

# 融合关联规则的 MOOC 资源众包平台任务分配算法

潘志宏<sup>1</sup>, 万智萍<sup>1</sup>, 谢海明<sup>2</sup>

(1. 中山大学新华学院, 广东 广州 510520;

2. 中国移动通信集团广东有限公司, 广东 广州 510623)

**摘要:**在高校大力发展 MOOC 平台背景下,为了提升学生的主动性、MOOC 平台资源库的丰富性,出现了众包协同构建资源的方法,让学生和教师构成学习共同体,在完成新知识学习的同时进行资源库建设。高校 MOOC 众包平台的任务就属于知识密集型任务,挑选合适的参与者直接关系到 MOOC 资源库构建质量。为了更好地构建高校 MOOC 资源平台,提出一种针对知识密集型众包任务的分配方案,它包含学生的准入筛选、预期工作能力评估两个阶段。首先利用改进 Apriori 课程关联算法对学生进行准入筛选;其次利用知识关联算法对学生预期工作能力进行评估并将众包任务分配到工作能力最合适的学生;最后对方案进行测试验证,结果表明该方案能够较好地提升学生挑选和任务分配的效果,促进构建更高质量的 MOOC 资源库。

**关键词:**知识密集型众包任务;MOOC 资源构建;任务分配;关联规则;Apriori 算法;预期工作能力评估

中图分类号:TP311

文献标识码:A

文章编号:1673-629X(2020)04-0189-06

doi:10.3969/j.issn.1673-629X.2020.04.036

## Task Allocation Algorithm in MOOC Resource Crowdsourcing Platform Combined with Association Rules

PAN Zhi-hong<sup>1</sup>, WAN Zhi-ping<sup>1</sup>, XIE Hai-ming<sup>2</sup>

(1. Xinhua College of Sun Yat-sen University, Guangzhou 510520, China;

2. China Mobile Group Guangdong Co., Ltd., Guangzhou 510623, China)

**Abstract:**In the context of vigorously developing MOOC platform in colleges and universities, in order to enhance the initiative of students and the richness of the MOOC platform resource, there is a crowdsourcing method for collaboratively building resources, which enables students and teachers to form a learning community and carry out the resource construction while learning the new knowledge. Since the task of college MOOC crowdsourcing platform is knowledge-intensive, the selection of appropriate participants is directly related to the construction quality of MOOC resource pool. In order to better build an excellent resource pool for college MOOC platform, we propose a task allocation scheme for knowledge-intensive crowdsourcing task, which includes two stages of student admission screening and expected work ability assessment. Firstly the course association rules based on improved Apriori algorithm are used to make the student admission judgment. Secondly, the knowledge association rules are used to evaluate the students' ability and assign the crowdsourcing tasks to students with the most appropriate work ability. Finally, the performance of the proposed scheme is tested and verified. The results show that it can improve the effect of student selection and task assignment, and promote the construction of higher quality MOOC resources.

**Key words:**knowledge-intensive crowdsourcing task; MOOC resource construction; task assignment; association rules; Apriori algorithm; expected work ability assessment

## 0 引言

近年来随着移动互联网、大数据、云计算等技术的

不断发展,国内外涌现大批 MOOC 平台,它让学习者根据不同的学习主题形成学习圈,并通过社交网络构

收稿日期:2019-04-22

修回日期:2019-08-26

网络出版时间:2019-12-05

基金项目:2016 广东省普通高校青年创新人才项目(自然科学)(2016KQNCX222);教育部产学研合作协同育人项目(201802153146, 201702071078);2016 年度广东省本科高校高等教育教学改革项目(2016SGJ002);中山大学新华学院教职工科研启动基金青年项目(2017QN006)

作者简介:潘志宏(1984-),男,硕士,高级工程师,讲师,CCF 会员(29357M),研究方向为群智计算、机器学习、物联网。

网络出版地址: <http://kns.cnki.net/kcms/detail/61.1450.TP.20191205.1133.050.html>

建学习群体。但目前大部分 MOOC 平台的课程资源仅依靠平台方来建设,学习者只需要按部就班地学习课程内容,这种模式具有一定的局限性。首先学习者会比较被动地学习平台所提供的资源,缺乏参与学习的主动性;其次平台学习资源仅依靠一方来维护,很难保证内容的多样性。因此为了提升学习者的主动性以及平台资源库的丰富性,有学者提出将众包理念融入 MOOC 平台<sup>[1-4]</sup>,让学习者和教师构成学习共同体,在完成新知识学习的同时进行资源库建设。通过众包协同构建 MOOC 资源的步骤如下:首先,教师通过 MOOC 平台发布任务,包含以校验学习效果为基础型任务、以构建 MOOC 资源为主的开放型任务、以评估他人任务完成效果为主的评估型任务等各种类型众包任务,这些设置合理的众包任务一方面让学生融入课程学习的全过程,另一方面协助 MOOC 平台资源的动态更新;其次,设计众包质量控制策略保证众包任务在实现教学功能的前提下帮助平台构建更高质量的 MOOC 资源也是非常重要的问题。

在众包研究中质量控制一直是重点和热点,主要集中在众包任务分配和众包任务评估两方面<sup>[5-7]</sup>,其中合理分配众包任务是获得优秀众包结果的核心基础,不同类型的平台有不同的任务分配策略。第一类众包平台通常被用来发布简单的微任务,对参与者的技能要求不高,它更多时候采用的是先来先得的任务分配方式,比如通过亚马逊 AMT 发布标注图片任务<sup>[7]</sup>;第二类众包平台虽然对参与者的知识技能要求不高,但是对参与者的位置、时间有严格要求,比如通过群智感知平台收集移动终端用户的感知位置、时间、照片等信息获得城市交通、环境等情况<sup>[8-9]</sup>;第三类是知识密集型众包任务,对参与者的知识技能水平都有特定的要求,比如众包翻译<sup>[10]</sup>、软件众包平台 TopCoder<sup>[11]</sup>等。高校 MOOC 资源众包平台上的任务就属于知识密集型任务,因此挑选合适的任务参与者至关重要,它关系到 MOOC 资源的构建质量。文中利用课程、学生知识技能与众包任务的关联关系来设计任务分配算法,并将任务分发至最合适的学生来完成,以构建更高质量的 MOOC 资源库。

主要贡献在于提出一种针对高校 MOOC 资源众包平台的任务分配方案,它包含学生的准入筛选、预期工作能力评估与任务分配两个阶段。方案首先设计基于改进 Apriori 的课程关联规则生成算法<sup>[12]</sup>对学生选课进行准入筛选;其次设计基于黄金标准和知识关联的学生预期工作能力评估与任务分配算法,将众包任务分配到工作能力最合适的学生;最后分别在任务课程关联、任务知识关联两个方面对方案的实现效果进行测试验证,测试结果表明,该方案能够较好地提高

MOOC 资源的构建效果。

## 1 高校 MOOC 资源众包平台的任务分配策略

挑选合适的任务参与者对知识密集型众包任务是非常关键的,它是获得高质量众包结果的重要保证。文中从学生选课准入筛选、学生能力评估和任务分配两个方面来设计 MOOC 资源众包平台的任务分配方案,并设计详细的算法流程。

### 1.1 构建高校 MOOC 资源平台的众包任务分配方案

一般众包平台都想用尽量少的成本雇佣合适的人以高质量地完成众包任务,但是高校 MOOC 资源众包平台在众包人员的选择上具有特殊性。首先平台强调参与广度,让更多学生能受惠于优质的学习资源;其次平台任务属于知识密集型众包任务,对知识和技能要求较高,所以必须对学生的技能进行一定的评估筛选,以确认学生更好地完成课程学习和构建 MOOC 资源的任务;最后普通众包平台通常会因为成本问题,在进行众包质量控制时尽量避免使用太多冗余人员来获得更高质量的结果,但这却是高校 MOOC 平台的优势,它没有成本限制,可以让更多的学生参与同样的任务,从而获得更高质量的 MOOC 资源库。

针对上述三个特点,首先提出基于改进 Apriori 的课程关联规则生成算法,在课程准入方面对学生进行初步筛选,其次提出黄金标准和知识关联的学生预期工作能力评估与任务分配算法,在不同知识块对学生预期任务工作能力进行多维度评估,保证他们有能力完成众包任务。图 1 给出了高校 MOOC 众包资源平台的任务分配工作流程,包含学生准入筛选、工作能力评估与任务分配两大模块。

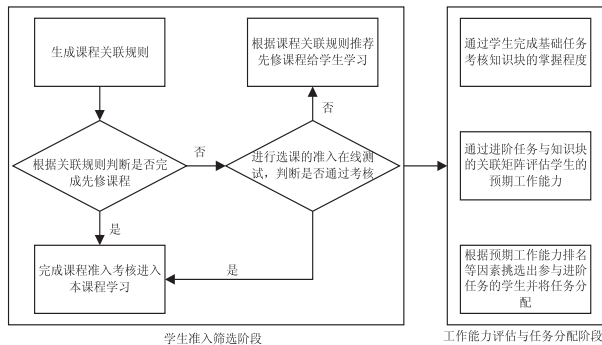


图 1 MOOC 平台任务分配工作流程

### 1.2 基于课程关联规则的学生准入筛选

学生准入筛选阶段,判断学生是否具备完成课程学习和完成相应众包任务的知识储备,一般情况下高校会采用人才培养方案的先修课程规则对学生进行初步过滤准入,这种通过静态课程设置规则进行简单限

制具有一定的缺陷,首先它不能很好地过滤出更多满足条件的学生,其次也不能较好地适应不断更新的高校 MOOC 平台的课程体系。针对这种情况,文中利用改进 Apriori 算法对不同课程成绩进行分析从而得到课程之间的关联规则<sup>[12]</sup>,一方面可以进行选课准入筛选,另一方面当学生没有完成相关先修课程或者没有通过课前能力测试时,可以通过课程关联规则推荐相关先修课程给学生学习。

首先高校 MOOC 平台与本校教务系统对接获取课程成绩数据源;接着进行课程关联规则生成,主要分为三个步骤:第一对成绩数据进行预处理;第二根据课程成绩数据找出有效的频繁项集;第三从生成的频繁项集中生成高置信度课程关联规则;最后将规则保存至对应数据库,以供用户准入过滤使用。详细流程如算法 1 所示,为了算法描述方便,做如下定义。

定义 1:数据项集  $I = \{i_1, i_2, \dots, i_m\}$ , 集合  $X \subseteq I$  为项集,  $T_n = \langle t, X \rangle$  为每个事务,  $D = \{T_1, T_2, \dots, T_n\}$  为事务数据库,  $I_m = \langle i_m, t_m \text{Set} \rangle$  ItemDB 为项目事务数据库,  $t_m \text{Set}$  表示包含  $i_m$  事务的所有事务 id 集合。

定义 2:  $X \rightarrow Y$  表示课程关联规则,支持度用  $\text{support}(X \rightarrow Y)$  表示,它代表  $X$ 、 $Y$  同时出现在  $D$  中的概率,计算公式如式(1)所示。其中  $\text{sup}(XY)$  表示  $XY$  在  $D$  中同时出现的次数,  $\text{Count}(D)$  表示  $D$  中事务的总数量。置信度用  $\text{conf}(X \rightarrow Y)$  表示,它代表  $X$  出现,  $Y$  也出现的概率,计算公式如式(2)所示。其中  $\text{sup}(X)$  表示  $X$  在  $D$  中出现的次数。

$$\text{support}(X \rightarrow Y) = P(X \cup Y) = \frac{\text{sup}(XY)}{\text{Count}(D)} \quad (1)$$

$$\text{conf}(X \rightarrow Y) = P(Y | X) = \frac{\text{sup}(XY)}{\text{sup}(X)} \quad (2)$$

定义 3:  $L_k$  表示频繁  $K$  项集,  $C_k$  表示候选项集,最小支持阈值为  $\text{minsup}$ ,最小置信度为  $\text{minconf}$ ,规则数据库为 RDB。

算法 1:基于改进 Apriori 的课程关联规则生成算法。

1:初始化:设置  $\text{minsup}$ ,  $\text{minconf}$  初始值。

2:数据采集与预处理:首先通过 API 接口对教务系统的学生课程成绩进行采集,将成绩进行离散化。文中只是为了分析学生学好某门课程能否学好另外一门课程,因此只分是否优秀等级,85 ~ 100 为优秀等级,用“1”表示,0 ~ 84 为不优秀等级,用“0”表示,将课程用  $KC_n (n \geq 1)$  表示,经过预处理后生成事务数据库  $D$ 。

3:产生有效的课程频繁项集。

将事务数据库  $D$  转化为垂直结构的项目事务数据 ItemDB,接着扫描项目事务数据 ItemDB 并根据最

小支持度获得频繁 1 项集  $L_1 = \text{find\_frequent\_1-itemsets}(\text{ItemDB})$

for(  $k = 2$ ;  $L_{k-1} \neq \emptyset$ ;  $k++$ ) do

$C_k = \text{Apriori\_gen}(L_{k-1})$  //生成新的候选项集

foreach  $i$  in ItemDB do //  $i$  表示 ItemDB 中每个项目事务

$C_i = \text{subset}(C_k, i)$  //获取  $i$  中包含的候选项集

foreach  $c$  in  $C_i$  do //  $c$  表示  $C_i$  中每个候选项

$c.\text{count}++$

end

$L_k = \{c \in C_k \mid c.\text{count} \geq \text{minsup}\}$

end

$L = \bigcup_k L_k$  //返回所有频繁项集  $k$

4:生成课程关联规则。

foreach  $L_k$  in  $L$ , 其中  $k \geq 2$  do

$A \leftarrow \{X \mid X \subset L_k, X \neq \emptyset\}$  //初始化前件集合

while  $A \neq \emptyset$  do

$X \leftarrow A$  中的最大元素

$A \leftarrow A \setminus X$  //从  $A$  中删除  $X$

$\text{conf} = \text{sup}(L_k) / \text{sup}(X)$

if(  $\text{conf} \geq \text{minconf}$ ) then

output  $X \rightarrow Y$ ,  $\text{support}(L_k)$ ,  $\text{conf}$  并存入 RDB

else

$A \leftarrow A \setminus \{W \mid W \subset X\}$  //从  $A$  删除  $X$  所有子集

end while

end for

算法性能优化分析:一般 Apriori 算法在计算项集时会扫描原始数据表,在面对大规模数据时会显得性能较低,文中主要从三个方面来优化算法性能问题:第一在算法步骤 2 中采用离散化对数据进行预处理并保存事务数据库,这样就会降低数据扫描和处理的速度;第二在算法步骤 3 中将事务数据库  $D$  转化为垂直结构的项目事务数据 ItemDB<sup>[13]</sup>,记录并统计每个 1 项集出现次数并根据最小支持度删除不满足的项集,从而生成新的数据库,然后用 Apriori 算法的连接,同时求连接项目的记录交集,重新生成记录数据库,这样就避免了重新扫描原始数据库,而且生成的频繁项目集越多,扫描的数据库就会越小,从而提高了运行效率;第三由于平台新增课程更新速度较小,在算法步骤 4 中将生成规则保存至规则数据库,只有新增课程时才会启动课程关联算法。从空间复杂度看,每次产生频繁项集后,数据库不断减小,减少占用内存空间;从时间复杂度看,改进 Apriori 算法节约了剪枝步骤中频繁项集的比较次数,大大提升了算法运行效率。

两种算法在不同最小支持度下的时间总开销对比如图 2 所示。



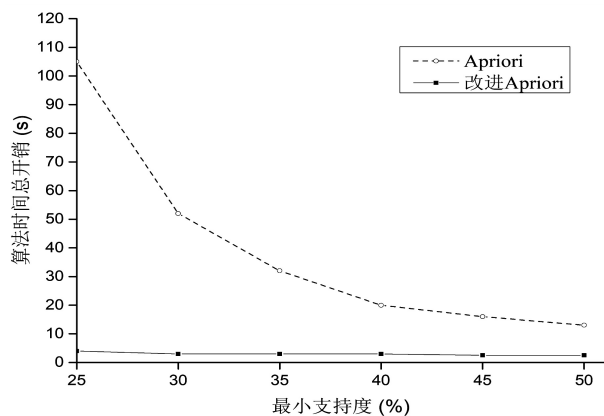


图 2 两种算法在不同最小支持度下的时间总开销对比

### 1.3 基于黄金标准和任务知识关联的学生预期工作能力评估与任务分配算法

众包平台都想获得更优秀的众包结果,因此非常关注任务参与者的工作能力与众包任务的匹配度,故设计分配算法将任务分配到匹配度最高的参与者就成为研究重点。文中提出了基于黄金标准和任务知识关联的学生预期工作能力评估与任务分配算法,首先是利用具有黄金标准答案的基础任务对学生预期工作能力进行评估,接着在学生参加某个进阶任务时需要对学生进行综合评估,评估重点为学生在与该进阶任务相关联的基础任务中的完成情况,以此确定是否参加进阶任务<sup>[14-15]</sup>。

高校 MOOC 资源平台有两类任务:第一类是课程学生都必须参与的基础任务,它由具有黄金标准答案的客观问题构成,可以通过它来评估学生预期的工作能力;第二类任务是通过预期评估挑选出来的学生才能参与的进阶任务,它大部分由不知道答案的开放性问题组成,这些设置合理问题是帮助构建丰富 MOOC 资源平台的关键。基础任务是按照课程的不同章节知识块来设置的,但是进阶任务既可以是单知识块内的任务,也可以是跨不同知识块的任务。由于进阶任务是构建 MOOC 资源的核心,在学生参与时需要评估他们的预期工作能力,并且进阶任务大多数是跨越不同知识块综合任务,所以必须全面考虑学生在该进阶任务相关联的基础任务中的综合表现。文中提出任务知识关联矩阵这个关键概念,并设计了能力评估和任务分配的详细算法流程。

定义 4:假设一门课程有  $m$  项进阶任务,用  $T_i (1 \leq i \leq m)$  表示每项进阶任务,用  $TDB_i (1 \leq i \leq m)$  表示存放每项进阶任务所有学生的能力评估得分,用  $WDB_i (1 \leq i \leq m)$  存放满足参与进阶任务条件的学生。

定义 5:一门课程有  $n$  个学生,用  $W_j (1 \leq j \leq n)$  表示每个学生。

定义 6:一门课程有  $k$  个知识块,每个知识块对应一组基础任务,其中用  $B_t (1 \leq t \leq k)$  表示知识块,用  $FT_t (1 \leq t \leq k)$  表示一组基础任务,知识块得分数据库  $BDB_t (1 \leq t \leq k)$ ,存储每个知识块所有学生的得分。

定义 7:知识块与进阶任务的关联矩阵,用  $R$  来表示关联矩阵,  $bt_{it} (1 \leq t \leq k, 1 \leq i \leq m)$  表示第  $t$  个知识块与第  $i$  项进阶任务的关系,取值为 0 或者 1,1 表示任务跟该知识块有关联,0 则表示两者无关联。

$$R = \begin{pmatrix} & T_1 & T_2 & \cdots & T_m \\ B_1 & bt_{11} & bt_{12} & \cdots & bt_{1m} \\ B_2 & bt_{21} & bt_{22} & \cdots & bt_{2m} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ B_k & bt_{k1} & bt_{k2} & \cdots & bt_{km} \end{pmatrix} \quad (3)$$

定义 8:筛选因子 FF (filter factor) 表示根据排名从学生中筛选出参加进阶任务的百分比,取值为  $(0 < FF < 1)$ ,筛选工作能力阈值 FWCT (filter work capacity threshold) 表示筛选出来参加进阶任务的学生预期工作能力得分不能低于这个阈值。

算法 2:基于黄金标准和任务知识关联的学生预期工作能力评估与任务分配算法。

- 1: 输入:  $T_i, W_j, B_t, FT_t, R, FF$
- 2: 输出:  $WDB_i (1 \leq i \leq m)$
- 3: 学生在完成每个知识块学习之后参加对应的一组基础任务  $FT_t (1 \leq t \leq k)$ ,系统自动评分并存放入  $BDB_t$  数据库
- 4: foreach  $T_i (1 \leq i \leq m)$  do
- 5: 从关联矩阵  $R$  中取出第  $i$  列
- 6: foreach  $B_t (1 \leq t \leq k)$  do
- 7: 读取  $bt_{it}$  用于判断进阶任务  $T_i$  与每个知识块  $B_t$  之间的关联
- 8: if ( $bt_{it} = 1$ )
- 9: 将  $BDB_t$  数据库对应第  $t$  知识块的所有学生得分加入  $TDB_i$  中
- 10: end if
- 11: end for
- 12: 将第  $i$  个进阶任务所有学生的能力评估得分  $TDB_i$  转为百分制得分进行排序
- 13: 从  $TDB_i$  中取出前  $n * FF$  个学生得分记录存入  $WDB_i$
- 14: 判断  $WDB_i$  是否有学生得分低于 FWCT,如果有将其从中剔除  $WDB_i$
- 15: end for

算法时间复杂度分析:假设对每个知识块关联处理时间为  $t$ ,一门课程包含  $m$  个进阶任务,每个进阶任务与  $k$  个知识块相关联,则算法计算时间总的开销为:

$T = m \times k \times t$ , 计算复杂度表示为  $O(N^2)$ 。由于  $k$  的取值相对较小, 一般一门课程大概分为 50–100 个知识块左右, 相对于一门课程上万甚至十万级别学生这个数量级来说,  $k$  可以忽略不计, 因此时间复杂度约等于  $O(N)$ , 算法运行时间总开销主要取决于选课学生数量。

2 实验结果及分析

本节主要在任务课程关联、任务知识关联两个方面对方案提出的两种算法进行验证, 利用教务数据和 MOOC 平台的使用数据进行测试, 并配置方案的最佳参数值, 最后与其他方法的应用效果进行测试对比。

2.1 融合课程关联规则的准入筛选实验

为了验证课程关联规则是否能灵活地解决学生选课的准入判断, 在 MOOC 平台选取三门课程, 对选课学生没有加入任何准入判断(NR), 但是对满足静态课程设置规则(SR)、提出的课程关联规则(AR)这两种规则的学生分别进行标注和跟踪, 并针对这三种情况下学生参加人数、完成基础任务平均得分两个维度进行测试对比, 结果如图 3 所示。

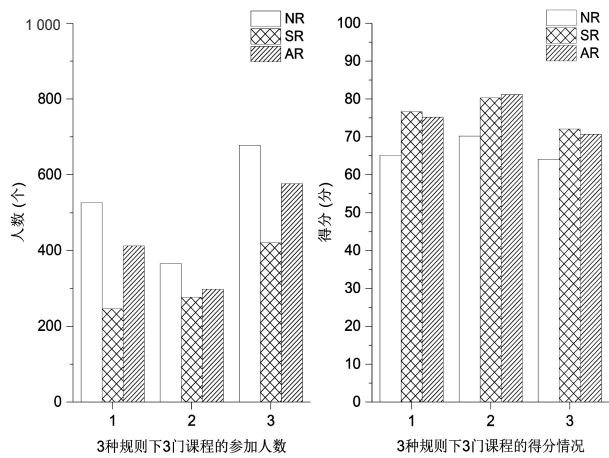


图3 学生准入筛选效果对比

从图 3 可以看出, 在没有任何准入过滤下, 虽然参加的学生人数是最多的, 但是学生的基础参差不齐, 完成任务平均分数最低; 静态课程规则由于是根据培养方案静态设置, 对学生过滤也是最严格的, 即使学生具备学习本课程的能力, 也会因为没有完成静态规则对应的先修课程直接被过滤掉, 所以它参加人数是最少的。文中提出的课程关联规则通过成绩挖掘出更多课程关联规则, 一方面能最大限度地满足学生选课的要求, 另一方面所选取的学生能很好地完成课程学习和众包任务, 静态规则和课程关联规则下学生完成基础任务能力差距微小。说明文中提出的算法在人数限制和能力限制这两方面做到了很好地均衡, 较好地满足了平台的实际需求。

2.2 融合知识关联的众包任务分配实验

在第 1.3 节中提出算法通过进阶任务与知识块之间的关联关系来评估学生预期工作能力从而进行任务分配, 下面通过两个实验测试算法效果。第 1 个实验主要是通过测试分配结果不断优化算法参数来获得更优性能。文中地众包任务分配算法有两个关键参数, 分别是筛选因子 FF 和筛选工作能力阈值 FWCT, 它们既要保证众包结果的效果, 也要尽量满足让更多学生参加进来, 这是 MOOC 资源众包平台与其他众包平台最大地不同。下面通过几组实验测试, 来选择更合适的参数值, 让算法更好地满足和均衡 MOOC 平台的两大要求。实验中选取两个参数的 3 组参数设置, 数据如表 1 所示, 选取一门课程中的 10 项进阶任务, 观察在这 3 组不同参数下, 文中提出的算法在每项任务上筛选出来的学生数量和完成每项进阶任务的平均得分。实验结果如图 4 所示, 可以看出, 在第  $PS_1$  与  $PS_2$  参数下, 所选学生完成进阶任务的得分差距非常小,  $PS_2$  参数下得分比  $PS_1$  参数得分略低 1.7% 左右, 但是  $PS_2$  参数下参与学生数量提高接近 10% 左右, 可以让更多学生参与学习任务中, 较好地达到了平台的两个要求;  $PS_3$  参数下虽然参加人数提高了, 但是得分明显降低较多, 10 项任务平均得分比  $PS_2$  参数下低 5.1% 左右。通过上面的分析结果可以得出, 算法在第 2 组参数下能取得不错的效果, 较好地满足平台需求。

表1 不同筛选因子和筛选阈值取值

参数组别 $PS_n$	筛选因子 FF	筛选工作能力阈值 FWCT
第 1 组 $PS_1$	0.15	80
第 2 组 $PS_2$	0.25	75
第 3 组 $PS_3$	0.35	70

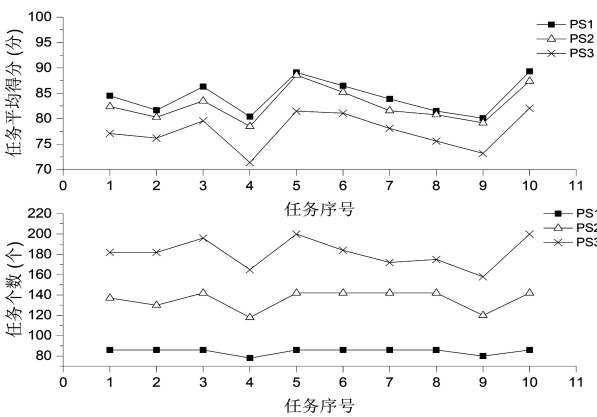


图4 不同参数下算法效果对比

第 2 个实验主要验证算法筛选出的学生的工作能力, 通过下面两种方法对比挑选出来的学生完成进阶任务的得分。方法 1 是所有任务都是直接根据学生在本课程所有知识块平均得分来分配, 方法 2 则是文中提出的通过计算每个任务在其关联知识块上面的得分

来挑选学生。实验中同样任意挑选某门课程的10项进阶任务,然后通过上述两种方法来挑选学生完成这10个进阶任务,接着对进阶任务的完成效果进行评估打分,结果如图5所示。可以看出,文中算法筛选的学生在每项任务的平均得分都要高于方法1,并且10项任务的平均得分高出5%左右,能较好地满足构建资源的需求。

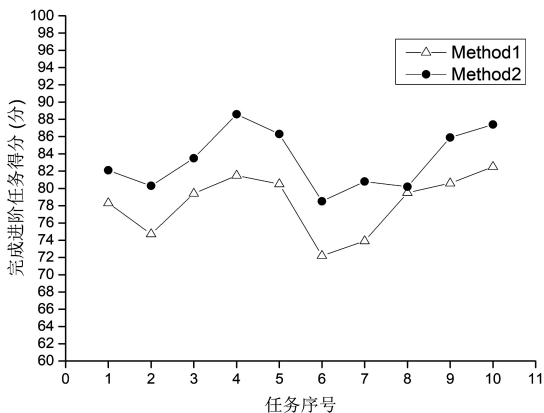


图5 筛选学生完成众包任务能力对比

### 3 结束语

在高校MOOC平台不断发展的背景下,利用众包策略构建MOOC学习平台的资源库,让学习者和教师构成学习共同体,在完成新知识学习的同时进行MOOC资源建设。为了解决这种知识密集型的众包任务,提出一种针对高校MOOC资源众包平台的任务分配方案,它包含学生的准入筛选、预期工作能力评估两个阶段。首先利用课程关联规则进行学生准入判断;其次利用知识关联进行能力评估与任务分配,最后对方案进行测试验证,结果表明该方案能够较好地保证学生的挑选和任务分配效果,促进构建更高质量的MOOC资源。

#### 参考文献:

- [1] MITROS P, SUN F. Creating educational resources at scale [C]//2014 IEEE 14th international conference on advanced learning technologies. Athens: IEEE, 2014: 16-18.
- [2] PORCELLO D, HSI S. Crowdsourcing and curating online education resources[J]. Science, 2013, 341(6143): 240-241.
- [3] 潘志宏, 万智萍, 谢海明. 大数据环境下高校智慧移动学习平台的构建研究[J]. 实验技术与管理, 2017, 34(4): 161-163.
- [4] WELD D S, ADAR E, CHILTON L, et al. Personalized online education—a crowdsourcing challenge[C]//Workshops at the twenty-sixth AAAI conference on artificial intelligence. Menlo Park, CA: AAAI, 2012: 1-31.
- [5] DANIEL F, KUCHERBAEV P, CAPPIELLO C, et al. Quality control in crowdsourcing: a survey of quality attributes, assessment techniques and assurance actions[J]. ACM Computing Surveys, 2018, 51(1): 1-40.
- [6] DANG D, LIU Y, ZHANG X, et al. A crowdsourcing worker quality evaluation algorithm on MapReduce for big data applications[J]. IEEE Transactions on Parallel & Distributed Systems, 2016, 27(7): 1879-1888.
- [7] IPEIROTIS P G, PROVOST F, WANG J. Quality management on Amazon mechanical turk[C]//ACM SIGKDD workshop on human computation. [s. l.]: [s. n.], 2010: 64-67.
- [8] HE S, SHIN D H, ZHANG J, et al. Near-optimal allocation algorithms for location-dependent tasks in crowdsensing[J]. IEEE Transactions on Vehicular Technology, 2017, 66(4): 3392-3405.
- [9] WANG X, ZHANG J, TIAN X, et al. Crowdsensing-based consensus incident report for road traffic acquisition[J]. IEEE Transactions on Intelligent Transportation Systems, 2018, 19(8): 2536-2547.
- [10] ZAIDAN O F, CALLISON-BURCH C. Crowdsourcing translation: professional quality from non-professionals[C]//Proceedings of the 49th annual meeting of the association for computational linguistics: human language technologies. [s. l.]: [s. n.], 2011: 1220-1229.
- [11] WANG X, WU W, HU Z. Evaluation of software quality in the TopCoder crowdsourcing environment[C]//2017 IEEE 7th annual computing and communication workshop and conference (CCWC). Las Vegas, NV: IEEE, 2017: 1-6.
- [12] BHANDARI P, WITHANA C, ALSADOON A, et al. Enhanced Apriori algorithm model in course suggestion system[C]//2015 international conference and workshop on computing and communication (IEMCON). Vancouver, BC: IEEE, 2015: 1-5.
- [13] VO B, HONG T P, LE B. DBV-Miner: a dynamic bit-vector approach for fast mining frequent closed itemsets[J]. Expert Systems with Applications, 2012, 39(8): 7196-7206.
- [14] LEE S, PARK S, PARK S. A quality enhancement of crowdsourcing based on quality evaluation and user-level task assignment framework[C]//2014 international conference on big data and smart computing (BIGCOMP). Bangkok: IEEE, 2014: 60-65.
- [15] HAN Y, WU W, PU Y. Task assignment of peer grading in MOOCs[C]//International conference on database systems for advanced applications. Berlin, German: Springer, 2017: 352-363.