

基于知识点网络的自动化专业学习路径推荐

刘 萌, 阎高伟, 续欣莹

(太原理工大学信息工程学院, 山西 太原 030024)

摘要: 推荐学习路径的目的在于引导学习者达成学习目标, 避免学习过程中“迷失”方向。为指导自动化专业学习, 以提高相关专业学习者的学习效率, 优化学习效果, 提出一种基于知识点网络的学习路径推荐方法。首先依据该专业的知识点及其知识关联构建有向知识点网络, 并以邻接矩阵存储网络的拓扑结构信息。然后基于目标知识, 通过可达矩阵与并行拓扑排序算法实现学习路径的提取与优化, 保证学习者“按需所取”的学习知识。最后针对选取的某一知识点推荐学习路径, 实验结果表明上述方法具有可行性。

关键词: 自动化专业; 知识点网络; 学习路径; 可达矩阵; 并行拓扑排序

中图分类号: TP391 **文献标识码:** B

Learning Path Recommendation for Automation Discipline Based on Knowledge Points – Based Network

LIU Meng, YAN Gao – wei, XU Xin – ying

(College of Information Engineering, Taiyuan University of Technology, Taiyuan 030024, China)

ABSTRACT: The purpose of learning path recommendation is to guide learners to achieve their learning goals, and to avoid learners lost in the learning process. In order to improve the efficiency of professional learning for automation discipline and optimize learning results, a learning path recommendation method via knowledge points – based network of automation discipline is presented. A directed network is established by the learning sequence of two knowledge points at first, and the topology of the network is stored in an adjacency matrix. Then the path for different learning objectives can be extracted depending on the reachability matrix of the network, and optimized through the parallel topological sort algorithm. After that the path can ensure learners fulfill their learning needs. Experiment results show that the proposed method is feasible.

KEYWORDS: automation discipline; knowledge points network; learning path; reachability matrix; parallel topological sort algorithm

1 引言

随着知识经济的发展, 当今社会对知识掌握程度提出了更高的要求: 知识的专业化、综合化与结构化。而学习效果的优劣既受到所获取的学习资源的影响, 更与学习活动中知识的序列设计密切相关。为此, 国内外学者们展开了对学习路径推荐的研究以期提升学习者的学习效果。Ahmad 介绍了利用蚁群优化算法与概念图结合思想为不同的学习群体推荐合适的学习路径^[1]; Hsieh 探讨了基于 Apriori 算法挖掘候选学习资源并提取其关键词构建知识关联结构, 从而利用形式概念分析决定学习路径上知识的排列^[2]; Lee 等通过追

踪优秀学生的学习日志从中找到使用最频繁的学习序列作为学习路径^[3]; 陈其晖等由交换子和交换序的概念对微粒群算法进行改进, 实现了学习状态空间中学习路径的最优化控制^[4]; 李孟等针对参与测试的学习者为其生成知识地图总路径集合, 设计基于反馈的、满足全局优化的学习路径^[5]; 林海平等采用本体论并结合知识结构图提出基于在线学习过程动态评估及拓扑排序的学习内容生成方法^[6]。文献梳理发现, 目前的学习路径推荐研究多是针对网络学习环境下, 海量学习资源导致的“认知过载”和“学习迷航”等问题, 不能较好地满足学习者对专业知识的学习需求。原因在于这些路径推荐方法所提供的是近似学习路径, 而在专业知识学习中, 明确的知识关联结构与学习顺序对学习具有重要影响。

收稿日期: 2015-09-15

知识点作为描述教学领域知识的完整教学单元,包括理论、原理、定义、范例和结论等^[7],充分挖掘和利用知识点联系是构建学习计划的必要考量。由于知识点间普遍存在支持与被支持联系且呈现出一种网络结构^[8],因此采用网络形式对知识点进行组织和管理能较大幅度的反映知识的本来结构。相应地随着知识点关联复杂性的提高,学习者依靠自身知识水平获得所需学习资源的难度也会增大。本文为提高自动化专业知识学习质量与效率,基于知识点网络的拓扑信息,首先采用图论中图的连通性分析方法,引入可达性矩阵以减小知识搜索空间,再通过有向图的并行拓扑排序算法使推荐的专业学习路径兼顾原有知识结构的同时实现路径的有序化,从而提供科学的学习指导。

2 学习路径推荐策略

基于知识点网络推荐自动化专业学习路径,首先要解决的是如何构建相应的有向知识点网络,以期科学客观地反映知识体系中的内在知识关联;其次是采用何种方法从复杂的知识点网络中提取出与目标知识点强相关的前期先学知识点并组成学习路径;最后确定如何对获取的学习路径进一步优化,得到最终的推荐学习路径。

2.1 有向知识点网络

知识点是学习活动中传递知识信息的基本单元,分布于课程的各个章节。知识点选取的基本原则是单独的知识点能体现知识内容本身的局部完整性,而全部知识点的集合又能保证专业知识体系的全局完整性。知识点关系是连接知识点的纽带,使分散的知识点形成关联的知识结构,为便于梳理知识关系将其划分为两类:先学关系与区域关系。先学关系表述知识点内容的前后衔接顺序,区域关系则反映知识点联系的紧密程度。若干知识点按其先学关系可构成有向知识点网络,知识点网络由区域关系又可进一步划分出不同的知识子网络。以太原理工大学自动化专业的教学大纲为基础,选取该专业35门课程中1422个知识点与4972条知识点先学关系构建有向知识点网络。

2.2 学习路径提取

任何一个知识点都不是孤立的,而是与若干其它知识点相关联。因而目标知识点的学习需要一定量的前期先学知识点的支持。学习路径提取的关键在于找出那些需要在学习目标知识点前开展学习的知识点,在网络中表现为有可达路径到达目标知识点,即依据先学关系可以通往目标知识点,并且这些知识点构成一个先学知识集合。先学知识集合与目标知识点组成知识群集合,集合中任意两元素间存在的先学关系组成先学关系集合。这两个集合最终构成了目标知识点网络,也是由知识点网络中提取出的目标知识点的学习路径。

2.3 学习路径优化

学习路径优化分为冗余归并和路径序化两部分。由于知识点关系很复杂,知识点间的可达路径不一定唯一。本文

在保证不阻断知识传递关系的前提下,按照知识点间的可达路径分布情况将学习路径进行冗余归并,以降低学习路径的交叉程度。合理遍历学习路径上的知识点是路径序化的主要目的,这里需要考虑的问题是如何将学习任务分步同时尽量避免产生学习断层。针对这个问题,本文引入有向图的并行拓扑排序算法序化学习路径,得到一个并行学习集合链。

2.4 相关定义

2.4.1 知识点网络的描述

定义1:(知识点) 用一个四元组(ID , $Subject$, $Course$, $Content$)表示知识点 p_i 的不同性质属性。 ID 是知识点的唯一标志号,其值与知识点的下标 i 一致; $Subject$ 为知识点的主题,是对知识点的一种高度概括表述; $Course$ 代表知识点的所属课程; $Content$ 叙述知识点包含的详细知识内容。

定义2:(先学关系) 知识点内容上的直接前后衔接顺序用先学关系 $FR(p_i, p_j)$ 表示。学习知识点 p_j 前需先学习知识点 p_i ,则 $FR(p_i, p_j) = 1$ 。

定义3:(区域关系) 知识点联系的紧密程度用区域关系 $CR(p_i, p_j)$ 表示。由于课程内部知识点关联性较强,设定隶属相同课程的知识点具有区域关系。两个不同的知识点 p_i, p_j 的 $Course$ 属性一致,则 $CR(p_i, p_j) = 1$ 。

定义4:(有向知识点网络) 知识点作为网络节点,知识点先学关系构建网络的有向连边,二者共同组成知识点网络 G 。

知识点网络的拓扑结构可抽象为图 $G = (V, E)$,其中 $V = \{v_1, v_2, \dots, v_n\}$ 是网络的知识节点集合, $E = \{(i, j) | i, j = 1, 2, \dots, n\}$ 是网络中知识节点连边的集合, (i, j) 对应先学关系 $FR(p_i, p_j) = 1$ 。在网络构建过程中设置知识关联权重为1,则知识点网络 G 的拓扑结构信息以邻接矩阵形式存储为

$$A = (a_{ij})_{n \times n} \in \{0, 1\}^{n \times n} \quad (1)$$

$$a_{ij} = \begin{cases} 1, & (i, j) \in E \\ 0, & (i, j) \notin E \end{cases} \quad (2)$$

其中 n 为网络节点数目1422。

2.4.2 知识点网络的路径

定义5:(可达路径) 知识点 p_i 到知识点 p_j 的有向路径称为知识点 p_i 到知识点 p_j 的可达路径,前者是后者的先学知识,每条可达路径的长度取决于路径中包含的先学关系的数目。

可达矩阵 $L = (l_{ij})_{n \times n}$ 用于表示网络中任意两知识节点间的可达情况,元素 l_{ij} 满足下列条件

$$l_{ij} = \begin{cases} 1, & r_{ij} = 1 \\ 0, & r_{ij} = 0 \end{cases} (i, j = 1, 2, \dots, n) \quad (3)$$

其中 $r_{ij} = 1$ 代表知识节点 v_i 与节点 v_j 间存在可达路径; $r_{ij} = 0$ 两节点间不存在可达路径。任意两知识点是否存在可达路径由Warshall算法求得。已知网络的可达矩阵中某一元素 $l_{ij} = 1$,若 $FR(p_i, p_j) = 1$,则两知识点间存在一条长度为1的直接可达路径,若改变先学关系使 $FR(p_i, p_j) = 0$, l_{ij} 的值

保持不变,则两节点间还存在一条长度大于1的间接可达路径。

定义6:(冗余归并)两知识点 p_i 与 p_j 间同时存在直接可达路径和间接可达路径,则直接可达路径成为冗余路径,删除相应的先学关系 $FR(p_i, p_j) = 1$,即为路径的冗余归并。

2.4.3 有向图并行拓扑排序的基本概念

在相关序关系论述中,反链是偏序集中的一个子集,该子集中的任何两个元素是无关(即没有偏序关系)。

定义7:(并行集合)设有向图 $K=(N, \leq)$, N 为图中顶点的集合, \leq 是由图中顶点间的边关系组成的一个偏序集合,对于 N 集合的所有反链集合 $\{s_1, s_2, \dots, s_m\}$ 中的任意两个反链 s_i 和 s_j ,如果满足条件:①设 $R(R \neq \emptyset)$ 为 K 中 s_i 到 s_j 边的集合 $\langle s_i, s_j \rangle, R \subseteq \leq s_i \times s_j$;② s_j 到 s_i 边的集合 $\langle s_j, s_i \rangle = \emptyset$;③ $s_i \cap s_j = \emptyset$,则称 s_i 与 s_j 为两个并行集合。

定义8:(并行集合链)设有向图 $K=(N, \leq)$, N 集合中所有反链集合 $\{s_1, s_2, \dots, s_i, \dots, s_m\}$ 的一个子集 $S=\{s_i, s_j, \dots, s_k\} (1 \leq i < j < k \leq m)$ 中,如果满足条件:①相邻的两个反链 s_i 和 $s_{i+1} (i=1, 2, 3, \dots, k-1)$ 分别属于两个不同的并行集合;② S 中所有反链元素的并集与 N 等价,则称 (S, \leq) 为 (N, \leq) 上的一个并行集合链,其中 $\leq s = \cup \langle s_i, s_{i+1} \rangle (i=1, 2, 3, \dots, k-1)$ 。

3 路径推荐算法

3.1 相关符号描述

表 FIDL—存放所有知识点

p—目标知识

表 TFRL—存放知识点间的所有先学关系

表 EFRL—存放提取出的先学关系

表 RFRL—存放冗余归并后保留的先学关系

$O(v_d)$ —目标知识节点的先学知识集合

O —目标知识节点的知识群集合

表 LIDL—存放 O 集合对应的知识点

表 LCRL—存放 O 集合对应知识点间区域关系

H —先学关系集合

G_d —目标知识点网络

PLSD—并行集合链

3.2 算法描述

输入: FIDL, p

输出: PLSD, LCRL

步骤1:由表 TFRL 和表 FIDL 中的数据构建知识点网络 $G=(V, E)$, G 的邻接矩阵为 A 。

步骤2:根据 p 和定义1 查询表 FIDL,找出目标知识点及其 ID 编号,以及对应的目标知识节点 v_d 。

步骤3:利用 Warshall 算法求得知识点网络的可达矩阵 L ,获得先学知识集合 $O(v_d) = \{v_j | l_{jd} = 1, p_d, v_j \in V\}$ 以及知识群集合 $O = \{v_d\} \cup O(v_d)$ 。

步骤4:扫描表 TFRL 提取先学关系集合 $H = \{FR(p_a,$

$p_b) = 1 | p_a, p_b \in O\}$,将集合 O 中节点对应的知识点存入表 LIDL,再对该表中知识点的 ID 重新编号为1至 $|O|$,集合 H 中元素相应改变并存入表 EFRL。

步骤5:查询表 LIDL,依据知识点的 Course 属性划分出区域关系,并存入表 LCRL。

步骤6:由表 LIDL 和表 EFRL 构成目标知识点网络 $G_d = (O, H)$, G_d 的邻接矩阵为 A_d ,可达矩阵为 L_d 。

步骤7:对目标知识点网络 G_d 进行冗余路径归并,从而约减集合 H 中的部分先学关系,集合 H 中保留下来的先学关系存入表 RFRL。

冗余归并过程伪代码描述:

1) 初始化: G_d 的连边数目 $e = |H|$,邻接矩阵 A_d ,可达矩阵 L_d 。

2) DO

3) 非重复的删除关系 $FR(p_i, p_j) = 1 \in H$,即

$$A'_d \leftarrow A_d (a_{ij} = 0);$$

4) $e = e - 1$;

5) 由 $A'_d \rightarrow L'_d$;

6) if $L'_d = L_d$

7) $A_d \leftarrow A'_d$;

8) $H \leftarrow H \setminus FR(p_i, p_j) = 1$;

9) end if

10) UNTIL $e = 0$

步骤8:由表 LIDL 和表 RFRL 构成冗余路径归并后的新目标知识点网络 $G'_d = (O, H)$, G'_d 的邻接矩阵为步骤7 冗余归并后得到的新 A_d ,再将先学关系集合 H 清空。

步骤9:利用定义7 和定义8,对新的目标知识点网络 G'_d 进行并行拓扑排序。

并行拓扑排序过程伪代码描述:

1) 初始化: G'_d 的邻接矩阵为 A_d , $m = 0$

2) DO

3) 计算 A_d 的各行元素和 $\text{sum}(v_i)$;

4) $m = m + 1$;

5) $S_m = \emptyset$;

6) for $i = 1: |O|$

7) if $\text{sum}(v_i) = 0$

8) 反链集合 $S_m \leftarrow S_m \cup \{v_i\}$

9) end if

10) end for

11) $O \leftarrow O \setminus S_m$;

12) 依次删除 S_m 中各元素在 A_d 中对应行和列;

13) UNTIL $|O| = 0$

步骤10:集合 $S = \{S_1, S_2, \dots, S_m\}$,查询表 RFRL,使先学关系集合 $H = \{FR(p_i, p_j) = 1 | p_i \in S_{k+1}, p_j \in S_k, k = 1, 2, 3, \dots, m\}$,得到并行集合链 $PLSD = (S, H)$ 。

4 实验结果及分析

为了验证基于知识点网络的自动化专业学习路径推荐

方法的可行性与推荐效果。本文利用 Pajek 软件平台展示相关的实验结果。图 1 为所构建的有向知识点网络的全局拓扑结构,图中有向箭头标识知识点间的先学关系,颜色标识

知识点间的区域关系。实验选择自动化专业课程《运动控制系统》中具有代表性的知识“双闭环调速系统动态分析与设计”作为目标知识点进而推荐学习路径。

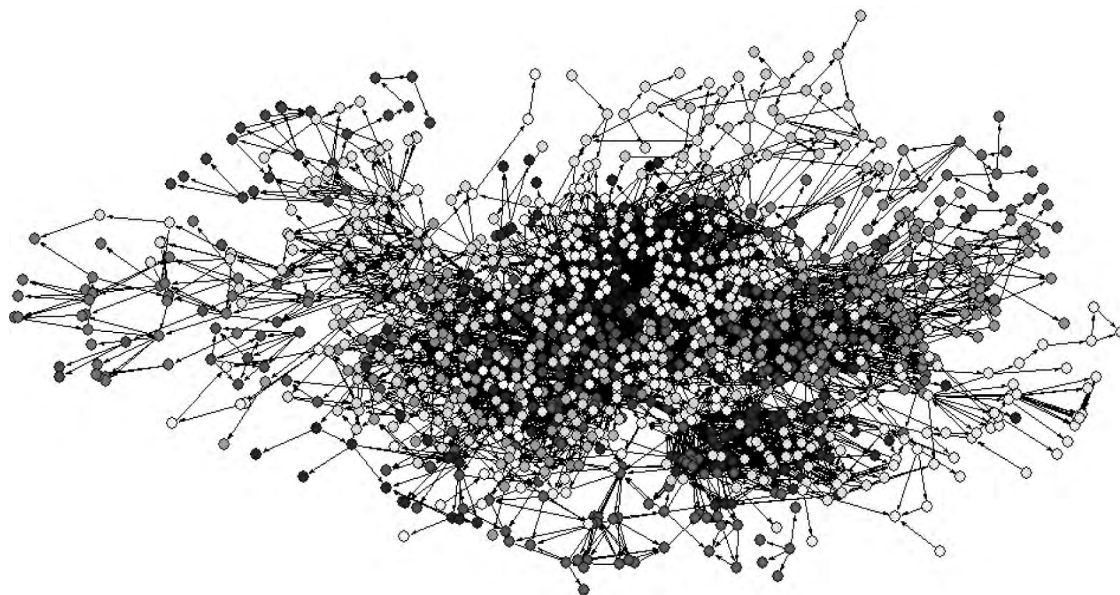


图 1 自动化专业知识点网络的全局拓扑结构

首先,依照前文提到的方法提取学习路径,结果表明提取的目标知识点网络规模相当大,涉及 15 门课程中的相关知识点。为能清楚地展示实验结果,摘除路径中涉及的数学和电路等基础课程,保留与目标知识点连接最紧密的《自动控制原理》、《电力电子技术》、《电力拖动与控制》、《运动控制系统》4 门课程中的 50 个先学知识节点和 78 条先学关系连边。图 2 给出了摘除部分先学知识节点后由图 1 所示的知识点网络中提取出的学习路径,并依据知识群集合中节点间的区域关系,用颜色标识不同课程中的知识点,数字标签为知识点的新 ID 编号。



图 2 目标知识点的学习路径提取

由于上述学习路径存在杂乱无序的缺点,能提供的指导信息仅限于知识间的局部关联,无法达成全局性的学习引

导。进一步对路径进行冗余归并,可简化路径并明确知识的传递关系。归并结果如图 3 所示,知识节点数目保持不变,共约减了 8 条知识点间的先学关系。



图 3 冗余归并后的学习路径

合理的知识结构和学习顺序直接影响学生的学习质量和效率。采用并行拓扑排序算法对学习路径序化处理得到一条并行学习链。目标知识点的学习路径序化处理结果如图 4,也是学习知识“双闭环调速系统动态设计与分析”的最终推荐路径。

并行学习链中的每一个知识点反链作为一个学习阶段,目标知识点的学习过程就被划分为不同的学习阶段,学生逐阶段学习,也可以参考知识点的课程分布,灵活的结合自身学习情况制定学习计划。此外,推荐的学习路径反映出知识

点的知识延伸情况,列出后继知识点,明晰当前知识对后续学习的影响,从而加强知识学习的深度。实验结果表明,基于有向知识点网络的目标知识学习路径推荐方法具有可行

性,借助知识关联推荐学习路径,将传统的知识传输转变为知识点网络构建,实现引导学习者达成学习目标并提升学习效率的目的。



图4 序化后的学习路径

5 结论

本文以有向知识点网络形式组织自动化专业知识体系中涵盖的知识资源,利用知识点关系由可达矩阵和并行拓扑排序算法实现目标知识的学习路径推荐。实验结果表明这种方法具有实用性,能较程度的降低学习活动中的盲目性和随意性,并针对性的指导自动化专业知识学习,提升学习者的学习效率。基于自动化专业知识点网络的学习路径推荐研究仍有很大空间。例如:引入知识节点间先学关系的不同权重,使推荐路径更贴近实际的学习需求;探究学习路径上知识节点缺失对目标知识点学习效果的影响,提供更多优化学习路径的考量因素等。

参考文献:

- [1] K Ahmad, B I Maryam, A E Molood. A novel adaptive learning path method[C]. E - Learning and E - Teaching (ICELET), 2013 Fourth International Conference on. IEEE, 2013: 20 - 25.
- [2] T C Hsieh, T I Wang. A mining - based approach on discovering courses pattern for constructing suitable learning path[J]. Expert systems with applications, 2010, 37(6): 4156 - 4167.
- [3] C H Lee, G G Lee, Y Leu. Analysis on the adaptive scaffolding learning path and the learning performance of e - learning[J].

WSEAS Transactions on information science and Applications, 2008, 5(4): 320 - 330.

- [4] 陈其晖, 凌培亮, 萧蕴诗. 基于改进微粒群优化的学习路径优化控制方法[J]. 计算机工程, 2008, 34(4): 190 - 192.
- [5] 李孟, 曹晟, 秦志光, 邱丹. 一种网络导航学习路径生成方法[J]. 电子科技大学学报, 2014, 43(4): 596 - 600.
- [6] 林海平, 檀小红, 申瑞民. 基于知识结构图的个性化学习内容生成算法[J]. 上海交通大学学报, 2010, 44(3): 418 - 42.
- [7] Miguel - Angel Sicilia, et al. Integrating descriptions of knowledge management learning activities into large ontological structures: A case study[J]. Data & Knowledge Engineering, 2006, 57(2): 111 - 121.
- [8] 谢深泉. 知识点及其网络的特性分析[J]. 软件学报, 1998, 9(10): 785 - 789.

【作者简介】



刘 萌(1989 -),女(汉族),河北徐水人,硕士研究生,主要研究领域为复杂网络的拓扑特性。
阎高伟(1970 -),男(汉族),山西洪洞人,教授,硕士研究生导师,主要研究领域为复杂网络控制。
续欣莹(1979 -),男(汉族),山西定襄人,副教授,硕士研究生导师,主要研究领域为自动化专业教学和科研工作。