的水平,表明此时的w指数能同时兼顾数量和质量,这也达到了预期的目标。

本文最后对w指数的计算方式及可用于的评价对象进一步拓展,提出还可将每个师生的投入天数、使用模块数作为w指数的质量维度;提出可以将课程、教师、学生作为评价对象,计算在线课程实施过程活

跃指数、辅导教师在线教学活跃指数、学生在线学习活跃指数。

正如h指数能够综合反映一位研究人员的发文量和被引次数,w指数则能够综合反映一所远程教育机构在一个学期中在线教学的参与师生数及师生行为次数。参与学与教的师生数量(广度)和师生学与教行为次数(深度)两个指标互相制衡:参与师生很多,但师生行为次数不高的机构,w指数一定不高;师生行为次数很高,参与师生数量不多的机构,w指数也会被参与师生数量所限,而不能得到高w值;只有当参与师生数量多,而且多数师生具有高行为次数时,才能获得高w指数。可见,用w指数评价机构教学绩效可以遏制片面追求参与师生数量的不良倾向,同时又能够激发机构探索深层次教学的热情。这就是w指数与其他单项机构教学评价指标相比所独具的高明和绝妙之处。我们有理由相信w指数将在引导远程教育机构在线教学不断发生并健康发展上发挥重要作用。

参考文献:

- [1] CNNIC 发布第 39 次《中国互联网络发展状况统计报告》
[J].中国信息安全,2017,(2):24.

- [2] 国务院印发《国家教育事业发展规划“十三五”规划》[N].人民日报,2017-01-20(5).
- [3] Lori Breslow, David E. Pritchard, Jennifer DeBoer, Glenda S. Stump, Andrew D. Ho and Daniel T. Seaton. Studying Learning in the Worldwide Classroom Research into edX's First MOOC
- [4] 樊超,宗利永. MOOC 在线学习行为的人类动力学分析[J]. 开放教育研究,2016,(2):53-58.
- [5] 贾积有,于悦洋. 学习活动指数 LAI 及在线学习活动指数 OLAI 的具体分析[J]. 中国远程教育,2017,(4):15-22.
- [6] 郑勤华,陈耀华,孙洪涛,陈丽. 基于学习分析的在线学习测评建模与应用——学习者综合评价参考模型研究[J]. 电化教育研究,2016,37(9):33-40.
- [7] 魏顺平,程罡. 数据驱动的教育机构在线教学过程评价指标体系构建与应用[J]. 开放教育研究,2017,23(3):113-120.
- [8] Hirsch, J.E. An index to quantify an individual's scientific research output[J]. Research Technology Management, 2005, 48 (6):62.
- [9] 金碧辉. 科学家为自己设计了一项评价指标:h 指数[J]. 科学观察,2006,(1):8-9.

(责任编辑:关 山)

A Study on Online Learning Active Index of Open and Distance Education Institutions

WEI Shunping

(IT Department, the Open University of China, Beijing 100039)

Abstract: In this paper, the author proposes w-index as active index used to evaluate online learning of open and distance education institutions, with inspiration from h-index which is used to evaluate researchers' performance. After the author puts forwards the definition of w-index and its calculation, w-index of 44 branches' online learning in the Open University of China are calculated based on 110 million action logs from LMS of OUC, which were produced in 2016. After that, some correlation test between w-index and other evaluation indexes are made. The correlation test shows that w-index is synthesis of the number of active learners and the average action number of each active learner.

Key words: open and distance education institutions; online learning; windex; active index; learning analytics; educational big data