

基于知识图谱的个性化学习资源推荐研究

黄华升

(贺州学院数学与计算机学院, 广西 贺州 542899)

摘要: 学习者在使用网络学习资源过程中存在着“信息过载”及“知识迷途”的问题,无法实现学习资源的个性化需求。针对问题与现状,本文分析了学习者个性化学习资源需求特征,指出了知识图谱管理学习资源的优点,分析了基于知识图谱的方法技术生成学习路径图的推荐技术研究,提出基于知识图谱的个性化学习资源推荐系统架构。为建立个性化学习资源推荐系统提供一种技术方案。

关键词: 知识图谱; 个性化学习资源; 推荐系统

中图分类号: TP399 **文献标识码:** A

Study on Personalized Learning Resource Recommendation Based on Knowledge Graph

HUANG Huasheng

(School of Mathematics and Computer Science, Huzhou University, Hezhou 542899, China)

Abstract: In the process of using online learning resources, learners have the problems of information overload and knowledge trek, which cannot meet the personalized needs of learning resources. In view of the problems and current situation, this paper analyzes the characteristics of learners' personalized learning resource needs, points out the advantages of knowledge graph in managing learning resources, analyzes the recommendation technology of learning road map generation based on the method and technology of knowledge graph, and proposes the architecture of personalized learning resource recommendation system based on knowledge graph. The study provides a technical scheme for establishing personalized learning resource recommendation system.

Keywords: knowledge graph; personalized learning resources; recommendation system

1 引言(Introduction)

随着互联网的发展,教育教学环境发生了巨大变化,推动了我国的教学形态改革,混合教育学习、在线自主学习、翻转教学等各种形态的学习模式深入发展。发展过程中重要的推动力是网络教学资源建设,经过多年的建设,我国的网络教学资源已经非常丰富,各种网络平台推出形式多样的教学资源,为教师、学习者提供了极大的便利,学习变得泛在化和移动化,已然成为“人人皆学、处处能学、时时可学”的新形态。

有研究发现,在使用网络资源学习过程中学习者容易产生信息过载与知识迷途的现象^[1]。在面对海量的学习资源的选择中,检索出来的资源会包含大量类似内容,学习产生信息

过载,同时也面临着面对海量的学习资源,不知道资源的学习顺序,产生学习知识迷途。

针对网络学习资源使用过程中的信息过载及知识迷途问题,本文提出基于知识图谱对从学习目标开始到学习路径图生成技术的研究,以解决学习者在学习过程中产生的信息过载与知识迷途问题。

2 相关概念(Related terms)

2.1 个性化学习资源

在校学生的学习路径都是由教学计划规范好,然后按部就班的分阶段传授给学生。学生按照这种逻辑严密的教学方案学习是能够基本满足专业发展的最低水平的。但是每个学习者都存在个体差异(如专业背景、认知水平等),在学习的过

程中对按部就班的教学会有跟不上进度的，也会有不满足既定学习目标的等各种情况，普适性的教育资源不能满足所有用户的个性化学习需求，这个时候学习者就需要满足自己特定需求的学习资源。

专业的能力发展落实到职业岗位能力需求会产生各种不同的学习路径。如同样是计算机专业的学生，如落实到java工程师、前端工程师等不同的岗位，对学习内容的取舍，知识点的广度和深度各不相同，这样就会造就不同的学习路径，需要不同的学习资源。

综合来说，学习者的个体差异、专业发展、兴趣爱好诸多因素导致对学习资源产生不同的需求，需要推荐恰当的学习资源内容。

2.2 知识图谱

知识图谱是由一条条知识组成，每条知识表示为一个实体、实体属性、实体关系三元组知识体系，所有数据被组织成一张有向图^[2]。它的基本思想来源于五六十年代所提出的一种知识表示形式——语义网络，近年来在搜索领域的热门话题。在学习资源检索领域引入知识图谱，可以把专业领域内容所有的知识点组织成一个专业知识图谱，图谱中的每个节点都表示一个知识单元，单元与单元之间通过有向边指出知识单元间的逻辑关系，这样的知识图谱具有具有明晰知识结构，预测知识前沿，揭示知识关系等特性关系。

利用专业知识图谱指导学习者学习具有全局性、引领性的优势，它可以使得学习者从一个全局的高度全面了解专业的知识体系、课程之间的关联性，它就像一张课程内容的地图，每个知识点都有地图中的位置，学习者可以根据知识在地图中的位置，查找知识地图的路线指引，结合自己目前的知识状况及自己的个人需求，来确定下一步的学习计划。

2.3 推荐系统

海量的资源伴随着个性化需求，产生了信息过载及知识迷途，为了缓解这些问题大量的研究人员在研究如何恰当的推荐信息问题，如同商场的优秀的导购员，能够非常熟悉所导航商品的特性，还能记住顾客的特征及了解顾客的潜在需求，为顾客推荐最合适商品。推荐系统的实质就是在海量的信息中，找到用户感兴趣的资源，并生成个性化推荐结果，为用户提供差异性的搜索服务。

推荐系统主要技术是依靠推荐算法来实现资源的查找推荐，主要的推荐算法包括协同过滤算法、基于内容的推荐算法和混合推荐算法。协同过滤算法主要对用户行为数据进行度量 and 打分，根据得分的数据确认用户的偏好，同时查找一组类似的用户组集合，然后找出该集合中的用户组偏好度高

但目标用户还没有涉及到的资源推荐给目标用户^[3]。基于内容的推荐算法是建立项目资源的特征库，分析对比候选资源的特征与用户的兴趣特征，为用户推荐相关的资源。混合推荐算法是一种将多种基本推荐算法混合在一起的推荐算法。在数字化学习资源推荐技术中，最常用的是协同过滤推荐，其次为混合推荐。

3 个性化学习资源推荐系统(Personalized learning resource recommendation system)

利用知识图谱构建的个性化学习资源推荐系统分四个步骤构成，建立专业知识图谱、学科资源本体模型、生成专业发展学习线路，建立个性化学习资源推荐系统。

3.1 建立专业知识图谱

学科专业知识图谱的基本结构由媒体资源层、逻辑关系层和节点显示层构成，媒体资源层链接到具体的各种学习资源，逻辑关系层把资源整合形成学科知识本体模型，节点显示层把学科知识关系用图谱的形式可视化地展示给用户。

媒体资源层、逻辑关系层和节点显示层三个层完成了资源整合到图谱呈现的过程^[4]。媒体资源层接受来自本地资源、scrom资源包、网络中的结构化与半结构化数据进行标签、关键字提取操作，经数据加工后实现学科资源的数据预处理。逻辑关系层对预处理后的媒体数据进行逻辑定义、关系描述、标签标注等操作，建构学科的专业本体库，同时进一步对各学科本体库进行加工整合和合并，形成学科资源本体模型。节点显示层把学科本体模型，各学科本体间的语义关联映射到知识图谱中，以网状结构呈现学科知识节点间的关联，把学科知识关系用图谱的形式可视化地展示给用户，实现学科知识图谱的全局关系网络呈现。

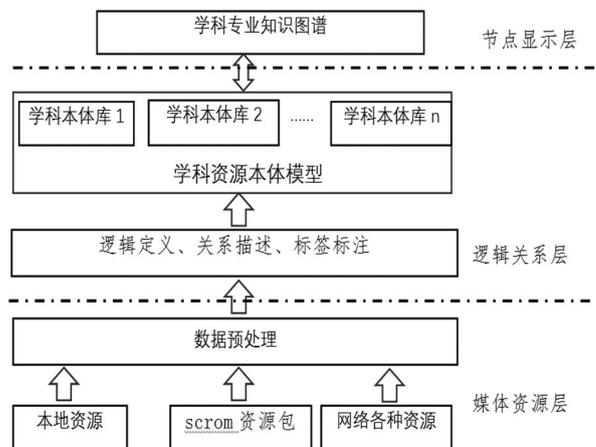


图1 专业知识图谱架构示意图

Fig.1 Schematic diagram of professional knowledge Graph

3.2 学科资源本体模型

在学科资源本体模型中采用知识点作为资源描述的基本

单位，每个知识点对应多个相关的学习资源，包括word、PPT、音视频等。若干个相关性较大的知识点组合成一个专题，而若干个专题组合成一个课程内容，各课程内容之间的关系按课程内容之间的相关性等参数来记录的。每个知识点包括一系列的属性标签，如媒体呈现方式(word、PPT、音视频等)、知识点要求的掌握程度(了解、理解、掌握)等。

知识点与知识点之间也包含各种关系属性，主要是相关性与先导性。相关性指知识点A与知识点B有某方面的关联，这种关联能使得在学习过程中增加新旧知识的联系，更容易理解。先导性是指，知识点A与知识点B的逻辑关系，若知识点A是建立在知识点B的基础上，称知识B为知识A的预备知识，要学习知识点A，须先学习预备知识点B；反之则是知识点A为知识点B的预备知识。

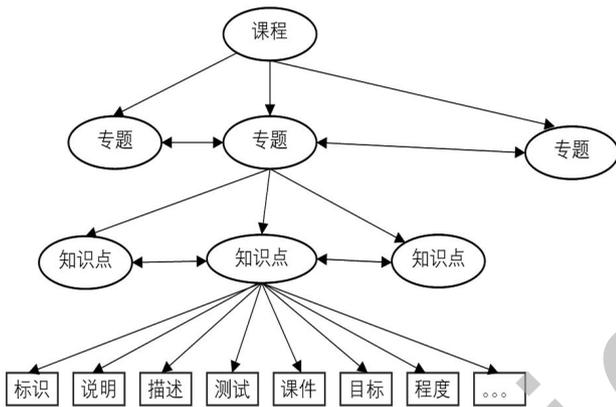


图2 学科资源本体模型

Fig.2 Subject resource ontology model

3.3 生成专业发展学习线路

协同过滤推荐系统使用算法通过对用户行为数据的挖掘找出用户的偏好特性，基于不同的偏好对用户进行群组划分，对用户推荐用户群组偏好相似的资源列表^[5]。协同过滤推荐的实现流程包含四个关键步骤：获取及表示用户信息、匹配学习资源、形成邻居和产生推荐。

表1 学习者—资源评分矩阵

Tab.1 Learner-Resource grading matrix

学习者 \ 学习资源	Res_1	Res_i	Res_j	Res_n
	User_1	R11	R1i	R1j
.....
User_i	Ri1	Rii	Rij	Rin
User_j	Rj1	Rji	Rjj	Rjn
User_n	Rn1	Rni	Rnj	Rnn

下面我们通过具体的步骤与算法来说明基于协同过滤算法的个性化推荐技术的实现^[6,7]。

第一步：将获取的数据以学习者—资源评分矩阵的形式表示。如表 1 所示，水平方向向 Res_n表示学习资源的个数，垂直方向User_1表示学习者的个数；第i行第j列的元素R代表第i个学习者对第j个资源的评分，R_{ij}的值表示学习者对资源的喜爱程度。

第二步：通过计算矩阵稀疏度来匹配学习资源。首先设定一个最小值 α ，稀疏度计算公式为： $Sparsity = E_n / U_n \times R_n$ ，其中E_n指学习者评价资源的数量，U_n指学习者的数量，R_n是学习资源的数量。只有当Sparsity > α 时，才能和学习者特征进行匹配从而为其推荐相应的学习资源，否则说明在线学习资源库评价矩阵过于稀疏，不能和学习者特征进行匹配。

第三步：寻找相似的学习者，形成邻居。计算学习者i和学习者j共同评分过的所有资源集合 $R_{ij} = R_i \cup R_j$ ，之后使用相似性度量方法计算学习者之间的相似性。基本思路是通过筛选得到学习者i和学习者j共同评分的所有资源集合，然后通过修正的余弦相似性度量方法计算两个用户之间的相似性。

第四步：生成推荐。计算邻居集中学习者i对学习资源s的预测评分值R_{i,s}，取R_{i,s}值最大的前N个资源推荐给学习者。

3.4 基于知识图谱的个性化学习资源推荐系统架构

所构建的基于知识图谱的个性化学习资源推荐系统架构主要提供个性化学习推荐和资源管理两个功能模块，总体架构如图3所示^[8]。学习单元测评用于诊断学习者当前的认知水平，主要借助习题测试、调查问卷来实现；学习者信息数据记录主要包括学习者学习记录、学习策略，以及学习路径等用户的特征信息。协同过滤推荐模块获取学习者当前的认知水平信息，与邻居用户群中提取学习记录集，结合知识图谱的知识逻辑关系，产生个性化资源推荐列表，并将资源推荐通过可视化的接口推送到用户页面。管理员或教师用户登录系统后，通过管理功能实时更新学习资源信息，如上传和编辑资源、修改学习资源的策略关系等。

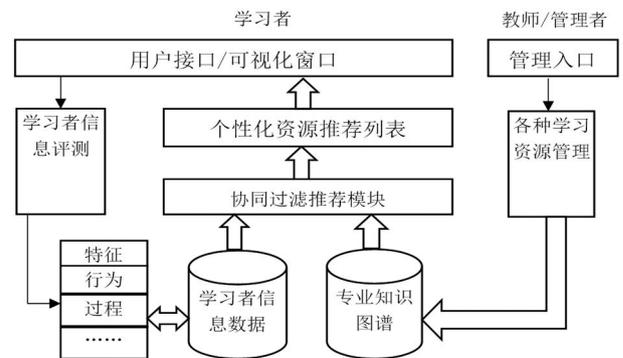


图3 个性化学习资源推荐系统架构

Fig.3 Architecture of personalized learning resource recommendation system