

一种基于相似学习者判定的个性化学习路径推荐及验证方法

冯舒, 祝义, 宋媚, 居程程

引用本文

冯舒, 祝义, 宋媚, 居程程. 一种基于相似学习者判定的个性化学习路径推荐及验证方法[J]. 计算机科学, 2023, 50(11A): 220900067-10.

FENG Shu, ZHU Yi, SONG Mei, JU Chengcheng. [Personalized Learning Path Recommendation and Verification Method Based on Similar Learners Determination](#) [J]. Computer Science, 2023, 50(11A): 220900067-10.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于进程代数的Yahalom协议正确性的形式化验证](#)

Formal Verification of Yahalom Protocol Based on Process Algebra
计算机科学, 2021, 48(6A): 481-484. <https://doi.org/10.11896/jsjcx.200500074>

[基于进程代数的Otway-Rees协议的形式化验证](#)

Formal Verification of Otway-Rees Protocol Based on Process Algebra
计算机科学, 2021, 48(6A): 477-480. <https://doi.org/10.11896/jsjcx.200500072>

[分级论辩系统的逻辑研究](#)

On Logic of Graded Argumentation System
计算机科学, 2020, 47(5): 27-31. <https://doi.org/10.11896/jsjcx.200200052>

[航天器图像压缩小波变换的FPGA设计](#)

FPGA Design of Wavelet Transform in Spatial Aircraft Image Compression
计算机科学, 2010, 37(9): 261-263.

[构件系统建模及其动态演化一致性验证方法](#)

Dynamic Evolution Consistency Verification Method for Component System Modeling
计算机科学, 2017, 44(11): 80-86. <https://doi.org/10.11896/j.issn.1002-137X.2017.11.012>

一种基于相似学习者判定的个性化学习路径推荐及验证方法

冯舒^{1,3} 祝义^{1,2,3} 宋媚^{1,3} 居程程^{1,3}

1 江苏师范大学计算机科学与技术学院 江苏 徐州 221116

2 南京航空航天大学高安全系统的软件开发与验证技术工业和信息化部重点实验室 南京 211106

3 江苏省教育信息化工程技术研究中心 江苏 徐州 221116

(ffeng@jsnu.edu.cn)

摘要 基于相似学习者判定方法由于具有轻量级的特点而被广泛用于个性化推荐领域,目前一般采用协同过滤等机器学习的方法,但此类方法并不能保证判定过程的可解释性以及判定结果的可信性。针对这一问题,提出一种基于相似学习者判定的个性化学习路径推荐及验证方法,采用进程互模拟的方式研究相似学习者的判定过程。首先,扩展 CCS(Calculus of Communication System)的行为特性,提出 LR-CCS(Learning Resources-Calculus of Communication System),用于建模学习者的学习行为序列;其次,通过进程代数中互模拟理论判定学习者学习行为序列相似性,提出学习行为序列强(弱)互模拟关系判定算法进行互模拟关系判定;再次,使用互模拟验证工具 MWB(Mobile Workbench)验证学习者学习行为序列相似性,得到满足互模拟关系的候选推荐路径,以保证判定结果的正确性;最后通过一个基于相似学习者的推荐系统实例验证了该方法的有效性。

关键词: 学习行为序列相似性;进程代数;CCS;互模拟

中图法分类号 TP311

Personalized Learning Path Recommendation and Verification Method Based on Similar Learners Determination

FENG Shu^{1,3}, ZHU Yi^{1,2,3}, SONG Mei^{1,3} and JU Chengcheng^{1,3}

1 School of Computer Science and Technology, Jiangsu Normal University, Xuzhou, Jiangsu 221116, China

2 Key Laboratory of Ministry of Industry and Information Technology for Software Development and Verification Technology of High Security Systems, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China

3 Jiangsu Education Information Engineering Technology Research Center, Xuzhou, Jiangsu 221116, China

Abstract The similarity-based learner determination method is widely used in the field of personalized recommendation due to its light weight. At present, machine learning methods such as collaborative filtering are generally used. However, such methods cannot guarantee the interpretability of the determination process and the availability of the determination results. To solve this problem, a personalized learning path recommendation and verification method based on similar learner determination is proposed, which uses the method of process bisimulation to study the determination process of similar learners. Firstly, the behavior characteristics of calculus of communication system(CCS) are extended, and learning resources-calculus of communication system(LR-CCS) is used to model the learning behavior sequence of learners. Secondly, the bisimulation theory of process algebra is used to determine the similarity of learners' learning behavior sequences, and the algorithms for determining the strong(weak) bisimulation relationship of learning behavior sequence is proposed. Thirdly, the bisimulation verification tool mobile workbench(MWB) is used to verify the similarity of the learner's learning behavior sequence, and the candidate recommended paths which satisfy the bisimulation relationship are obtained to ensure the correctness of the judgment result. Finally, a case study of a recommender system based on similar learners verifies the effectiveness of this method.

Keywords Similarity of learning behavior sequence, Process algebra, CCS, Bisimulation

1 引言

研究人员的关注。人工智能技术应用于教育领域的本质是数据驱动的算法应用^[1]。在个性化推荐领域,文献[1]提到在真实的教育场景中,同一类型的学习者往往能达到近似的学习

目前,“人工智能+教育”的模式受到教育工作者以及

基金项目:国家自然科学基金(62077029);CCF-华为创新研究计划资助(CCF-HuaweiFM202209);南京航空航天大学基本科研业务费科研基地创新基金(NJ2020022);未来网络科研基金项目(FNSRFP-2021-YB-32);江苏师范大学研究生科研创新计划(2021XKT1384)

This work was supported by the National Natural Science Foundation of China (62077029), CCF-Huawei Populus Grove Fund (CCF-HuaweiFM202209), Open Project Fund of Key Laboratory of Safety-Critical Software Ministry of Industry and Information Technology (NJ2020022), Future Network Scientific Research Fund Project (FNSRFP-2021-YB-32) and Graduate Science Research Innovation Program of Jiangsu Normal University(2021XKT1384).

通信作者:祝义(zhuy@jsnu.edu.cn)

效果。基于此,很多研究者认为同一类型的学习者对目标学习者具有参考意义。其中,基于聚类以及协同过滤算法^[2-5]的个性化推荐方法应用广泛。此外,随着智慧教育越来越受到人们的重视,深度学习^[6]等新兴方法应运而生。此类基于教育大数据的人工智能方法由于其成本低、可移植性强而应用广泛,但也存在诸多问题:(1)“黑盒”算法导致教育有被算法控制的可能;(2)算法偏见导致教育弊端被放大;(3)教师过度依赖算法有导致专业技能丧失的风险。导致这些问题的根本原因是算法的输入和输出过程是黑盒的,很难从模型内部的数据交互及处理逻辑给出直观易懂的解释,缺乏数学理论支撑^[7]。同时,算法也不具备解释处理过程的能力,而适当的解释可以说服用户尝试学习该资源,增加用户的信任,并帮助用户做出更好的决策^[8]。一方面,目前的深度学习模型过于复杂,有太多的参数需要学习,需要花大量的时间。另一方面,相似性作为推荐的核心环节,一直是众多学者关注的重点研究方向^[9]。因此,本文针对基于相似学习者的推荐方法,主要从增加处理过程逻辑性的角度出发,从而保证判定过程的可解释性以及判定结果的可信性。

形式化方法可将可信性转化为形式化规范和约束,并通过模型检查等技术对其验证,从而保证其可信性^[10]。近年来形式化方法在验证可信性方面取得了许多成果。例如,Goodfellow等^[11]提出利用线性假设来解释对抗样本的存在性,并指出神经网络在高维空间中利用线性性质才是导致对抗样本存在的真正原因;Zhang等^[12]提出了一种更严格的线性逼近方法,用于卷积神经网络(CNN)的鲁棒性验证;Shi等^[10]提出一种基于形式化和逻辑推理方法的机器学习可解释性方法,用于解释随机森林的预测结果。进程代数是形式化方法中比较有代表性的建模与分析方法,经典进程代数包括 CCS(Calculus of Communication System)^[13],CSP(Communicating Sequential Processes)^[14]和 LOTOS(Language of Temporal Ordering Specification)^[15]等。其中通信演算系统 CCS 是 Robin Milner^[13]在 1980 年提出的一种描述并发交互

系统的方法,该方法能够对学习行为序列中事件的顺序、并发、同步等问题进行描述。因此,本文采用 CCS 对学习行为序列建模,为使用互模拟理论判定学习者学习行为序列的相似性提供模型支持,保证模型可信性。

互模拟^[16]是并发理论中最重要的概念之一,一般而言,若两进程满足互模拟关系,即可判定这两个进程是相似的。互模拟分为强互模拟以及弱互模拟^[17]。强互模拟要求行为要一一对应等价,但在某些情况下,无需考虑系统的内部动作,因此就产生了弱互模拟的概念。互模拟是判定进程相似性方面的有效研究方法,但它缺乏对学习资源信息的描述,而学习资源信息又是描述学习者学习行为序列不可或缺的元素。因此,本文在互模拟理论上增加了学习资源信息描述。

通过以上分析,本文针对基于相似学习者的推荐系统中的相似性判定过程可解释性差以及判定结果可信性差的问题,提出了一种基于相似学习者判定的个性化学习路径推荐及验证方法,技术路线如图 1 所示,包括建模部分、互模拟关系判定部分以及互模拟关系验证 3 部分。该方法首先扩展 CCS 的行为属性,提出学习资源-通信演算系统 LR-CCS,使其具有描述学习资源动作特性的能力;其次,扩展进程行为等价关系互模拟等价的行为特性,提出学习行为序列强互模拟以及学习行为序列弱互模拟方法;再次,通过验证工具 MWB 验证其相似性,若不满足互模拟关系则进行容错处理,修改 LR-CCS 模型;最后得到满足相似性的候选推荐路径。

本文以提供可靠的形式化模型为基础,采用逻辑推理的方式提高过程可解释性,通过相应的验证工具保证推荐结果的可信性,主要贡献包括:

- (1)扩展进程代数 CCS,对学习行为序列进行建模,为学习者相似性判定提供模型支持;
- (2)扩展互模拟理论,判定学习者学习行为序列相似性,为学习者相似性判定提供理论支持;
- (3)使用互模拟验证工具 MWB 验证学习者学习行为序列相似性,为学习者相似性判定提供工具支持。

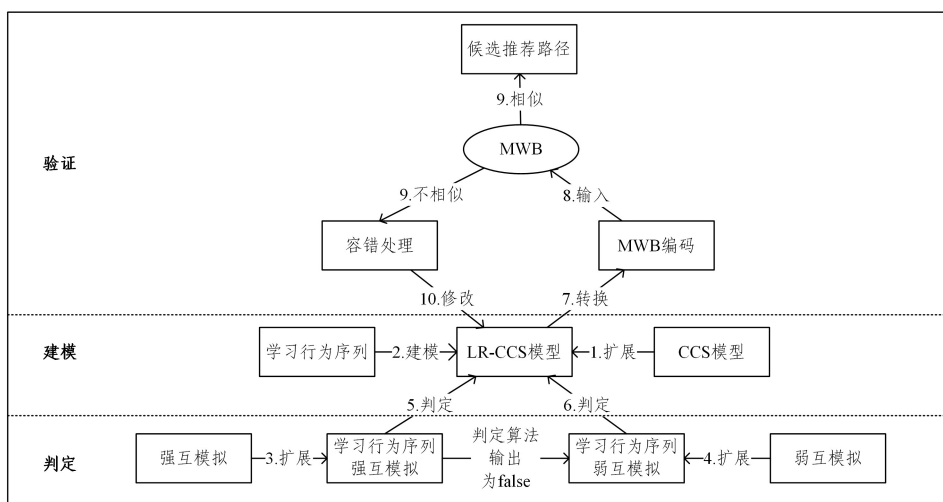


图 1 技术路线

Fig. 1 Technology roadmap

2 基本定义

有关学习资源以及行为的序列数据大量存在,由于其中蕴含对学习分析有益的信息,因此,将学习行为序列作为学

习者相似性分析的指标。在进行学习行为序列相似性判定时存在由于非关键行为导致的学习行为序列不一致的情况,因此,本章首先介绍了学习行为序列、关键学习行为序列的概念以及关键学习行为的诊断方法。

2.1 学习行为序列

具有 m 个学习资源的学习资源集 $Res = \{r_i | 0 \leq i \leq m, i \in N\}$, 具有 n 个学习行为的学习行为集 $Act = \{a_j | 0 \leq j \leq n, j \in N\}$, 由此可得学习资源动作集为 $RAct = Res \times Act = \{\langle r_i, a_j \rangle | r_i \in Res, a_j \in Act\}$ 是一个有序对的集合, 代表学习者在学习资源 r_i 上执行 a_j 等一系列学习行为。学习行为序列 LBS (Learning Behavior Sequences) 是由一组按时间先后顺序排列的学习资源及学习行为的有序对组成。因此, 在 LBS 中的任意元素 $\langle r_i, a_j \rangle$, 都有 $\langle r_i, a_j \rangle \in RAct$ 。

2.2 关键学习行为序列

2.2.1 关键学习行为序列定义

关键行为可以探索和解读学习者与学习环境交互内在机制, 并分析影响学习者绩效表现的重要学习过程^[17]。此外, 先点击学习资源 r_0 还是先下载学习资源 r_0 的顺序不同对分析学习者相似性没有太大意义, 因此对于仅是非关键行为不一致的行为序列, 也可判定其相似性。

具有 l 个关键学习行为的关键行为集 $KAct = \{b_k | 0 \leq k \leq l, k \in N, l \leq n\}$, $KAct \subseteq Act$, $|KAct| \leq |Act|$ 。特殊情况下, $\forall a_j \in Act, \forall b_k \in KAct$, 当 $j = k$ 时都有 $a_j = b_k$, 此时 $Act = KAct$ 。关键学习资源动作集为 $RKAct = Res \times KAct = \{\langle r_i, b_k \rangle | r_i \in Res, b_k \in KAct\}$, $RKAct \subseteq RAct$, 代表学习者在学习资源 r_i 上执行 b_k 等一系列学习行为。关键学习行为序列 (Key Learning Behavior Sequences) 是指在学习行为序列基础上只由关键行为构成的子序列。因此, 在 KLBS 中的任意元素 $\langle r_i, b_k \rangle$, 都有 $\langle r_i, b_k \rangle \in RKAct$ 。

2.2.2 关键学习行为诊断

有研究指出通过识别学习活动中的高频转换行为发现学习过程中的关键学习行为是目前应用比较广泛的学习行为诊断方法^[18], 因此, 本文将学习行为序列中的高频转换行为作为关键学习行为。

近年来, 滞后序列分析法 LSA (Lag Sequential Analysis) 逐渐成为分析在线学习行为序列的有效工具, 由 Sackett 于 1978 年提出, 主要用于检验一种行为发生以后另一种行为出现的概率以及是否具有显著性^[19]。LAS 可以统计学习行为之间的转换频次, 揭示学习过程中关键的学习行为序列^[20]。

滞后序列分析法首先要确定研究问题, 然后筛选行为数据, 进行行为编码, 接着通过 GSEQ (General Sequential Querier)^[21] 软件生成的调整后残差表绘制行为转换图, 解读行为模式图, 最后进行相应的分析。但本文主要是通过 GSEQ 软件得到学习者行为频次关系, 因此在进行分析的时候只需要得到行为转换频率表即可。主要过程包括: (1) 按 GSEQ 要求的格式输入所有行为编码; (2) 系统编译, 生成 MDS 文件; (3) 进行行为序列分析, 输出行为转换频率表。

3 学习资源-通信演算系统

3.1 LR-CCS 语法

定义 1 学习资源-通信演算系统 LR-CCS 可定义为:

$$P ::= 0 | \langle r, a \rangle. P | P + Q | P | Q | P \setminus RA | A | P[f]$$

其中, 0 代表空进程, 即 P 进程不执行任何操作; $\langle r, a \rangle. P$ 代表进程的前缀操作, 其中, $\langle r, a \rangle$ 表示进程在学习资源 r 上执行动作 a , a 包括输入动作 c 、输出动作 \bar{c} 以及内部动作 τ 这 3 类动作; $P + Q$ 代表进程的非确定性选择; $P | Q$ 代表进程的

并发; $P \setminus RA$ 代表限制, 产生进程 P 把学习资源动作集合 RA 中的动作隐藏起来; A 代表进程标识符, 表示递归定义的进程; $P[f]$ 代表换名, 将 P 中的学习资源动作按照 f 给出的映射关系转换为其他学习资源动作。

3.2 LR-CCS 的学习资源标号迁移系统

文献^[22]将 CCS 的语义定义为标号迁移系统 LTS (Labeled Transition Systems), 进程中一个事件的执行在标号迁移系统里可以看作一个迁移。我们在研究 LR-CCS 时引入了学习资源标号迁移系统 LRLTS (Learning Resources-Labeled Transition Systems) 的概念。

定义 2 一个通信演算系统 CCS 的语义描述为一个标号迁移系统 $LTS = (S, S_0, Act, \rightarrow)$ 。 S 代表状态集合; S_0 代表初始状态; Act 代表动作集合; \rightarrow 代表迁移关系, $S_1 \xrightarrow{\alpha} S_2$ 表示 S_1 所代表的进程执行完动作 α 变成 S_2 所代表的进程。

定义 3 一个学习资源-通信演算系统 LR-CCS 的语义描述为一个学习资源标号迁移系统 $LRLTS = (S, S_0, RAct, \rightarrow)$ 。 S 代表状态集合, 包含所有的状态; $S_0 \in S$ 代表初始状态; $RAct$ 代表学习资源动作集; $\rightarrow \subseteq S \times RAct \times S$ 表示迁移关系。

一个学习资源-通信演算系统被描述为一个学习资源标号迁移序列

$$\bar{\omega} = (\langle r_0, a_0 \rangle, \langle r_1, a_0 \rangle, \dots, \langle r_m, a_0 \rangle, \langle r_0, a_1 \rangle, \langle r_1, a_1 \rangle, \dots, \langle r_m, a_1 \rangle, \dots, \langle r_m, a_n \rangle)$$

3.3 LR-CCS 的操作语义

下面给出一组 LR-CCS 的操作规则。限于篇幅, 所有操作语义均省略了对称部分。

前缀算子: $\langle r, a \rangle. P \xrightarrow{\langle r, a \rangle} P'$ 表示学习资源动作顺序执行的情况。进程 P 执行 $\langle r, a \rangle$ 后演变为进程 P' , 其中 $\langle r, a \rangle \in RAct$ 。

$$\text{选择算子: } \frac{P \xrightarrow{\langle r, a \rangle} P'}{P + Q \xrightarrow{\langle r, a \rangle} P'} \text{ 与 } \frac{Q \xrightarrow{\langle r, a \rangle} Q'}{P + Q \xrightarrow{\langle r, a \rangle} Q'}$$
 表示进程 P 与进程 Q 之间存在非确定性选择。

$$\text{并发组合算子: } \frac{P \xrightarrow{\langle r, a \rangle} P'}{P | Q \xrightarrow{\langle r, a \rangle} P' | Q} \text{ 与 } \frac{Q \xrightarrow{\langle r, a \rangle} Q'}{P | Q \xrightarrow{\langle r, a \rangle} P | Q'}$$
 表示进程 P 与进程 Q 异步并发执行;

$\frac{P \xrightarrow{\langle r, a \rangle} P', Q \xrightarrow{\langle r, a \rangle} Q'}{P | Q \xrightarrow{\tau} P' | Q'}$ 表示进程 P 与进程 Q 同步并发执行。

$$\text{限制算子: } \frac{P \xrightarrow{\langle r, a \rangle} P'}{P \setminus RA \xrightarrow{\langle r, a \rangle} P' \setminus RA}$$
 表示学习资源动作集合 RA 中被隐藏的学习资源动作不被系统外所见, 本文中指隐藏非关键学习行为。

$$\text{换名算子: } \frac{P \xrightarrow{\langle r, a \rangle} P'}{P[f] \xrightarrow{f(\langle r, a \rangle)} P'[f]}$$
 表示将换名函数作用在进程 P 上所得到的进程, 表明换名对转换关系的影响。

$$\text{常量算子: } \frac{P \xrightarrow{\langle r, a \rangle} P'}{A \xrightarrow{\langle r, a \rangle} P'}$$
 表示进程 P 执行 $\langle r, a \rangle$ 后演变为进程 P' , 则用 P 定义的进程常量 A 也有这样的迁移。

3.4 精化关系

下面证明 CCS 的语义模型是 LR-CCS 的子语义模型。

我们在之前的研究中定义了进程代数的精化关系^[23-24],参考这种精化关系,本文定义了 CCS 的语义模型与 LR-CCS 语义模型之间的精化关系。首先定义 LR-CCS 语义模型,其次通过定理证明通信演算系统 CCS 是学习资源-通信演算系统 LR-CCS 的子语义模型。

定义 4 假设任意一个 A 类具有语义模型 M_A 的通信演算系统 P_A 能够通过一个精化关系得到另一个 B 类具有语义模型 M_B 的通信演算系统 P_B ,那么称语义模型 M_B 是语义模型 M_A 的子语义模型^[25], P_A 到 P_B 的精化关系如下:

$$\frac{P_B \text{ Sat } S_B \text{ in } M_B}{P_A \text{ Sat } S_A \text{ in } M_A} [P_B \sqsubseteq P_A]$$

其中, S_A 是进程 P_A 的语义, S_B 是进程 P_B 的语义。

定义 5 一个学习资源-通信演算系统 CCS 可接受的语言为其所有学习资源标号迁移序列的集合。

定理 1 通信演算系统 CCS 的语义模型是学习资源-通信演算系统 LR-CCS 的子语义模型。

证明:令 P_{CCS} 是一个通信演算系统,由定义 2 知 CCS 是一个标号迁移系统 $LTS = (S, S_0, Act, \rightarrow)$,假设其可接受的语言为 L ,根据通信演算系统构造一个学习资源-通信演算系统 P_{LR-CCS} ,由定义 3 可知, P_{LR-CCS} 是一个学习资源标号迁移系统 $LRLTS = (S, S_0, RAct, \rightarrow)$,假设其可接受的语言为 L' ,此时,在 L 中任取一个标号迁移序列 $\theta = (a_0, a_1, a_2, \dots, a_n)$,在 L' 中都有一个唯一的学习资源标号迁移序列

$$\bar{\omega} = (\langle r_0, a_0 \rangle, \langle r_1, a_0 \rangle, \dots, \langle r_m, a_0 \rangle, \langle r_0, a_1 \rangle, \langle r_1, a_1 \rangle, \dots, \langle r_m, a_1 \rangle, \dots, \langle r_m, a_n \rangle)$$

与之对应,因此 $Act \subseteq RAct, P_{LR-CCS}$ 到 P_{CCS} 存在精化关系。

$$\frac{P_{CCS} \text{ Sat } S_{CCS} \text{ in } M_{CCS}}{P_{LR-CCS} \text{ Sat } S_{LR-CCS} \text{ in } M_{LR-CCS}} [P_{CCS} \sqsubseteq P_{LR-CCS}]$$

因此通信演算系统 CCS 的语义模型是学习资源-通信演算系统 LR-CCS 的子语义模型,证毕。

由定理 1 可知,LR-CCS 上所有与功能性质相关的模型检测都可以通过 CCS 的模型检测工具完成,而 LR-CCS 上仅需要完成在学习资源行为属性下互模拟关系的判断。

4 学习行为序列互模拟

学习行为序列互模拟是由经典的进程代数互模拟理论引入到学习行为序列相似性判定场景下得到的。在传统的功能模型中,只需要考虑变迁时的动作是否互模拟即可,而本文判定的是学习行为序列之间的互模拟关系,需要考虑学习资源动作匹配。因此本文在经典的互模拟定义上加入了学习资源属性,提出包含行为以及学习资源信息的学习行为序列强互模拟关系以及学习行为序列弱互模拟关系。在此基础上将学习行为序列相似性分为学习行为序列强相似性、学习行为序列弱相似性。

定义 6(学习行为序列强相似性) 具有 n 个学习者的学习者集合 $Stu = \{stu_i \mid 0 \leq i \leq n, i \in N\}$, $\forall stu_i, stu_j \in Stu, i \neq j$,当 LBS_{stu_i}, LBS_{stu_j} 满足强互模拟关系时,则 stu_i, stu_j 满足强相似性。

定义 7(学习行为序列弱相似性) 对于不满足学习行为序列强相似性的 m 个学习者组成的学习者集合 $WStu = \{\omega stu_i \mid 0 \leq i \leq m, i \in N\}$, $\forall \omega stu_i, \omega stu_j \in WStu, i \neq j$,当 $KLBS_{\omega stu_i}, KLBS_{\omega stu_j}$ 满足弱互模拟关系时,则 $\omega stu_i, \omega stu_j$ 满足弱相似性。

4.1 学习行为序列强互模拟

学习资源标号迁移系统 LRLTS 包含包括目标学习者在内的所有学习者的学习行为序列。具有 h 条路径的路径集 $Path = \{Path_p \mid 0 \leq p \leq h\}, p \in N, Path_p$ 代表学习资源标号迁移系统 LRLTS 中的第 p 条路径。接下来进行两两路径的互模拟关系分析,得到所有满足互模拟关系的候选推荐路径。因学习者可能存在学习行为序列完全一致的情况,所以我们允许两个学习者在同一条路径上进行学习行为序列强互模拟关系分析;反之,选择两条不同路径进行学习行为序列弱互模拟关系分析。

4.1.1 学习行为序列强互模拟定义

定义 8 在一个学习资源标号迁移系统 LRLTS 中,任取两条路径 $Path_g$ 和 $Path_d, 0 \leq g \leq h, 0 \leq d \leq h, g \in N, d \in N$ 分别代表第 g 条路径和第 d 条路径。对于一个二元关系 $R \subseteq Path_g \times Path_d, R_s$ 是学习行为序列强互模拟关系,对于任意的状态对 s 和 t ,

(1) 若 sR_t ,当 $s \xrightarrow{\langle r,a \rangle} s'$ 时,那么存在某个 t' 使得 $t \xrightarrow{\langle r,a \rangle} t'$ 和 $s'R_t'$ 成立。

(2) 若 sR_t ,当 $t \xrightarrow{\langle r,a \rangle} t'$ 时,那么存在某个 s' 使得 $s \xrightarrow{\langle r,a \rangle} s'$ 和 $s'R_t'$ 成立。

若两个状态满足以上条件,且 $(s,t) \in R_s$,则可判定 s 和 t 是强互模拟等价的,记为 $s \sim t$ 。

学习行为序列强互模拟通过比较可观察学习资源动作来判断是否等价。两个模型的等价是由模型中状态之间的等价关系决定的,而状态的等价是通过每次执行由学习资源动作触发的迁移来判定的。

4.1.2 学习行为序列强互模拟判定算法

文献[26]提出 on the fly 算法验证互模拟关系,但仅仅适用于纯动作触发的情况。因此,本文将学习资源的概念引入到学习行为序列互模拟关系判定过程中。首先定义 $Path_g \times Path_d$ 是基于同步动作得到的 $Path_g$ 和 $Path_d$ 的积。状态对 $(s,t) \in Path_g \times Path_d, s \in Path_g, t \in Path_d$ 。当状态对 (s,t) 的迁移可以被同一个学习资源动作触发,代表状态是互模拟的。若两个状态非互模拟,标记为 $fail$,令 $Q \subseteq (Path_g \times Path_d) \cup \{fail\}, \langle r,a \rangle, \langle r,b \rangle, \langle r_1,a \rangle, \langle r_2,a \rangle, \langle r_2,b \rangle \in RAct$,接下来定义相应的转换规则。

$$(1) \frac{(s,t) \in Q, s \xrightarrow{\langle r,a \rangle} s', t \xrightarrow{\langle r,a \rangle} t'}{(s',t') \in Q, \{(s,t) \xrightarrow{\langle r,a \rangle} (s',t')\}}$$

学习资源相同,但行为不同;

$$(2) \frac{(s,t) \in Q, s \xrightarrow{\langle r,a \rangle} s', t \xrightarrow{\langle r,b \rangle} t'}{(s',t') \in Q, \{(s,t) \xrightarrow{\emptyset} fail\}}$$

行为相同,且学习资源相同;

$$(3) \frac{(s,t) \in Q, s \xrightarrow{\langle r_1,a \rangle} s', t \xrightarrow{\langle r_2,a \rangle} t'}{(s',t') \in Q, \{(s,t) \xrightarrow{\emptyset} fail\}}$$

行为不同,且学习资源不同;

$$(4) \frac{(s,t) \in Q, s \xrightarrow{\langle r_1,a \rangle} s', t \xrightarrow{\langle r_2,b \rangle} t'}{(s',t') \in Q, \{(s,t) \xrightarrow{\emptyset} fail\}}$$

行为、学习资源都不相同。

以上 4 种情况,只有第一种满足强互模拟关系。若其他 3 种情况中的任意一种发生,算法便认为这两个状态是非强

互模拟的,并将非强互模拟状态对存入 $nobisimu$ 中,反之,存入 $bisimulation$ 中。集合 $visited$ 代表所有已经访问过的状态, $st = ((s, t), list)$ 存储当前分析的执行序列, (s, t) 代表状态对, $list$ 代表此状态对的直接后继,并通过函数 $succeed$ 获得。

$$succeed(s, t) = \{ \langle (r, a), (s', t') \rangle, s \xrightarrow{\langle r, a \rangle} s' \wedge t \xrightarrow{\langle r, a \rangle} t' \}$$

接下来给出具体的学习行为序列强互模拟判定算法。

算法 1 学习行为序列强互模拟判定算法

输入:状态对 (s_0, t_0)

输出:true or false, $nobisimu$, $bisimulation$

过程:

Repeat

$st := ((s_0, t_0), succeed(s_0, t_0))$;

$visited := \{(s_0, t_0)\}$;

$nobisimu = \emptyset$;

 while $st \neq \emptyset$

$((s_0, t_0), list) := top(st)$;

 if $list = \emptyset$

$pop(st)$;

 then backtrack;

 else

$remove(s_0', t_0')$ in $list$;

 if $(s_0, t_0) \xrightarrow{\langle r, a \rangle} (s_0', t_0')$

 if $(s_0', t_0') \notin visited$

$insert(s_0', t_0')$ in $visited$;

$push((s_0', t_0'), succeed(s_0', t_0'))$ in st ;

 end if

 elseif $s_0 \xrightarrow{\langle r, a \rangle} s_0', t_0 \xrightarrow{\langle r, a \rangle} t_0'$

$(s_0, t_0) \xrightarrow{0} fail$;

$nobisimu = nobisimu \cup \{(s_0, t_0)\}$;

 return false;

 elseif $s_0 \xrightarrow{\langle r, a \rangle} s_0', t_0 \xrightarrow{\langle r, b \rangle} t_0'$

$(s_0, t_0) \xrightarrow{0} fail$;

$nobisimu = nobisimu \cup \{(s_0, t_0)\}$;

 return false;

 elseif $s_0 \xrightarrow{\langle r, a \rangle} s_0', t_0 \xrightarrow{\langle r, b \rangle} t_0'$

$(s_0, t_0) \xrightarrow{0} fail$;

$nobisimu = nobisimu \cup \{(s_0, t_0)\}$;

 return false;

 end if

 end if

 end while

$bisimulation = bisimulation \cup \{(s_0, t_0)\}$

 return true;

fail isn't reachable, $Path_j \sim Path_k$.

接下来本文给出了算法可终止性、正确性以及时间复杂度的证明。

定理 2 在学习行为序列强互模拟判定过程中,总是存在一个状态对 (s, t) 使得学习行为序列强互模拟判定算法终止。

证明:算法开始时集合 $nobisimu$ 和 $bisimulation$ 都为空集。状态对的后继状态将由函数 $succeed$ 获得,并且通过 if 函数根据学习资源动作的不同检验互模拟关系,将检验出的

非互模拟的状态对存入 $nobisimu$ 集合中,互模拟的状态对存入 $bisimulation$ 集合中。

算法考虑的是一个有限状态的系统,因此 $visited$ 集合中的状态对是有限的, $succeed$ 函数以及 if 判断函数也不会被无限次调用。综上所述,在执行完所有被调用的 $succeed$ 函数后,整个算法终止运行。

定理 3 在学习行为序列强互模拟判定过程中,总是存在一个状态对 (s, t) ,当且仅当 s 与 t 是互模拟的,学习行为序列强互模拟判定算法才会返回 true。

证明:(反证法)假设状态对 (s, t) 是非互模拟的状态,算法返回 true。

当状态对 (s, t) 是非互模拟的状态,即 (s, t) 触发状态迁移的学习资源动作不同,对应算法 1 中 $ifelse$ 语句对应行。如果存在动作或者学习资源不同,则状态对 (s, t) 将变为 fail 状态,进而存入 $nobisimu$ 集合中,返回 false。

综上所述,假设不成立,所以学习行为序列强互模拟判定算法的正确性得以保证。

在学习行为序列强互模拟判定算法执行过程中,假设 n_1 代表 $Path_g$ 的状态数, n_2 代表 $Path_d$ 的状态数, n 代表 $Path_g \times Path_d$ 的状态数,其中 $n \leq n_1 \times n_2$ 。由此可知,判定互模拟关系的状态对最多为 n ,判定过程最多需要执行 n 次。因此,学习行为序列强互模拟判定算法的时间复杂度为 $O(n)$ 。

4.2 学习行为序列弱互模拟

4.1 节从学习行为序列强互模拟角度出发,对可观测的学习资源动作的等价展开研究。但对学习行为序列进行相似性分析时,往往需要隐藏一些对学习行为分析效果影响不大的非关键学习行为。本文将这些非关键学习行为使用内部动作来刻画,用符号 τ 来表示。具体定义如下:

定义 9 在一个学习资源标号迁移系统 LRLTS 中,任取两条路径 $Path_g, Path_d, 0 \leq g \leq h, 0 \leq d \leq h, g \in N, d \in N$ 分别代表第 g 条路径和第 d 条路径。对于一个二元关系 $R \subseteq Path_g \times Path_d, R_w$ 是学习行为序列弱互模拟关系。对于任意的状态对 s 和 $t, \Rightarrow (\xrightarrow{\langle r, a \rangle})^* \xrightarrow{\tau} (\xrightarrow{\langle r, a \rangle})^*$ 代表弱迁移。

(1)若 $sR_w t$,当 $s \xrightarrow{\langle r, a \rangle} s'$ 时,那么存在某个 t' 使得 $t \Rightarrow t'$ 和 $s'R_w t'$ 成立。

(2)若 $sR_w t$,当 $t \xrightarrow{\langle r, a \rangle} t'$ 时,那么存在某个 s' 使得 $s \Rightarrow s'$ 和 $s'R_w t'$ 成立。

若两个状态满足以上条件,即,如果 $(s, t) \in R_w$,则可判定 s 和 t 是弱互模拟等价的,记为 $s \approx t$ 。

由上述定义可知,若要证明两个 LR-CCS 是互模拟的,首先要构造一个二元关系 R ,使得 $(s, t) \in R$,然后再利用定义证明即可。

为了将学习行为序列强互模拟算法应用在判定学习行为序列弱互模拟关系上,本文将学习资源标号迁移系统进行饱和和处理,因篇幅限制,过程不在此赘述,具体可参考文献[27]。

5 实例

5.1 基于相似学习者的推荐系统设计

基于相似学习者的推荐是经典的推荐方法,因此本文以

基于相似学习者的推荐系统(CRS)为例展开论述,架构图如图2所示。

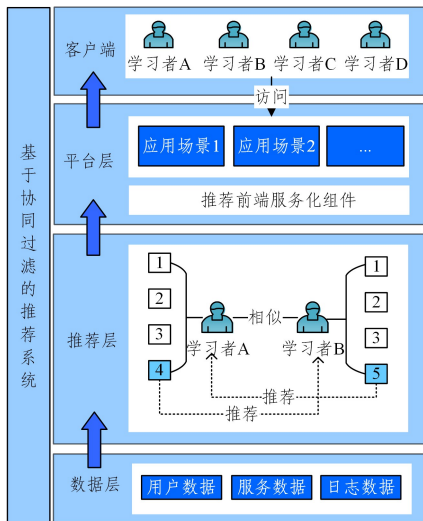


图2 基于相似学习者的推荐系统架构图

Fig.2 Architecture diagram of recommendation system based on similar learners

该系统自下而上分为4层,各层的主要作用如下:

数据层(DATA LAYER):数据层包括用户数据及用户行为数据等,如用户的浏览、点击等。

推荐层(RECOMMENDATION LAYER):主要包括一些推荐算法(基于相似学习者的算法),其中相似学习者判定是推荐的基础,它包括相似学习者判定模块(SLD)以及推荐模块(RM)两个模块。相似学习者判定模块的正确性决定了推荐模块的结果可信性。

平台层(PLATFORM LAYER):平台层包括应用场景模块以及推荐前端服务化组件。

用户层(CLIENT):用户层包含一系列用户,系统通过平台层向用户展示信息,用户根据需求决策要访问的应用场景,并做出相应的访问动作。

基于相似学习者的推荐系统由这几个模块协同完成推荐工作。其中,推荐层中的相似学习者判定是其子过程,获取学习者的学习行为序列是判定相似性的基础。

5.2 学习行为序列获取

5.2.1 研究对象和数据来源

本研究以江苏师范大学2021-2022第二学期选取《软件工程》必修课的在线行为数据为研究对象,从45名学生中,以获得有效期末成绩为标准(成绩85分以上),筛选出27名学习者。所有数据来源于希冀平台数据库记录的课程及学习行为数据,选择了5种行为(见表1)和5种学习资源(见表2),并根据课程大纲的要求,参考“教育目标分类理论认知分类”,确定相关知识点的学习目标。

表1 学习行为编码

Table 1 Learning behavior coding

编码	学习行为	说明
wt	观看视频	通过希冀平台查看章节视频
vw	浏览文档	通过希冀平台查看文档资源
vp	浏览PPT	通过希冀平台查看PPT资源
pw	参与作业	在希冀平台上参与作业
cm	评论	学习者对各类资源进行点评

表2 学习资源编码

Table 2 Learning resource coding

编码	学习资源名称	对应的知识点	学习目标
r_0	软件工程概述资料	软件工程定义、软件危机等	领会
r_1	软件需求方法资料	需求初始化、结构化的需求分析、面向对象的需求分析等	应用
r_2	UML需求建模案例资料	UML动态模型、UML静态模型、UML功能模型等	应用
r_3	面向对象的设计资料	面向对象的分析建模、面向对象的设计建模	应用
r_4	测试试题资料	测试概述、黑盒测试等	应用

5.2.2 学习行为序列选取

为了更好地说明本文学习行为序列的选取,首先介绍字母表的相关概念。字母表中的字符必须具备两个特点:

(1)整体性。例如{aa, bb, cc}中的aa, bb, cc均是单个字符。

(2)可辨认性。字母表中的字符是独一无二的,可以明确区分。

由5.2.1节可知,学习行为集 $Act = \{wt, vw, vp, pw, cm\}$,学习资源集 $Res = \{r_0, r_1, r_2, r_3, r_4\}$,二者的笛卡尔乘积用 $Res \times Act$ 表示,是一个有序对的集合,那么

$$Res \times Act = \{ \langle r_0, wt \rangle, \langle r_0, vw \rangle, \langle r_0, vp \rangle, \langle r_0, pw \rangle, \langle r_0, cm \rangle, \langle r_1, wt \rangle, \langle r_1, vw \rangle, \langle r_1, vp \rangle, \langle r_1, pw \rangle, \langle r_1, cm \rangle, \langle r_2, wt \rangle, \langle r_2, vw \rangle, \langle r_2, vp \rangle, \langle r_2, pw \rangle, \langle r_2, cm \rangle, \langle r_3, wt \rangle, \langle r_3, vw \rangle, \langle r_3, vp \rangle, \langle r_3, pw \rangle, \langle r_3, cm \rangle, \langle r_4, wt \rangle, \langle r_4, vw \rangle, \langle r_4, vp \rangle, \langle r_4, pw \rangle, \langle r_4, cm \rangle \}$$

由于其中的每一个元素对都是独一无二的。因此,

$$\Sigma = Act \times Res = \{ \langle r_0, wt \rangle, \langle r_0, vw \rangle, \langle r_0, vp \rangle, \langle r_0, pw \rangle, \langle r_0, cm \rangle, \langle r_1, wt \rangle, \langle r_1, vw \rangle, \langle r_1, vp \rangle, \langle r_1, pw \rangle, \langle r_1, cm \rangle, \langle r_2, wt \rangle, \langle r_2, vw \rangle, \langle r_2, vp \rangle, \langle r_2, pw \rangle, \langle r_2, cm \rangle, \langle r_3, wt \rangle, \langle r_3, vw \rangle, \langle r_3, vp \rangle, \langle r_3, pw \rangle, \langle r_3, cm \rangle, \langle r_4, wt \rangle, \langle r_4, vw \rangle, \langle r_4, vp \rangle, \langle r_4, pw \rangle, \langle r_4, cm \rangle \}$$

是一个字母表,是由学习资源动作组成的非空有穷集合。

$$\Sigma^+ = \{ \langle r_0, wt \rangle, \langle r_0, vw \rangle, \langle r_0, vp \rangle, \langle r_0, pw \rangle, \langle r_0, cm \rangle, \langle r_1, wt \rangle, \langle r_1, vw \rangle, \langle r_1, vp \rangle, \langle r_1, pw \rangle, \langle r_1, cm \rangle, \langle r_2, wt \rangle, \langle r_2, vw \rangle, \langle r_2, vp \rangle, \langle r_2, pw \rangle, \langle r_2, cm \rangle, \langle r_3, wt \rangle, \langle r_3, vw \rangle, \langle r_3, vp \rangle, \langle r_3, pw \rangle, \langle r_3, cm \rangle, \langle r_4, wt \rangle, \langle r_4, vw \rangle, \langle r_4, vp \rangle, \langle r_4, pw \rangle, \langle r_4, cm \rangle, \langle r_0, wt \rangle \langle r_0, wt \rangle, \langle r_0, wt \rangle \langle r_0, vw \rangle, \dots, \langle r_0, wt \rangle \langle r_0, wt \rangle \langle r_0, wt \rangle, \dots \}$$

Σ^+ 是所有可能学习行为序列的集合,所有的学习行为序列都可由字母表中的这几个元素组成。为了简化分析过程,在分析时,只考虑学习资源第一次出现的情况,给出4个学习者的学习行为序列。其中, S_{ij} 代表第*i*个学习者的第*j*个状态。

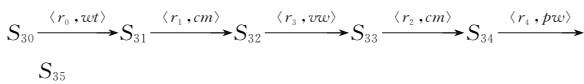
学习者STU1的学习行为序列:

$$S_{10} \xrightarrow{\langle r_0, wt \rangle} S_{11} \xrightarrow{\langle r_1, vw \rangle} S_{12} \xrightarrow{\langle r_2, vp \rangle} S_{13} \xrightarrow{\langle r_3, cm \rangle} S_{14} \xrightarrow{\langle r_4, pw \rangle} S_{15}$$

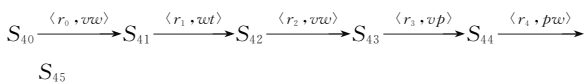
学习者STU2的学习行为序列:

$$S_{20} \xrightarrow{\langle r_0, wt \rangle} S_{21} \xrightarrow{\langle r_1, vw \rangle} S_{22} \xrightarrow{\langle r_2, vp \rangle} S_{23} \xrightarrow{\langle r_3, cm \rangle} S_{24} \xrightarrow{\langle r_4, pw \rangle} S_{25}$$

学习者 STU_3 学习行为序列:



学习者 STU_4 学习行为序列:



接下来根据 27 名学生的在线行为数据获得关键行为。按照 GSEQ 要求的格式输入所有行为编码(见图 3),生成行为序列频次表(见图 4)。由图 4 可知,行为编码为 wt 以及 pw 的学习行为频次远高于其他行为,因此将这两种行为作为关键行为。

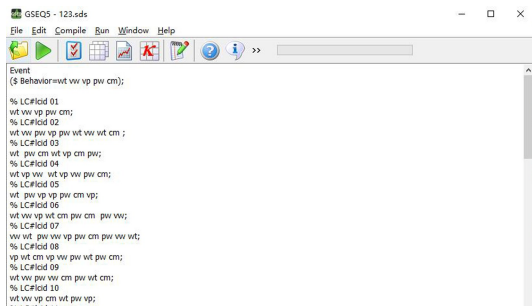


图 3 行为序列编码

Fig. 3 Behavioral sequence coding

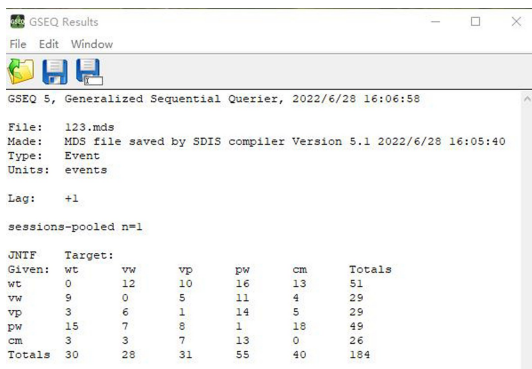


图 4 行为频次

Fig. 4 Behavior frequency

5.3 学习行为序列互模拟关系分析

5.3.1 学习行为序列互模拟关系判定

由于判定学习者相似性是推荐系统的关键,接下来针对推荐层 RECOMMENDATION LAYER 中的相似学习者判定过程 SLD 进行分析。SLD 为当前相似学习者判定过程模型,包含强互模拟关系分析模块 SA 和弱互模拟关系分析模块 WA。其中,学习行为集 $Act = \{wt, vw, vp, pw, cm\}$,学习资源集 $Res = \{r_0, r_1, r_2, r_3, r_4\}$,学习资源动作集 $RAct = \{\langle r_0, wt \rangle, \langle r_0, vw \rangle, \langle r_1, vw \rangle, \langle r_1, cm \rangle, \langle r_1, wt \rangle, \langle r_2, vp \rangle, \langle r_2, cm \rangle, \langle r_2, vw \rangle, \langle r_3, vw \rangle, \langle r_3, cm \rangle, \langle r_3, vp \rangle, \langle r_4, pw \rangle\}$,关键学习资源动作集 $RKAct = \{\langle r_0, wt \rangle, \langle r_1, wt \rangle, \langle r_4, pw \rangle\}$,则相似性判定过程模型的进程表达式为:

$$SLD = SA + WA$$

$$SA = STU1 | STU2 | STU3 | STU4$$

$$STU1 = \langle r_0, wt \rangle. \langle r_1, vw \rangle. \langle r_2, vp \rangle. \langle r_3, cm \rangle. \langle r_4, pw \rangle$$

$$STU2 = \langle r_0, wt \rangle. \langle r_1, vw \rangle. \langle r_2, vp \rangle. \langle r_3, cm \rangle. \langle r_4, pw \rangle$$

$$STU3 = \langle r_0, wt \rangle. \langle r_1, cm \rangle. \langle r_3, vw \rangle. \langle r_2, cm \rangle. \langle r_4, pw \rangle$$

$$STU4 = \langle r_0, vw \rangle. \langle r_1, wt \rangle. \langle r_2, vw \rangle. \langle r_3, vp \rangle. \langle r_4, pw \rangle$$

$$WA = STU1' | STU2' | STU3' | STU4'$$

$$STU1' = STU1 \setminus \{\langle r_1, vw \rangle, \langle r_2, vp \rangle, \langle r_3, cm \rangle\}$$

$$STU2' = STU2 \setminus \{\langle r_1, vw \rangle, \langle r_2, vp \rangle, \langle r_3, cm \rangle\}$$

$$STU3' = STU3 \setminus \{\langle r_1, cm \rangle, \langle r_3, vw \rangle, \langle r_2, cm \rangle\}$$

$$STU4' = STU4 \setminus \{\langle r_0, vw \rangle, \langle r_2, vw \rangle, \langle r_3, vp \rangle\}$$

$$STU1 \setminus \{\langle r_1, vw \rangle, \langle r_2, vp \rangle, \langle r_3, cm \rangle\} = \langle r_0, wt \rangle. \tau. \tau. \tau. \langle r_4, pw \rangle$$

$$STU2 \setminus \{\langle r_1, vw \rangle, \langle r_2, vp \rangle, \langle r_3, cm \rangle\} = \langle r_0, wt \rangle. \tau. \tau. \tau. \langle r_4, pw \rangle$$

$$STU3 \setminus \{\langle r_1, cm \rangle, \langle r_3, vw \rangle, \langle r_2, cm \rangle\} = \langle r_0, wt \rangle. \tau. \tau. \tau. \langle r_4, pw \rangle$$

$$STU4 \setminus \{\langle r_0, vw \rangle, \langle r_2, vw \rangle, \langle r_3, vp \rangle\} = \tau. \langle r_1, wt \rangle. \tau. \tau. \langle r_4, pw \rangle$$

其中,SA 进程是 $STU1, STU2, STU3$ 以及 $STU4$ 的进程并发,各个进程的执行由相应的资源动作有序对顺序执行触发;WA 进程是 $STU1', STU2', STU3'$ 以及 $STU4'$ 的进程并发。 $STU1', STU2', STU3'$ 以及 $STU4'$ 是在 $STU1, STU2, STU3$ 以及 $STU4$ 进程基础上隐藏非关键行为得到的。

为了简化状态空间大小,文中仅画出了 SA 的学习资源标号迁移系统图(见图 5),通过 4.1 节学习行为序列强互模拟判定算法判定互模拟关系。

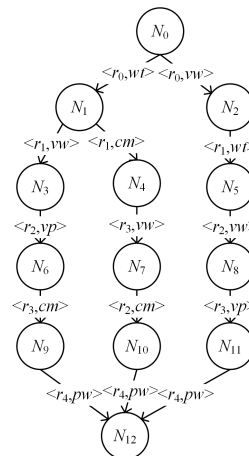
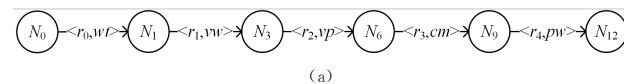


图 5 SA 的学习资源标号迁移系统图

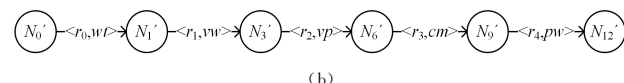
Fig. 5 Learning resources-labeled transition systems graph of SA

图 5 包含了 $STU1$ 等 4 位学习者的学习行为序列。我们可以在 LR-CCS 模型上进行路径上互模拟关系的检验,即学习行为序列互模拟关系检验。接下来给出学习者 $STU1$ 和 $STU2$ 所对应的学习资源标号迁移系统图的路径,为了区分 $STU1$ 以及 $STU2$,用 N_i, N_i' 分别表示 $STU1$ 以及 $STU2$ 的路径。

下面证明学习者 $STU1$ 与 $STU2$ 满足强互模拟关系。先证明图 6(a)被图 6(b)强互模拟。



(a)



(b)

图 6 示例路径

Fig. 6 Example path

- (1) 第一步,图 6(a)中有 $N_0 \xrightarrow{\langle r_0, wt \rangle} N_1$; 图 6(b)中存在 $N_1',$ 使得 $N_0' \xrightarrow{\langle r_0, wt \rangle} N_1';$
- (2) 第二步,图 6(a)中有 $N_1 \xrightarrow{\langle r_1, vw \rangle} N_3$; 图 6(b)中存在 $N_3',$ 使得 $N_1' \xrightarrow{\langle r_1, vw \rangle} N_3';$
- (3).....
- (4) 第五步,图 6(a)中有 $N_9 \xrightarrow{\langle r_1, pw \rangle} N_{12}$; 图 6(b)中存在 $N_{12}',$ 使得 $N_9' \xrightarrow{\langle r_1, pw \rangle} N_{12}';$
- (5) 第六步, N_{12} 和 N_{12}' 均为结束标识符。

综上所述, S_{12}, S_{12}' , 关系 $R_s = \{(N_0, N_0'), (N_1, N_1'), (N_3, N_3'), (N_6, N_6'), (N_9, N_9'), (N_{12}, N_{12}')\}$ 。再证明关系 R 的逆 $R_s^{-1} = \{(N_0', N_0), (N_1', N_1), (N_3', N_3), (N_6', N_6), (N_9', N_9), (N_{12}', N_{12})\}$ 是强互模拟, 即, 图 6(b) 被图 6(a) 强互模拟。

- (1) 第一步,图 6(b)中有 $N_0' \xrightarrow{\langle r_0, wt \rangle} N_1';$ 图 6(a)中存在 $N_1,$ 使得 $N_0 \xrightarrow{\langle r_0, wt \rangle} N_1;$
- (2) 第二步,图 6(b)中有 $N_1' \xrightarrow{\langle r_1, vw \rangle} N_3';$ 图 6(a)中存在 $N_3,$ 使得 $N_1 \xrightarrow{\langle r_1, vw \rangle} N_3;$
- (3).....
- (4) 第五步,图 6(b)中有 $N_9' \xrightarrow{\langle r_1, pw \rangle} N_{12}';$ 图 6(a)中存在 $N_{12},$ 使得 $N_9 \xrightarrow{\langle r_1, pw \rangle} N_{12}.$

综上所述, $N_0, R_s^{-1} N_0'$, 因此 $N_0 \sim N_0'$, 故学习者 $STU1$ 和 $STU2$ 满足强互模拟关系, 且满足学习行为序列强相似性。学习者 $STU3, STU4$ 等的互模拟关系判定过程和上文的判定过程类似, 由于篇幅原因不做过多的赘述, 接下来使用互模拟验证工具验证四者之间的学习行为序列相似性。

5.3.2 学习行为序列互模拟关系验证

本文采用验证工具 MWB 对学习者的学习行为序列进行互模拟关系验证, 具体过程如下:

- (1) 将进程表达式表示为 MWB 编码。
- (2) 运用互模拟验证命令验证不同学习者的进程表达式是否满足互模拟关系。

我们将 SA 过程的 $STU1$ 等 4 个学习者的进程表达式转换为 MWB 编码, 并输入到 MWB 工具中(见图 7), 采用命令 eq 进行强互模拟关系检验, 由图 8 可知, 只有 $STU1$ 与 $STU2$ 是满足强互模拟关系的, 它们属于强相似学习者。

```
C:\Windows\system32\cmd.exe
D:\迅雷下载\mwb-进程代数工具安装文件\mwb-进程代数安装文件>sm1 @SMLload=mwb.x86-win32
The Mobility Workbench
(MWB 99, version 4.136, built Thu Jan 08 13:13:41 2004)
MWB>agent STU1 (r0wt, r1vw, r2vp, r3cm, r4pw)=r0wt. r1vw. r2vp. r3cm. r4pw. STU1 <r0wt, r1vw, r2vp, r3cm, r4pw>
MWB>agent STU2 (r0wt, r1vw, r2vp, r3cm, r4pw)=r0wt. r1vw. r2vp. r3cm. r4pw. STU2 <r0wt, r1vw, r2vp, r3cm, r4pw>
MWB>agent STU3 (r0wt, r1cm, r3vw, r2cm, r4pw)=r0wt. r1cm. r3vw. r2cm. r4pw. STU3 <r0wt, r1cm, r3vw, r2cm, r4pw>
MWB>agent STU4 (r0vw, r1wt, r2vw, r3vp, r4pw)=r0vw. r1wt. r2vw. r3vp. r4pw. STU4 <r0vw, r1wt, r2vw, r3vp, r4pw>
```

图 7 学习行为序列的 MWB 编码

Fig. 7 MWB coding of learning behavior sequences

```
MWB>eq STU1 (r0wt, r1vw, r2vp, r3cm, r4pw) STU2 (r0wt, r1vw, r2vp, r3cm, r4pw)
The two agents are equal.
Bisimulation relation size = 5.
MWB>eq STU1 (r0wt, r1vw, r2vp, r3cm, r4pw) STU3 (r0wt, r1cm, r3vw, r2cm, r4pw)
The two agents are NOT equal.
MWB>eq STU1 (r0wt, r1vw, r2vp, r3cm, r4pw) STU4 (r0vw, r1wt, r2vw, r3vp, r4pw)
The two agents are NOT equal.
MWB>eq STU3 (r0wt, r1cm, r3vw, r2cm, r4pw) STU4 (r0vw, r1wt, r2vw, r3vp, r4pw)
The two agents are NOT equal.
```

图 8 学习行为序列强相似性判定结果图

Fig. 8 Strong similarity judgment result graph of learning behavior sequence

接下来进行弱互模拟关系的检验, 首先将 WA 过程中的 $STU1'$ 等 4 个学习者的进程表达式转换为 MWB 编码, 并输入到 MWB 工具中(见图 9), 采用命令 weq 进行弱互模拟关系检验, 由图 10 可知, 只有 $STU1$ 与 $STU3$ 是满足弱互模拟关系的, 属于弱相似学习者, $STU1$ 与 $STU2$ 是强互模拟关系, 因此 $STU2$ 与 $STU3$ 也属于弱相似学习者。

```
MWB>agent STU1' (r0wt, r4pw)=r0wt. t. t. r4pw. STU1' <r0wt, r4pw>
MWB>agent STU2' (r0wt, r4pw)=r0wt. t. t. r4pw. STU2' <r0wt, r4pw>
MWB>agent STU3' (r0wt, r4pw)=r0wt. t. t. r4pw. STU3' <r0wt, r4pw>
MWB>agent STU4' (r1wt, r4pw)=t. r1wt. t. t. r4pw. STU4' <r1wt, r4pw>
```

图 9 关键学习行为序列 MWB 编码

Fig. 9 MWB coding of key learning behavior sequences

```
MWB>weq STU1' (r0wt, r4pw) STU3' (r0wt, r4pw)
The two agents are equal.
Bisimulation relation size = 9.
MWB>weq STU1' (r0wt, r4pw) STU4' (r1wt, r4pw)
The two agents are NOT equal.
MWB>weq STU3' (r0wt, r4pw) STU4' (r1wt, r4pw)
The two agents are NOT equal.
```

图 10 关键学习行为序列弱相似性判定结果

Fig. 10 Weak similarity judgment result graph of key learning behavior sequences

综上所述 $STU1$ 与 $STU2$ 的学习行为序列满足强相似性, $STU1, STU2$ 与 $STU3$ 的学习行为序列满足弱相似性, 而 $STU4$ 与其他学习者不相似。

经过互模拟关系验证, 得到路径 $\{N_0, N_2, N_5, N_8, N_{11}, N_{12}\}$ 不满足互模拟关系, 说明该路径不会被推荐给学习者, 因此进行容错处理, 删除此路径, 得到修正后的学习资源标号迁移系统图(见图 11), 此时所有路径都是候选推荐路径。

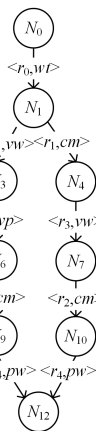


图 11 修正后的学习资源标号迁移系统图

Fig. 11 Modified graph of learning resources-labeled transition systems

6 相关工作

根据相似性进行个性化推荐研究一直备受关注。文献[28]采用最近邻居的卷积深度学习模型, 经过计算相似度, 得到最近邻居集; 文献[29]运用深度神经网络来挖掘用户和物品之间的复杂关系, 计算目标用户和历史用户的相似性, 捕捉用户交互项目的复杂过程; 文献[30]提出了一种新的基于深度神经网络的协同过滤推荐引擎, 采用神经架构进行偏好学习和用户表示, 能够深度挖掘每个用户在社交网络中的互动; 文献[31]提出了一种基于图的 ICF (Item-based Collaborative Filtering) 方法 (graph-ICF), 利用图结构的信息聚合和传播

特性,挖掘出现有 ICF 模型所忽略的项目关系。以上方法都无法保证推荐过程的可解释性以及推荐结果的可信性。而导致这一问题的根本原因是以上方法缺乏数理逻辑支撑。本文在解决学习者相似性问题上另辟蹊径,通过逻辑性较强的形式化方法对学习行为相似性进行判定与验证。

形式化方法在资源建模方面的相关研究众多。文献[32]提出了一种刻画资源使用和消耗的演算系统 RUCCS(Resources Usage and Consumption Calculus System)并对其行为理论、逻辑特征等进行了研究;文献[33-34]在进程代数中引入资源概念,提出了一个通信共享资源的代数(Algebra of Communicating Shared Resources, ACSR),为计算机系统资源建模;以上文献中的资源主要指计算机中供进程共享的资源;另外,文献[35]将资源定义为可再生资源、不可再生资源和双重约束资源,并制定了块结构资源受限调度过程的抽象模型,用于活动和资源建模;文献[36]针对时空约束下的特殊资源,如停车场中的停车位、线路中的一段铁轨等,提出了时空资源通信顺序进程(Duration-Space Resource, TCSP),用于时空约束下的资源安全性分析。文献[37]定义了虚拟定时环境演算,模型中的资源配置基于作为本地资源的虚拟时间片。与以上方法不同的是,本文的资源是指学习者学习过程中的学习对象-学习资源,现有模型并不适合本文的研究,而学习资源是学习行为序列的重要组成部分,且若对学习行为序列的描述不够准确,可能导致分析结果错误。因此,本文提出了一种基于学习行为序列的建模语言 LR-CCS,为学习行为序列互模拟关系分析提供可靠的模型支持。

针对互模拟关系的研究,近些年来成果丰富。文献[38]提出了一种名为 on the fly 的局部算法,只需要验证相关 s 和 t 的转移行为;文献[39]提出了准局部算法,对于简单的 LTS 和确定型 LTS,准局部算法具有更好的时间复杂度;文献[40]改进准局部算法,使其可以验证一般的 LTS。另外,互模拟在多个领域都有广泛应用。文献[41]提出了一种有效的算法来解决确定有限结构上的互模拟关系的问题,该算法对显式和隐式(符号)表示都有效;文献[42]提出了一种可以计算概率标记转换系统的互模拟算法,具有非确定性选择和离散概率选择;文献[43]介绍了基于并行关系的符号互模拟算法及其推导;文献[44]通过进程代数互模拟定义了神经网络之间的等价概念,引入一个与给定的神经网络在语义上等价的简化版网络。但以上互模拟算法都是围绕进程之间的行为展开,都没有考虑学习资源对互模拟关系的影响。因此,本文在传统互模拟基础上扩展学习资源信息,提出学习行为序列互模拟关系判定算法,并使用 MWB 工具保证其可信性。

基于上述调研和分析,本文提出了一种基于相似学习者判定的个性化学习路径推荐及验证方法,以提供可靠的形式化模型为基础,采用逻辑推理的方式提高过程可解释性,并通过相应的验证工具保证推荐结果的可信性。

结束语 本文针对现有个性化推荐中的可解释性差以及结果可信性差的问题,提出了一种基于相似学习者判定的个性化学习路径推荐及验证方法。首先定义学习行为序列的概念;其次在进程代数通信演算系统 CCS 中引入学习资源属性,提出学习资源-通信演算系统 LR-CCS,使之能够形式化建模学习行为序列;再次,通过进程代数中互模拟理论进行学习行为序列相似性判定,并使用 MWB 来验证

学习者学习行为序列相似性;最后通过一个基于相似学习者的推荐系统实例验证了该方法的有效性。

但本文方法还有一定的局限性,需进一步研究:

(1)本文方法在状态空间过大的情况下存在一定缺陷,在后期工作中需要增强处理状态空间爆炸的能力。

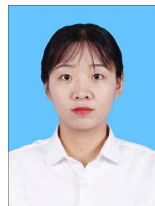
(2)本所提出的学习行为序列强互模拟以及学习行为序列弱互模拟都不能表示进程相似程度。接下来我们将刻画出相似学习者之间的量化关系,得到相似度。

(3)推荐时应该考虑到知识点之间存在顺序关系,一个知识点是否可以学习往往取决于另一些知识点是否学习过。因此在下一步工作中将考虑对候选推荐路径进行知识点顺序关系的验证,从而进一步提升推荐结果的正确性以及可信性。

参考文献

- [1] YUN Y, DAI H, ZHANG P Y, et al. State-of-the-Art Survey of Personalized Learning Path Recommendation [J]. Journal of Software, 2022, 33(12): 4590-4615.
- [2] JAIN A, NAGAR S, SINGH P K, et al. EMUCF: Enhanced multistage user-based collaborative filtering through non-linear similarity for recommendation systems [J]. Expert Systems with Applications, 2020, 161: 113724.
- [3] SARWAR B, KARYPIS G, KONSTAN J, et al. Item-based collaborative filtering recommendation algorithms [C] // Proceedings of the 10th international conference on World Wide Web. 2001: 285-295.
- [4] TEWARI A S. Generating items recommendations by fusing content and user-item based collaborative filtering [C] // Procedia Computer Science, 2020, 167: 1934-1940.
- [5] YU M, QUAN T, PENG Q, et al. A model-based collaborative filtering algorithm based on stacked AutoEncoder [J]. Neural Computing and Applications, 2022, 34(4): 2503-2511.
- [6] WANG X, KADIOĞLU S. Modeling uncertainty to improve personalized recommendations via Bayesian deep learning [J]. International Journal of Data Science and Analytics, 2021, 16(2/3): 191-201.
- [7] TAN W Z. The Algorithmic Risk of Artificial Intelligence in Education [J]. Open Education Research, 2019, 25(6): 20-30.
- [8] LI L. Design and Implementation of Recommendation Algorithm Based on Post Hoc Interpretability [D]. Taiyuan: Shanxi University, 2021.
- [9] LIU H L, GUO Y, MA J. Research Progress of Similarity Algorithm in Collaborative Filtering [J]. Computer Engineering and Applications, 2022, 58(13): 27-35.
- [10] MA S C, SHI J Q, HUANG Y H, et al. Minimal-unsatisfiable-core-driven Local Explainability Analysis for Random Forest [J]. Journal of Software, 2022, 33(7): 2447-2463.
- [11] GOODFELLOW I J, SHLENS J, SZEGEDY C. Explaining and Harnessing Adversarial Examples [C] // ICLR. 2015.
- [12] WU Y, ZHANG M. Tightening Robustness Verification of Convolutional Neural Networks with Fine-Grained Linear Approximation [C] // The Thirty-Fifth AAAI Conference on Artificial Intelligence (AAAI-21). 2021.
- [13] MILNER R. A Calculus of Communicating Systems [C] // Springer Berlin Heidelberg, 1980.
- [14] HOARE C A R. Communicating sequential processes [J]. Com-

- munications of the ACM, 1978, 21(8):666-677.
- [15] ISO. ISO 8807. LOTOS Standard[OL]. http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=16258.
- [16] MILNER R. Communication and concurrency[M]. Englewood Cliffs: Prentice hall, 1989.
- [17] FIORE M, CATTANI G L, WINSKEL G. Weak bisimulation and open maps. Proceedings[C]// 14th Symposium on Logic in Computer Science(Cat. No. PR00158). IEEE, 1999: 67-76.
- [18] MA Z Q, WANG Y C, YUE Y Z, et al. Research on Online Learning Behavior Pattern from the Perspective of Learning Analysis[J]. Modern Distance Education, 2019(6): 35-44.
- [19] BAKEMAN R, GOTTMAN J M. Observing interaction: An introduction to sequential analysis [M]. Cambridge university press, 1997.
- [20] ZHANG H, CUI Y L, YU L Y, et al. Study of Classroom Event Logic Graph of Intelligent Teaching based on Method of Data Mining[J]. Journal of Distance Education, 2020, 38(2): 80-88.
- [21] BAKEMAN R, QUERA V. Analyzing interaction: Sequential analysis with SDIS and GSEQ [M]. Cambridge University Press, 1995.
- [22] LI Y, LI X. A Verification Tool MWB for CCS[J]. Communications Technology, 2005(S1): 178-180.
- [23] ZHU Y, XIAO F X, ZHOU H, et al. Method for Modeling and Analyzing Software Energy Consumption of Embedded Real-Time System[J]. Journal of Computer Research and Development, 2014, 51(4): 848-855.
- [24] CHEN X Y, ZHU Y, ZHAO Y, et al. Hybrid AADL Modeling and Model Transformation for CPS Time and Space Properties Verification[J]. Journal of Software, 2021, 32(6): 1779-1798.
- [25] FEI Y, ZHONG L, JHA N K. An energy-aware framework for dynamic software management in mobile computing systems [J]. ACM Transactions on Embedded Computing Systems (TECS), 2008, 7(3): 1-31.
- [26] FERNANDEZ J C, MOUNIER L. Verifying Bisimulations “On the Fly”[C]// FORTE. 1990: 95-110.
- [27] LEI G Q. Comparison and Application of Bisimulation Verification Algorithm[M]. East China Normal University, 2018.
- [28] ZHENG G, YU H, XU W. Collaborative filtering recommendation algorithm with item label features[J]. International Core Journal of Engineering, 2020, 6(1): 160-170.
- [29] CHEN J, WANG X, ZHAO S, et al. Deep attention user-based collaborative filtering for recommendation[J]. Neurocomputing, 2020, 383: 57-68.
- [30] ZHOU W, DU Y, DUAN M, et al. NtCF: Neural Trust-Aware Collaborative Filtering Toward Hierarchical Recommendation Services[J]. Arabian Journal for Science and Engineering, 2022, 47(2): 1239-1252.
- [31] LIU M, LI J, LIU K, et al. Graph-ICF: Item-based collaborative filtering based on graph neural network[J]. Knowledge-Based Systems, 2022, 251: 109208.
- [32] WANG L S. Research on Theory and Application of Access Control Based on Resources [D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2010.
- [33] LEE I, PHILIPPOU A, SOKOLSKY O. Resources in process algebra[J]. The Journal of Logic and Algebraic Programming, 2007, 72: 98-122.
- [34] LEE I, PHILIPPOU A, SOKOLSKY O. A general resource framework for real-time systems[C]// Proceedings of RISSEF’02, LNCS 2941. 2002: 234-248.
- [35] BĂDICĂ A, BĂDICĂ C, LOGOFĂTU D, et al. Modeling block structured project scheduling with resource constraints[C]// International Conference on Large-Scale Scientific Computing. Cham: Springer, 2019: 484-492.
- [36] CHEN X Y, ZHU Y, ZHAO Y, et al. Modeling and safety verification method for CPS time and topology constrained resources [J]. Journal of Software, 2022, 33(8): 2815-2838.
- [37] JOHNSEN E B, STEFFEN M, STUMPF J B. Virtually timed ambients: A calculus of nested virtualization[J]. Journal of Logical and Algebraic Methods in Programming, 2018, 94: 109-127.
- [38] FERNANDEZ J C, MOUNIER L. Verifying bisimulations on the fly[C]// Proceedings of the 3rd International Conference on Formal Description Techniques for Distributed Systems and Communication Protocols. North-Holland Publishing Co., 1990: 95-110.
- [39] DU W J, DENG Y X. A quasi-local algorithm for checking bisimilarity[C]// Proceedings of the IEEE International Conference on Computer Science and Automation Engineering, 2011: 1-5.
- [40] ZHENG X L, DENG Y X, FU C, et al. Extension and Implementation of the Quasi-Local Algorithm for Checking Bisimilarity [J]. Journal of Software, 2018, 29(6): 1517-1526.
- [41] DOVIER A, PIAZZA C, POLICRITI A. An efficient algorithm for computing bisimulation equivalence[J]. Theoretical Computer Science, 2004, 311(1/2/3): 221-256.
- [42] GROOTE J F, RIVERA VERDUZCO J, DE VINK E P. An efficient algorithm to determine probabilistic bisimulation[J]. Algorithms, 2018, 11(9): 131.
- [43] HUYBERS R, LAARMAN A. A Parallel Relation-Based Algorithm for Symbolic Bisimulation Minimization [C] // International Conference on Verification, Model Checking, and Abstract Interpretation. Cham: Springer, 2019: 535-554.
- [44] PRABHAKAR P. Bisimulations for Neural Network Reduction [C] // International Conference on Verification, Model Checking, and Abstract Interpretation. Cham: Springer, 2022: 285-300.



FENG Shu, born in 1996, postgraduate, is a member of China Computer Federation. Her main research interests include adaptive learning, formal methods and software engineering.



ZHU Yi, born in 1976, Ph.D, professor, is a member of China Computer Federation. His main research interests include software engineering, formal methods, software reliability, adaptive learning and cyber-physical fusion systems.