

基于二进制微分进化算法的学习资源推荐方法

王文举 窦曙光 王鸾熠 姜中敏

(上海理工大学 出版印刷与艺术设计学院,上海 200093)

摘要 为解决当前方法学习资源推荐精度低、时效性不高、无法适应学习者动态调整学习目标的问题,提出了一种二进制的微分进化算法并以此为基础进行学习资源推荐的方法.该法对学习者和学习资源进行数学建模,依据推荐资源需要最佳匹配学习者需求的约束条件,将学习资源推荐转化成了最优化求解问题;根据课程知识点的逻辑组织结构图与学习资源的十字链表存储结构图,快速判定目标函数的关键参数值;采用所提出的一种二进制微分进化算法用于二进制数据表征的学习资源个体变量的变异、交叉、选择操作进行迭代运算直至满足条件找到学习资源最佳的选取组合.实验结果表明:所提出的学习资源推荐方法收敛速度优于以微粒群为基础的学习资源推荐方法,可方便学习者对学习目标范围动态灵活选取,所推荐的学习资源可完全满足不同学习者对学习资源难易度差异化的需求.

关键词 微分进化算法;学习资源;推荐服务;十字链表

中图分类号 TP391.9; G434 文献标识码 A 文章编号 1672-4321(2018)01-0125-07

Learning Resource Recommendation Method Based on Binary Differential Evolution Algorithm

Wang Wenju, Dou Shuguang, Wang Luanyi, Jiang Zhongmin

(Shanghai University of Science and Technology, Shanghai 200093, China)

Abstract To address the problem of low accuracy and low effectiveness of current methods which can't adapt to learners' dynamic adjustment of learning goals, a binary differential evolution algorithm is proposed as a basis for learning resource recommendation. The learners and learning resources are mathematically modeled. According to the recommended conditions of learning resources should best match the learners' requirements, the learning resources recommendation is transformed into the optimal solution problem. According to the logical organization chart of the course knowledge point and the storage structure diagram of the cross-linked list of learning resources, the key parameter value of the objective function can be quickly determined. A binary differential evolution algorithm proposed is used to make the binary individual variables of the learning resource variance, crossover, selection, until the configured conditions are met to find the best selection of learning resources. The experimental results show that the convergence rate and stability of the proposed method are better than the recommended method of learning resources based on the particle swarm optimization algorithm, which can facilitate the learners to choose the learning target range dynamically and make the recommended learning resources be satisfied to different learners' demand for differentiated learning resources.

Keywords differential evolution algorithm; learning resource; recommendation service; cross-linked list

学习资源的个性化推荐技术历经了用户建模研究、推荐内容研究、推荐策略研究三个发展阶段^[1].在用户建模方面,国外的相关研究有:2014年,Rtili, M.K 通过收集学习者的交互痕迹并进行处理对学

习者进行建模,以便通过一组相互作用的代理自动地提出适合他们需要的教育资源,但推荐精度不高^[2].2017年,Tarus, J.K 针对用户需求提出了一种基于用户偏好的异构教育资源推荐系统,系统实现

收稿日期 2017-10-10

作者简介 王文举(1979-)男,讲师,博士,研究方向:计算机辅助教学、机器学习, E-mail: wangwenju666@163.com

基金项目 新闻出版广电总局重点实验室 2017 年度招标课题(ZBKT201710);上海市属高校应用型本科试点专业建设项目(10-16-309-001)

高效但推荐质量不高^[3].国内:2013年,涂金龙使用新近打开的标签,构建用户(学习者)的兴趣模型,使学习资源推荐会随着用户兴趣的改变而发生变化,但该推荐方法的时效性不高^[4].2017年,张小雪采用学习者自我评价结合 Felder-silverman 量表的方式分析学习者的学习特征,从而构建出学习者的相关学习模型,有针对性地进行学习推荐.但该方法不仅需要学习者在线学习资源个性化推荐服务模型的建构,还需花费大量时间进行学习自我测评^[5].

在内容研究方面,国外的相关研究有:2013年, Salehi M 采用了学习资源的属性以及推荐过程中学习者访问资源的序列模式进行学习资源推荐.该方法引入了学习树(LT),考虑了资源的显式多属性、推荐精度高,但未考虑学习者的知识基础,推荐资源不能与学习者的接受能力相匹配^[6].2016年, Alinani K 利用本体论的领域知识和学习顺序存取模式挖掘系统来推荐学习资源,但推荐准确性不高^[7].国内:2014年,徐守坤对学习资源使用本体进行建构并使用语义推理丰富了资源推荐结果集,但难以对一个多语义联系学习资源进行准确的本体建构与初始化^[8].2016年,刘萌针对专业学习深入分析专业课程的知识,采用可达矩阵和并行拓扑排序方法进行学习资源路径的推荐,但该方法知识点梳理需人工处理具有一定的局限性并不适合用于在线资源推荐^[9].

在推荐策略上:2014年,张海东采用关联规则挖掘和相似度的方法,确定任意课程或资源之间的关联,并向中小学生进行学习资源推荐,但推荐的精度难以保证^[10].2016年,程春雷基于知识关系概念作为语义基本单元建立关系概念语义标识模型,用于web个性化学习资源推荐,但此方法在模型及参数的动态调整实现上较为困难^[11].朱夏提出了基于协同过滤构建个性化推荐学习资源推荐方法,该方法推荐精度和效率较高,但因依靠学习行为的历史纪录尚不能解决好冷启动问题^[12].2017年, Khosravi H 使用一种基于矩阵分解的协同过滤算法,为个别学生提供个性化的建议,以解决他们的兴趣和当前的知识差距,但算法的普适性不强^[13].2013年, Salehi M 将学习者的学习资源隐含或潜在属性的权重重视为遗传算法的染色体,然后根据历史评分对权重进行优化确定所推荐的学习资源,但系统的可伸缩性较差,实时性也不高^[14].2014年,杨超分别对学习者,学习资源进行特征描述,将学习资源推荐转化成了多目标,最优化问题,进而使用粒子群优化算法

进行求解形成最优推荐策略,但学习目标选取范围无法动态调整且算法复杂度较高,不适宜在线推荐^[15].

已有推荐策略存在学习目标动态调整困难、实时性较差的缺陷,本文结合上述推荐技术研究三个发展阶段的算法特性,对学习者、学习资源的内容进行建模,基于二进制微分进化算法提出了个性化的学习资源推荐方法.

1 问题表征

1.1 学习者模型构建

对于某门课程假设有几位学习者在共同在线学习.为了提高推荐精确度和服务效率,必须考虑学习者的认知能力,学习目标等核心特征要素.对学习者的进行建模所需的、具体的参数表征形式如下.

- (1) $S_j (1 \leq j \leq n)$ 表示第 j 位学习者.
- (2) $C_{A_j} (1 \leq j \leq n)$ 表示学习者 S_j 的认知能力.
- (3) $L_{T_j} (1 \leq j \leq n)$ 表示学习者 S_j 的学习目标.
- (4) $T_{\max_j} (1 \leq j \leq n)$ 表示学习者 S_j 预计学习某门课程在线资源时所投入时间的上限.

1.2 学习资源模型构建

假定一门课程拥有 M 个知识点, N 个在线学习资源.与推荐相关的学习资源模型构建的参数表征形式如下.

- (1) $k_{P_m} (1 \leq m \leq M)$ 表示某一门课程所具有的第 m 个学习知识点.
- (2) $C_{S_i} (1 \leq i \leq N)$ 表示某一门课程所拥有的第 i 个学习资源.
- (3) $R_i (1 \leq i \leq N)$ 表示第 i 个学习资源所涵盖的学习知识点. $R_i = \{r_{i1}, r_{i2}, r_{i3}, \dots, r_{im}\}, 1 \leq m \leq M$, 如果 $r_{im} = 1$ 表示第 i 个学习资源 C_{S_i} 是覆盖知识点 k_{P_m} 的, 否则 $r_{im} = 0$.
- (4) $S_i _D (1 \leq i \leq N)$ 表示第 i 个学习资源 C_{S_i} 的难度.
- (5) $T_i (1 \leq i \leq N)$ 表示学习第 i 个学习资源 C_{S_i} 所花费的时间.

1.3 最优推荐策略的目标函数表征

学习资源推荐方法的本质在于应该满足如下基本条件的最优化策略选取,即是一个多目标的最优化问题.

- (1) 推荐的学习资源所包含的知识点范围一定是学习者要掌握的知识点(即达到学习目标).
- (2) 学习资源知识点的难度要适合学习者的认

知能力.

(3) 推荐学习资源的总学习时间应该在学习者预计学习时间的上限之下.

据此目标函数可以表征为如下形式.

设变量 x_{ij} , $1 \leq i \leq N$, $1 \leq j \leq n$, 则有:

$$x_{ij} = \begin{cases} 1 \\ 0 \end{cases} . \quad (1)$$

如果学习资源 C_{S_i} 被选中推荐给学习者 S_j , 则 $x_{ij} = 1$; 否则 $x_{ij} = 0$.

根据上述约束条件(1) 构造目标函数 F_1 :

$$F_1 = \left(\sum_{i=1}^N \sum_{L=1}^M x_{ij} |r_{iL} - L_{T_{jL}}| \right) / \sum_{i=1}^N x_{ij} , \quad (2)$$

式中 N 表示一门课在线学习资源的数目, M 表示一门课程所具有的知识点数目. r_{iL} 表示第 i 个学习资源 C_{S_i} 是否覆盖知识点 k_{P_L} ; $L_{T_{jL}}$ 表示第 j 位学习者是否将第 L 个知识点作为自身的学习目标.

根据上述约束条件(2) 构造目标函数 F_2 :

$$F_2 = \left(\sum_{i=1}^N x_{ij} |S_i - D - C_{A_j}| \right) / \sum_{i=1}^N x_{ij} , \quad (3)$$

此式表示学习资源的难易度要尽量符合学习者的学习能力. $S_i - D$ 表示第 i 个学习资源的难度, C_{A_j} 表示

第 j 位学习者的学习能力.

根据上述约束条件(2) 构造目标函数 F_3 :

$$F_3 = \min \left(\sum_{i=1}^N t_i x_{ij} - T_{\max_j} \rho \right) , \quad (4)$$

函数 F_3 表示推荐资源的总学习时间满足学习者预计投入时间的上限要求. 若学习资源的学习时间超越学习者设置的学习时间上限 T_{\max} , 根据此式则 $F_3 = 0$, 系统凭此将对该学习资源不再进行推荐.

最终的目标函数可记为函数 F :

$$F(X_{ij}) = \min \left(\sum_{k=1}^3 W_k F_k \right) , \quad (5)$$

W_k 是 F_k 的加权系数.

1.4 算法参数的确定

(1) r_{iL} 、 $L_{T_{jL}}$ 参数的确定.

一门课程可有 M 个知识点. 根据学习规律, 知识点具体分布在各个章节中, 且相互之间有先后的学习顺序, 所以 M 个知识点是一个网络结构. 知识点的逻辑组织结构图见图 1. 根节点为课程, 树的第 2 层为课程的各章, 第 3 层为每章的各节, 第 4 层以下是具体的知识点按照学习顺序而组织的分布层次图.

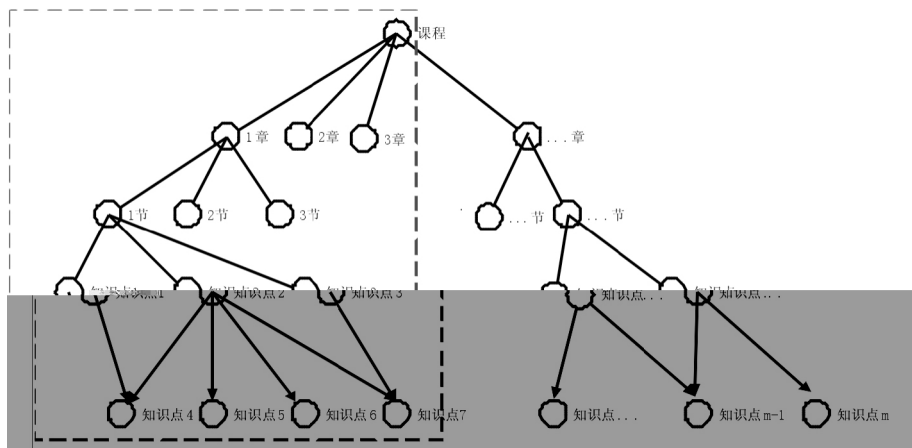


图 1 知识点的逻辑结构组织图

Fig.1 Logical structure organization of knowledge point

根据知识点的逻辑组织结构图, 规范整理学习资源采用十字链表构建学习资源存储结构图. 由于文章空间所限, 可取图(1) 虚线框标识部分相对应的十字链表构建如图 2.

结合各种格式的资源所覆盖知识点的范围, 在组织教学过程中 PPT、word 文本格式资源往往放在章层面上的资源节点内进行管理; 视频资源往往放在节层面上的资源节点内进行管理; 动画 flash 资源往往放在知识点所在的资源节点内进行管理. 根据知识点逻辑组织结构图, 以章 -> 节 -> 知识点的

顺序遍历基于十字链表的学习资源存储结构图, 分别依次对章 / 节 / 知识点中的知识点、学习资源进行顺序编号, 并统计好章 / 节 / 知识点所存放的学习资源编号的下限 N' , 编号的上限 M' . 由此种存储结构快可速判别该学习资源属性 R_i 集合中的 r_{iL} 具体的赋值. 例如学习资源 i 存放在某章 / 节所在层面的节点时, 该学习资源 C_{S_i} 将覆盖该节点章 / 节下面的所有知识点 (记录知识点编号 L'), 则该学习资源属性 R_i 集合中的 r_{iL} 自动赋值为 1. 当学习者在知识点逻辑组织结构图中点选学习内容. 例如章, 节及

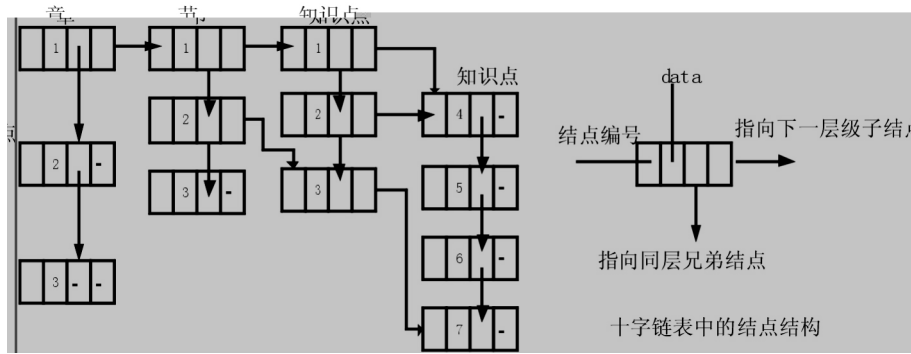


图 2 知识点的逻辑结构组织图

Fig.2 Logical structure organization of knowledge point

具体的各知识点时,软件系统将自动记录下点选节点以及点中所存放的所有学习资源编号的下限 N' 和编号上限 M' 。同时根据“ $L_{T_{jL}}$ 表示第 j 位学习者是否将第 L 个知识点作为学习的目标”的定义对 $L_{T_{jL}}$ 赋值。如果学习者在知识点逻辑组织结构图中已经点选某节点,则该节点及以该节点为前序所连的知识点相对应的 $L_{T_{jL}}$ 赋值为 1。将 R_{iL} 、 $L_{T_{jL}}$ 代入 F_1 , 用于最佳学习资源的推荐。通过上述方式,根据学习者的学习需求可实现学习资源动态推荐,缩小推荐资源的范围,节省运算时间。

(2) S_{i_D} 、 C_{A_i} 参数的确定。

S_{i_D} 表示第 i 个学习资源 C_{S_i} 的难易度。教师根据教学经验,将对所有学习资源进行难易度甄别。难易度一般分为容易、中等、难三个级别可用 1、2、3 整数数值分别相对应表征。可在链表中关于资源节点中 $data$ 域信息中给体现出来,见图 2。

C_{A_i} 表示学习者 S_i 的认知能力,可以分为低、中、高三个级别,由整数数值 1、2、3 分别相对应。学习者的自我认知能力,可在学习资源推荐之前由用户自我选择进行认知能力的自我定位。

2 基于二进制微分进化算法的学习资源的最优推荐

根据式(5),学习资源的最优推荐问题可以转化为基于约束条件的最优化求解问题,解决此类问题的算法有神经网络、遗传算法、微粒群算法等,相比上述算法微分进化算法具有突出的全局优化性能而且求解过程简单、受控参数少^[16],是解决学习资源推荐的最佳选择。

在微分进化算法中 $X(time)$ 表示迭代到第 $time$ 次时的种群规模,种群中的个体总数目记作 N_{sum} ,可行解空间的维数记作 D 。种群可表示为:

$$X(time) = \{x_1^{time}, x_2^{time}, x_3^{time}, \dots, x_{N_{sum}}^{time}\}, \quad (6)$$

经过第 $time$ 次迭代后,其中第 U 个个体可以表示为:

$$x_u^{time} = \{x_1^{time}, x_2^{time}, x_3^{time}, \dots, x_D^{time}\}, 1 \leq u \leq N_{sum}. \quad (7)$$

本文提出的基于微分进化算法进行学习资源的推荐方法中:一个个体就是一门课程一系列学习资源 C_{S_i} 向学习者 S_j 推荐与否的组合表示; D 表示与一门课程相关的所有学习资源统计数目,据此 $D = N$;第 U 个个体可以表示为:

$$x_u^{time} = \{x_{1j}^{time}, x_{2j}^{time}, x_{3j}^{time}, \dots, x_{Nj}^{time}\}, \quad (8)$$

其中 x_{ij}^{time} 表示学习资源 C_{S_i} 向学习者 S_j 推荐与否, $1 \leq i \leq N, 1 \leq j \leq n, x_{ij}^{time} = 0$ 表示学习资源 C_{S_i} 不向学习者 S_j 推荐; $x_{ij}^{time} = 1$ 表示学习资源 C_{S_i} 向学习者 S_j 进行推荐。

二进制差分进化算法具体实现步骤如下。

(1) 初始化参数:初始化迭代次数 $time = 0$, 最大迭代次数 N_{max} , 种群规模数 N_{sum} 。

(2) 生成初始种群:对第 u 个个体 $x_u^0 = \{x_{1j}^0, x_{2j}^0, x_{3j}^0, \dots, x_{Nj}^0\}, 1 \leq u \leq N_{sum}, 1 \leq i \leq N$, 第 i 个分量 x_{ij} 由 $rand$ 函数随机生成的 0 或 1 进行赋值,由此生成初始种群 $X(0) = \{x_1^0, x_2^0, x_3^0, \dots, x_{N_{sum}}^0\}$ 。将式(5)中的 F 作为适应度函数,由此计算每个个体的适应度值。

(3) 迭代次数 $time = time + 1$ 。

(4) 对种群中个体 u 依次编号,以编号 $u = 1$ 的个体为研究对象。

(5) 变异操作:在 $(time - 1)$ 种群所包含的个体中随机选择 3 个不同的个体 $x_{A_1}^{time-1}, x_{A_2}^{time-1}, x_{A_3}^{time-1}, 1 \leq A_1, A_2, A_3 \leq N$, 按照:

$$\bar{x}_u^{time-1} = x_{A_1}^{time-1} + F'(x_{A_2}^{time-1} - x_{A_3}^{time-1}), \quad (9)$$

执行变异操作 F' 称为缩放因子,取值可有 0、0.5、1、2 四种选择由 $randt$ 决定, $randt$ 是 $\{1, A\}$ 之间的随机整数, +, - 均为二进制加减操作。当 $F = 0.5$ 时,

第

(x)

时

体 C

ran

密

片

g (C
所

的收敛性
验证本
设置种群
是从“遗传学
B”学习研
标且其

系... 5
目... 计算...
能... “1”

习资源推荐算法的运算性能.基于同一硬件开发环境并使用基于微粒群优化推荐算法^[15]和本文提出的基于微分进化的推荐学习算法,分别做了 50 次实验进行了对比测试,并记录迭代次数与平均终止运算时间,见表 1 中的数据.

当学习资源选取范围(学习目标)具体到章/节,甚至具体到知识点时,基于粒子群优化算法的推荐算法并不支持,因为该推荐算法未采用层次逻辑组织结构和基于十字链表的存储结构,无法在众多知识点中选取学习目标的覆盖范围.当章/节/知识点为 50 个全部集中在“计算机图形学 B”第 3 章第 1 节中时,粒子群算法无法运行;而本文所提算法基于十字链表存储结构,可支持用户动态选取学习目标——当学习目标覆盖范围为章或节,甚至一个具体的知识点时,本算法可向学习者推荐最佳学习资源,见表 1.

当学习目标覆盖范围为课程全部知识点时,两种学习资源推荐方法均支持.对于两门不同课程,当待推荐学习资源数目由 100 增大到 200 时或由 300 增大到 420 时,相应的迭代次数与平均运算终止时间也相应增加.表明算法的复杂度变得越来越高.两种算法相比较,对于同样数量的学习资源(如学习

资源数量为 420 时),本文所提出的推荐算法找到最佳学习资源所花费的时间为 9394.327ms 是微粒群算法推荐学习资源所花费时间 12697.674ms 的 2/3,见表 1.由此可见在算法性能上本文所提出的学习资源推荐算法在运行速度上优于微粒群学习资源推荐算法.

3.3 推荐算法应用效果

为了评估本算法的实际应用效果,本研究对 2014 级印刷工程专业的 45 名学生进行了问卷调查.从学习资源选取的灵活性、推荐资源是否与学习目标主题相符、难易度是否适合自身需求定位、学习资源是否丰富,通过推荐的学习资源学习是否达到学习目标,这五项进行打分(每项满分 100 分),得到的平均分结果如图 5 所示.统计分数表明本推荐算法在推荐资源与学习目标的相关度,学习资源选取的灵活性、学习资源的丰富性、推荐资源难易度符合学习者自身定位需求上,学生给出的评价较好都在 90 分之上.说明本文提出的推荐学习资源算法学习目标资源的选取范围是灵活的,可大至整个课程,可小至具体的某个知识点,所推荐的学习资源针对性强,能够完全满足不同学习者对学习资源难易度差异化的需求.

表 1 学习资源算法性能比较

Tab.1 Performance comparison of learning resource algorithms

目标选取范围		算 法				
		基于粒子群优化算法		本文微分进化算法		
选取范围	覆盖的资源数目	运算时间/ms	迭代次数	运算时间/ms	迭代次数	
章/节/知识点	50	/	/	133.077	40	
全部 知识 点	面向对 象程序	100	534.324	82	408.327	62
	200	3832.452	278	2875.368	208	
	计算机 图形学	300	8593.456	430	6826.535	346
	420	12697.674	525	9394.327	405	

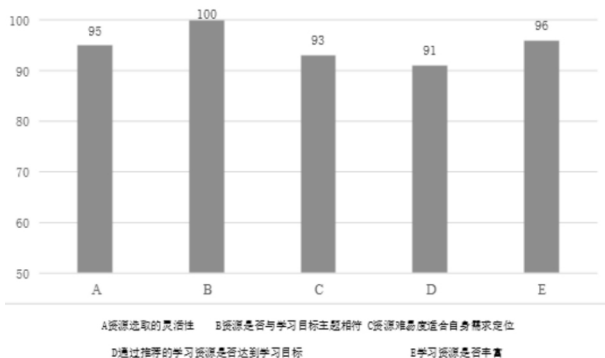


图 5 反馈结果统计图

Fig.5 Statistical chart of feedback results

4 结语

本文采用基于二进制的微分进化算法帮助学习者网络在线从多种多样的学习资源中挑选出适合自身需求的学习资源.实验结果表明该方法支持学习者对学习目标的灵活选取,运算速度快;推荐结果精度高,能够满足不同学习者对学习资源难易度差异化的需求.

参 考 文 献

- [1] 潘澄,陈宏. 我国学习资源个性化推荐研究进展[J]. 现代教育科学, 2015, (04): 31-34+37.
- [2] Rtili M K, Khaldi M, Dahmani A. Modeling Approach to Learner Based Ontologies for the Recommendation of Resources in an Interactive Learning Environments [J]. Journal of Emerging Technologies in Web Intelligence, 2014, 6(3): 340-347.
- [3] Tarus J K, Niu Z, Yousif A. A hybrid knowledge-based recommender system for e-learning based on ontology and sequential pattern mining [J]. Future Generation Computer Systems, 2017, 72: 37-48.
- [4] 涂金龙,涂风华. 一种综合标签和时间因素的个性化推荐方法[J]. 计算机应用研究, 2013, 30(4): 1044-1047, 1054.
- [5] 张小雪,张立国. 在线学习资源个性化推荐服务模型的构建[J]. 中国医学教育技术, 2017, 31(2): 172-176.
- [6] Salehi M. Application of implicit and explicit attribute based collaborative filtering and BIDE for learning resource recommendation [J]. Data and Knowledge Engineering, 2013, 87: 130-145.
- [7] Alinani K, Alinani A, Liu X, et al. Heterogeneous educational resource recommender system based on user preferences [J]. International Journal of Autonomous and Adaptive Communications Systems, 2016, 9(2): 20-39.
- [8] 徐守坤,孙德超,石林等. 基于语义推理的学习资源推荐[J]. 计算机工程与设计, 2014, 35(4): 1496-1501.
- [9] 刘萌,阎高伟,续欣莹. 基于知识点网络的自动化专业学习路径推荐[J]. 计算机仿真, 2016, 33(6): 180-184.
- [10] 张海东,倪晚成,赵美静. 面向基础教育阶段的教学资源推荐系统[J]. 计算机应用, 2014, 34(11): 3353-3356, 3364.
- [11] 程春雷,夏家莉. 关系概念的Web资源语义标识模型研究[J]. 计算机科学与探索, 2016, 10(08): 1092-1103.
- [12] 朱夏,宋爱波,东方. 云计算环境下基于协同过滤的个性化推荐机制[J]. 计算机研究与发展, 2014, 51(10): 2255-2269.
- [13] Khosravi H, Cooper K, Kitto K. RiPLE: Recommendation in Peer-Learning Environments Based on Knowledge Gaps and Interests [EB/OL]. CORR abs/1704.00556, 2017.
- [14] Salehi M, Kamalabadi I N, Ghouschi M B G. An Effective Recommendation Framework for Personal Learning Environments Using a Learner Preference Tree and a GA [J]. IEEE Transactions on Learning Technologies, 2013, 6(4): 350-363.
- [15] 杨超. 基于粒子群优化算法的学习资源推荐方法[J]. 计算机应用, 2014, 34(5): 1350-1353.
- [16] VStorn R, Price K. Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces [J]. Journal of Global Optimization, 1997, 11(4): 341-359.

(责任编辑 雷建云)