

王凌峰. 预印本自组织同行评议中“论文-评议人匹配”规则的性能仿真[J]. 中华医学图书情报杂志, 2018, 27(5): 13-16, 74.

DOI: 10.3969/j.issn.1671-3982.2018.05.003

· 研究与探讨 ·

预印本自组织同行评议中“论文-评议人匹配”规则的性能仿真

王凌峰

[摘要] 预印本数据库引入同行评议是近年来国际学术出版研究领域的一个热点问题,对预印本采用何种同行评议模式存在不同观点。自组织同行评议是预印本同行评议模式的一种新设计,论文与评议人匹配规则是自组织同行评议的核心环节,直接决定了自组织同行评议的性能表现。对笔者以前提出的自组织同行评议机制中的“论文与评议人匹配规则”进行了改进,并基于真实数据对改进规则性能表现进行了仿真。

[关键词] 开放存取;预印本;自组织同行评议;论文-评议人匹配;性能仿真

[中图分类号] G252.8; G250.74

[文献标志码] A

[文章编号] 1671-3982(2018)05-0013-04

Performance simulation of "paper-reviewer matching" rules in self-organized peer review of preprints

WANG Ling-feng

(Guilin University of Electronic Science and Technology Business School, Guilin 541004, Guangxi Zhuang Autonomous Region, China)

[Abstract] Introduction of preprints database into peer review has become a hotspot problem in recent international academic publication research field. However, the points of view are different in terms of what kind of preprint peer review model should be used. Self-organized peer review is a newly designed preprint peer review model, and paper-reviewer matching rules are its core points which directly determine the performance of self-organized peer review. The paper-reviewer matching rules in the self-organized peer review mechanisms proposed before by the author were thus revised and simulated.

[Key words] Open access; Preprint; Self-organized peer review; Paper-reviewer matching; Performance simulation

预印本是开放存取的主要实现方式之一。20世纪年代以来,信息技术不断普及、学术界竞争日益激烈,研究者希望论文尽快公开,保护发现优先权、商业学术出版机构定价过高等多种因素推动预印本快速发展^[1]。目前预印本在学术出版中的论文储存与优先权确认功能已经得到充分证明,近年来预印本领域关注焦点是引入同行评议^[2]。全球最大

的多学科预印本数据库 arXiv 在 2017 年接受了 Heising-Simons Foundation 约 32 万美元的资助,启动了“Next Generation arXiv or arXiv-NG”计划,其中一个备受关注的议题即为是否启动同行评议。

“预印本+同行评议”模式在多方面优越于“期刊+同行评议”模式,主要体现在“预印本+同行评议”模式是“以文评文”,而不是通过期刊影响因子“以刊评文”,从而消除了论文评价中的影响因子搭便车现象^[3],彻底解决了论文刊发滞后问题(由于期刊版面有限,很多论文被录用后需要等待数月甚至半年才能刊发,如《中文社会科学引文索引》(Chinese Social Science Citation Index, CSSCI)收录的期刊《统计与决策》的官网显示,其 2017 年 5 月

[基金项目] 国家自然科学基金“E-prints 预印本数据库环境下开放同行评议机制研究”(71563007)

[作者单位] 桂林电子科技大学商学院,广西 桂林 541004

[作者简介] 王凌峰(1976-),男,四川仁寿人,博士后,教授,硕士生导师,研究方向为科技创新与管理、人力资源管理。

录用的稿件要等到 2018 年 10 月才能刊发。而预印本同行评议一结束,即可正式发布稿件,不用等待版面,消除了一篇稿件被多本期刊轮流审稿现象(一篇投稿往往不会一投即中,而是先后投多家期刊后,才被录用,而一期刊审稿周期往往为 1~2 月甚至更长,这就导致稿件最终发表的滞后期较长,而预印本同行评议结束时,论文即可正式发表,稿件不用被反复轮流审稿)。

在 CNKI 期刊论文、学位论文数据库中使用“预印本+同行评议”进行摘要检索,得到的检索结果只有 5 条,其中只有 1 篇最相关文献^[4]。英文文献方面,在 WoS 数据库中使用“pre-print+ peer review”为检索词,进行主题检索,得到 57 条结果,大部分为讨论预印本构建同行评议的必要性、可能性以及建议预印本采用社区同行评议模式。

关于预印本采用何种同行评议模式,目前最有代表性的观点是社区同行评议(Community Review 或 Reader Review)。社区评议的要点是作者向预印本数据库提交论文后,读者(主要是属于同一学科领域的注册研究者或者论文作者)自由评议,预印本后台程序根据某个规则将众多读者给出的评议信息整合为一个最终的评议结果^[5]。社区评议的主要缺点是评议参与度高度分化,一部分论文可能很长一段时间内都无人问津,如果评议结果被主流学术界接受后,用于职称晋升、经费申请等现实利益场合,则可能出现大面积作弊行为,且不易控制。因此,社区评议不宜作为预印本的主要同行评议模式,只可以作为辅助、补充^[6]。

文献[5]提出了预印本的自组织同行评议新模式,其基本思路是同一个领域的预印本注册者之间按照一套规则进行匹配后,互相评议彼此的论文。“论文-评议人匹配规则”对自组织同行评议的性能表现有决定性影响。由于同行评议实验成本较高,故先以仿真方法开展研究为宜^[7]。文献[8]提出了 4 个“论文与评议人匹配规则”,仿真结果为规则 4 最佳。由于篇幅限制,本文暂不列出规则 4 的具体内容,读者可自行下载文献[8]细读。

规则 4 的规则设计及仿真主要有 3 个不足:一是没有考虑到一名评议人评议完一篇论文后,应有适当休息期,不能评完一篇马上又评第二篇;二是洛

特卡定律的参数设定组合中,漏掉了约束条件“不同篇数论文作者所占百分比之和为 1”,导致仿真中引入了很多不存在的参数组合;三是直接给定不同水平评议人对不同等级论文的评议正确率,主观性较大。

针对上述 3 点不足,本文提出了改进后的规则 5。规则 5 对规则 4 的具体改进包括引入了评议人休息期、考虑了约束条件“不同篇数论文百分比之和为 1”,排除了不满足该条件的不合理参数组合、通过控制评议人与被议论文间等级差确保评议质量,而不是直接给定不同水平评议人对不同等级论文的评议正确率。

本文主要提出了“论文-评议人”匹配规则的操作步骤,说明了仿真变量、仿真变量的取值范围与取值依据,给出了论文与评议人匹配规则、评议速度与评议效度的性能评价指标,讨论分析了仿真结果,总结了贡献,指出了需要进一步研究的问题。

1 规则 5 要点与性能评价指标

1.1 规则 5 要点

规则 5 的要点主要包括:评议人资格等级比被评论文等级最多低 1 个等级;有多个满足条件的评议人组合,优先匹配评议任务次数与已完成评议任务次数差最大者;如果还有多个组合,选择休息期累积时间最长的组合或随机选一个;如果没有满足条件的组合,等待下一轮匹配。

所谓“评议任务次数与已完成评议任务次数差最大”的含义是:假设一篇文章安排 3 个人评议,有两个评议人组合,分别是甲乙丙 3 人、ABC 3 人。其中,甲提交系统 5 篇、乙提交 3 篇、丙提交 2 篇,那么甲乙丙需要完成的评议任务即各自提交论文数量乘以 3,因为每篇论文都需要 3 个人评议,分别是 15、9、6 次评议任务。假设甲、乙、丙分别评议了 1 篇,则甲乙丙组合还需要评议 14+8+5 即 27 次。类似地,ABC 分别提交 1、2、3 篇论文,评议任务次数分别为 3、6、9 次,3 人分别评议了 1 篇,则 ABC 还需要评议 2+5+8 即 15 篇。由于甲乙丙组合评议差 27 大于 ABC 的评议任务差 15,故优先向甲乙丙组合分配等待评议的论文。

“评议人的评议等级比被评论文等级最多低 1 个等级”意味着论文由和论文等级基本相当或更

高等级的评议人进行评议,故能以较大概率确保评议质量。评议质量很难直接测度,间接测度更可行。该指标值低于 70% 意味着 3 个评议人中有两个评议人的资格等级比被评议论文低了两个等级,评议质量低的概率会很高。

1.2 匹配规则的性能评价指标

“论文-评议人”匹配规则的性能评价指标包括论文评议完成比例、单篇论文评议速度两个方面。评议完成比例是指一个仿真周期(365 天)内各等级

论文完成评议的比例,评议速度是指一个仿真周期(365 天)内各等级论文的单篇论文评议时间的均值,每一篇论文以评议人中最晚完成时间计算。

评议质量的保证通过直接要求评议人资格等级满足一定条件实现,不必额外讨论。

2 规则 5 的性能仿真参数说明及取值

规则 5 的性能评价仿真变量以及各变量取值见表 1,各变量取值的依据主要来自科学计量学的相关理论与规律以及部分客观数据。

表 1 规则 5 的性能仿真变量与取值

变量名	变量取值范围	变量取值依据
一篇论文评议人数	1 篇论文由 3 人进行评议	国际通行做法
仿真时间长度	365 天	
论文作者数量(同时也是论文评议人)	500 人、2000 人两个取值水平	
论文总数量		给定论文作者总数、洛特卡分布的参数后,论文总数也就自动确定,不必单独给出
论文水平分布	第 5 等级论文占全部论文数的 5% 第 4 等级论文占全部论文数的 10% 第 3 等级论文占全部论文数的 15% 第 2 等级论文占全部论文数的 30% 第 1 等级论文占全部论文数的 40%	数据来自文献[9]及文献[10]
“论文数-作者”分布比例(即每年撰写 1 篇、2 篇到 6 篇论文的百分比)	由于各篇论文百分比之和等于 1,故 N 取值确定后 C 对应确定,N 取值在 1.25 到 3.5 之间	多学科领域的多数作者一年通常最多写 6 篇论文,超出则质量难以保证(参数来自文献[11]、文献[12])
每个作者的论文评议等级	该作者论文水平最高等级	该指标描述一个作者评价别人论文的水平高低
一个仿真周期 365 天中每个月论文提交数量分布	1 月 7%、2 月 7%、3 月 11%、4 月 21%、5 月 24%、6 月 6%、7 月 8%、8 月 4%、9 月 3%、10 月 1%、11 月 6%、12 月 3%	“中国科技论文在线工作”人员提供的“中国科技论文在线”2017 年各月份发表论文数据
评议速度(即评议完一篇论文所需天数)	评议时间 15 天以内占 14%、15~30 天占 35%、31~45 天占 24%、46~60 天以内占 15%、60~90 天占 12%	数据来自文献[13],61 天以上的合并,最长时间设置为 90 天,各类内部的个体评议速度按照均匀分布随机选择一个天数
休息期	10 天或 20 天	评议人完成一篇论文评议后需要休息一段时间
预印本注册者状态更新周期	每天定时更新一次	
“论文-评议人”的匹配周期	每天定期定时匹配一次	从时间顺序上,必须先更新注册者评议状态,确定评议人没有处于休息期、没有正在进行评议后,才能进行“论文-评议人”匹配

论文总数量在论文作者中的分布情况按照广义洛特卡分布进行分配。1926 年美国学者洛特卡提出“撰写 x 篇论文作者大约是写 1 篇论文科学家人数的 $1/x^2$ ”,后被推广为广义洛特卡规律,表达式为

$F(x) = C/x^n$, $F(x)$ 是撰写 x 篇论文的作者占全部作者的百分比。考虑到不同学科研究人数差异、不同学科研究周期与论文产出数量差异以及学术传播途径差异等因素, n 取值一般在 1.2 到 3.8 之间波动。

如果低质量论文数量较大,则洛特卡定律有效性下降。洛特卡提出该定律的时代,研究者数量很少、论文发表数量也少,与今天有很大不同。文献[14]对科学网博客发表的网络学术信息作者分布规律进行统计后发现,博文与作者分布不服从洛特卡定律,因为博文发表要求比正规论文低很多,但有基本质量要求的成果均与洛特卡定律吻合,如专利与作者关系,显然专利有基本质量要求^[15]。文献[16]对 Emerald 数据库进行了研究,发现“论文-作者”分布同样符合洛特卡定律。

3 仿真结果

使用 Python 3.4 语言编程,程序在 Windows 环境中运行。本文没有推断统计进行数据分析,原因是仿真结果特征非常明显,描述统计分析已经足够。今后研究可以采用推断统计方法进行更细致的分析。程序运行 100 次,得到结果的均值,由于篇幅原因,没有列出全部结果。整理后的仿真结果见表 2。

表 2 规则 5 的性能仿真结果

	3 名评议人+规则 4	3 名评议人+规则 5
5 等级完成率	95%	93%
4 等级完成率	95%	94%
3 等级完成率	小于 82%	93%
2 等级完成率	小于 82%	94%
1 等级完成率	小于 82%	92%
5 等级评议速度	28 天	46 天
4 等级评议速度	28 天	47 天
3 等级评议速度	56 天	49 天
2 等级评议速度	56 天	48 天
1 等级评议速度	56 天	49 天

仿真主要结果如下。

一年内评议完成比例方面,规则 4 中的 1~5 等级论文差异较大,4、5 等级论文完成率均为 95%,1、2、3 等级论文的完成率均小于 82%;规则 5 各等级论文的等级完成率则比较均衡,平均为 92%。

评议速度方面,规则 4 的高等级论文评议速度比规则 5 快 2 周,规则 5 的低等级论文评议速度比规则 4 快 1 周,规则 5 的各等级论文评议速度均稳定在 7 周左右。

评议质量方面,规则 4 与规则 5 定义不同,但都

要求评议人组合满足设定条件,故评议质量在各自定义下保持较高水平。

总的来说,规则 5 仿真性能表现比规则 4 更加均衡、稳定。需要说明的是,规则 5 的高等级论文评议速度慢于规则 4 的原因是规则 4 没有考虑评议人的休息期因素。另外,根据前面的假设,评议速度在 46~90 天之间的评议人占全部评议人的 27%,几乎占三分之一,完全可以采取措施加快速度,将评议速度减少到 30~45 天。因此,本文建议,某一次没有按时完成评议的作者,其作为第一作者或者通讯作者提交预印本要求评议的论文、参与署名的论文在 1 个月内暂时不予评议,这样将会激励大部分作者按时完成评议。

从出版速度看,出版速度=评议速度+版面等待时间。规则 5 的评议速度平均为 57 天,预印本版面等待时间为 0。国内学术期刊评议速度一般为 2 个月(60 天),稿件录用后 2 个月内刊发的期刊尤其核心期刊是罕见的,高影响因子期刊甚至可能录用半年后才有版面刊发。因此,“预印本+自组织同行评议”模式的出版速度比“期刊+同行评议”模式快。

4 结语

文献[5]首次提出了“预印本+自组织同行评议”模式,文献[8]提出了自组织同行评议中的 4 个“论文-评议人”匹配规则并进行了性能仿真。本文对文献[8]提出的规则 4 进行了 3 点改进,改进后的规则 5 考虑了评议人休息期等更加符合实际情况的内容,仿真结果表明规则 5 性能表现比规则 4 更加稳定、均衡,可用于预印本构建同行评议。

(致谢:Python 仿真程序由桂林电子科技大学计算机与信息安全学院黄琦编写。)

【参考文献】

- [1] Michael P. Taylor. Opinion: academic publishing is broken. [EB/OL]. (2012-03-19) [2018-03-20]. <https://www.the-scientist.com/news-opinion/opinion-academic-publishing-is-broken-41231>.
- [2] Curry S. Peer review, preprints and the speed of science [EB/OL]. (2015-09-07) [2018-03-20]. <https://www.theguardian.com/science/occams-corner/2015/sep/07/peer-review-preprints-speed-science-journals>.

视频操作,甚至抄袭其他学生的作业,他们并没有思考相关知识和数据库的操作过程。同时也说明实验课教学和考核方法仍然需要改进,这也正是我们下一步工作的方向。

文献检索课程计算机辅助教学”管理系统,实现实验课内容与考核的在线化;优化 UI 设计,帮助学生更好地掌握医学信息的检索与利用。

【参考文献】

[1] 陈 界,冯泽伦.我国 20 余年医学文献检索教学的经验与当前的问题及其建议[J].中华医学图书情报杂志,2005,14(3):46-49.

[2] 张 波.《信息检索与利用》对当代大学生提高信息素养重要性[J].科学中国人,2016(6):322.

[3] 彭丁晋,黄泽智.从试卷综合分析谈医学教育教学质量评价[J].国际检验医学杂志,2009,30(12):1214-1214.

[4] 唐莉萍.基于多元化考核模式的文献检索课程改革与实践[J].卫生职业教育,2017,35(17):136-139.

[5] 张华亮.基于 Linux 自由及开源代码的 Android 操作系统[J].计算机与网络,2016,42(3-4):108-111.

[6] 李世川. MariaDB 应用全攻略[J].网络运维与管理,2015(6):137-142.

[7] 王 芳.当前流行 Web 开发语言:PHP[J].信息系统工程,2014(5):30.

[8] 杨 楠,吴 凌,王千祥. POP-PHP:支持 PHP 应用的在线集成开发环境[J].计算机科学,2014,41(9):38-44.

[9] 徐文亮.在线考试系统的研究与实现[D].昆明:云南大学,2012:23-24.

[收稿日期:2018-04-22]

[本文编辑:黄思敏]

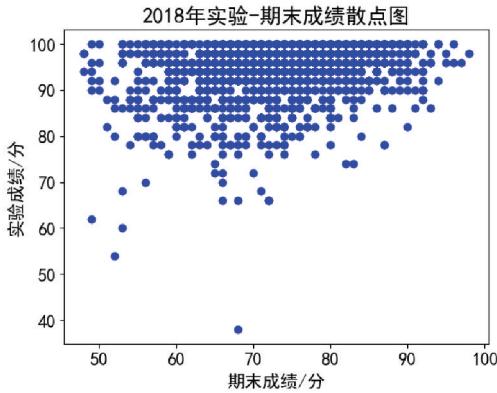


图 8 2018 年实验-期末成绩分布

6 结语

“医学文献检索”在线考试系统的建设与应用,实现了传统笔试向上机考试的平稳过渡,减轻了教师的阅卷压力,能够从考试中获取更多的信息以促进教学改革。同时在线考试系统的实施也反映出一些问题,如学生实验学习主动性不强、实践操作能力较低等。我们将进一步完善该系统,拟建立“医学

(上接第 16 页)

[3] 王凌峰.论文评价中影响因子搭便车现象[J].现代情报,2013,33(2):12-14.

[4] 王凌峰,孙英潮. E-prints 预印本数据库的自组织同行评议模式设计[J].现代情报,2016,36(5):9-15.

[5] Alfaro L, Faella M. True review: a proposal for post-publication peer review[EB/OL]. (2017-05-09) [2018-03-20]. <https://export.arxiv.org/pdf/1608.07878>.

[6] García J Rodríguez-Sánchez R, Fdez-Valdivia J. Problems with open participation in peer review[J]. Scientometrics, 2017, 112(3): 1881-1885.

[7] Allesina S. Modeling peer review: an agent-based approach[J]. Ideas in Ecology & Evolution, 2012, 5(2): 27-35.

[8] 王凌峰,姚依楠.自组织同行评议机制中 4 个论文-评议者匹配规则设计[J].情报杂志,2018,37(3):142-145,151.

[9] 毛国敏,蒋知瑞,任 蕾,等.期刊论文被引频次的幂律分布研究[J].中国科技期刊研究,2014,25(2):293-298.

[10] 刘 琳.情报学文献作者分布规律研究(1991-2007)[J].情报科学,2008,26(11):1694-1698.

[11] 邱均平.信息计量学[M].武汉:武汉大学出版社,2007:169-172

[12] 陈道兰.广义洛特卡分布特征参数的稳定性研究:以《情报理论与实践》为例[J].情报理论与实践,2014,37(8):28-31.

[13] 朱乾坤,石红青.从审稿统计数据看审稿人的选择[J].编辑学报,2010,22(2):151-153.

[14] 邱均平,余 凡.网络学术信息作者分布规律研究:以科学网博客为例[J].图书情报工作,2011,55(20):15-18.

[15] 张海燕.基于文献计量学的洛特卡定律研究综述[J].中华医学图书情报杂志,2013,22(8):18-21.

[16] Rowlands I. Emerald authorship data, Lotka's law and research productivity[J]. Aslib Proceedings, 2013, 57(1): 5-10.

[收稿日期:2018-03-20]

[本文编辑:黄思敏]