



计算机科学

COMPUTER SCIENCE

基于二分图的个性化学习任务分配

谭珍珠, 姜文君, 任演纳, 张吉, 任德盛, 李晓鸿

引用本文

谭珍珠, 姜文君, 任演纳, 张吉, 任德盛, 李晓鸿. [基于二分图的个性化学习任务分配](#)[J]. 计算机科学, 2022, 49(4): 269-281.

TAN Zhen-qiong, JIANG Wen-Jun, YUM Yen-na-cherry, ZHANG Ji, YUM Peter-tak-shing, LI Xiao-hong.

[Personalized Learning Task Assignment Based on Bipartite Graph](#)[J]. Computer Science, 2022, 49(4): 269-281.

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)

[基于图像分块与特征融合的户外图像天气识别](#)

Outdoor Image Weather Recognition Based on Image Blocks and Feature Fusion

计算机科学, 2022, 49(3): 197-203. <https://doi.org/10.11896/jsjcx.201200263>

[跨领域文本的可迁移情绪分析方法](#)

Transferable Emotion Analysis Method for Cross-domain Text

计算机科学, 2022, 49(3): 218-224. <https://doi.org/10.11896/jsjcx.210400034>

[结合特征融合和注意力机制的微表情识别方法](#)

Micro-expression Recognition Method Combining Feature Fusion and Attention Mechanism

计算机科学, 2022, 49(2): 4-11. <https://doi.org/10.11896/jsjcx.210900028>

[空间众包任务的路径动态调度方法](#)

Dynamic Task Scheduling Method for Space Crowdsourcing

计算机科学, 2022, 49(2): 231-240. <https://doi.org/10.11896/jsjcx.210400249>

[蒙汉神经机器翻译研究综述](#)

Survey of Mongolian-Chinese Neural Machine Translation

计算机科学, 2022, 49(1): 31-40. <https://doi.org/10.11896/jsjcx.210900006>

基于二分图的个性化学习任务分配

谭珍琼¹ 姜文君¹ 任演纳² 张吉³ 任德盛⁴ 李晓鸿¹

1 湖南大学信息科学与工程学院 长沙 410082

2 香港教育大学特殊教育及辅导学系 香港 810014

3 之江实验室 杭州 310012

4 香港中文大学信息工程学系 香港 999077

(tanzhenqiong@hnu.edu.cn)

摘要 学习是一种复杂的事件。个体的学习效果受多方面因素的影响,且不同个体有不同的学习习惯,学生通常难以根据自身学习特点合理规划学习时间。虽然目前有关任务管理方面的研究提出了一些具有通用性的理论管理策略,但其忽略了个体间的差异性;另外,现有研究不能提供一种计算方法来形成具体的任务管理方案。针对上述问题,文中通过数据分析找出学习效率与时间因素的关联性,从而理解学生的学习特征,量化出个性化的学习效率;使用二分图的方法构建学习任务分配场景,根据不同的学习目标设计自适应效用函数,并基于此提出了一种基于迁移学习的动态分配算法 TLTA,用于为学生制定合理的任务分配方案。在真实的学生数据集上进行了大量实验,验证了所提方案的有效性及其适用性。

关键词: 二分图;任务分配;时间因素;学习效果;迁移学习

中图法分类号 TP391

Personalized Learning Task Assignment Based on Bipartite Graph

TAN Zhen-qiong¹, JIANG Wen-Jun¹, YUM Yen-na-cherry², ZHANG Ji³, YUM Peter-tak-shing⁴ and LI Xiao-hong¹

1 School of Computer Science and Electronic Engineering, Hunan University, Changsha 410082, China

2 Department of Special Education and Counselling, The Education University of Hong Kong, Hong Kong 810014, China

3 Zhejiang Lab, Hangzhou 310012, China

4 Department of Information Engineering, The Chinese University of Hong Kong, Hong Kong 999077, China

Abstract “Learning” is a complex event. Individual’s learning effect is affected by many factors. Moreover, different individuals have different learning habits. Therefore, it is challenging for students to plan their learning schedule reasonably according to their own characteristics. Although some general theoretical strategies for task management have been proposed, the differences among individuals are usually neglected. Furthermore, existing research cannot provide a calculation method to form a specific task management schedule. To this end, this paper tries to explore students’ learning characteristics by deeply studying the relation between learning efficiency and time factor through data analysis. Based on this, it quantifies personalized learning efficiency. Furthermore, it exploits the bipartite graph method to construct the learning task assignment scenario, and designs adaptive utility function according to different learning goals. Then, a dynamic allocation algorithm TLTA based on transfer learning is proposed to formulate a reasonable schedule for students. Finally, a large number of experiments are carried out on real learning datasets, and the results validate the effectiveness and applicability of the proposed work.

Keywords Bipartite graph, Task allocation, Time factor, Learning effect, Transfer learning

1 引言

在学习活动中,学生往往有大量的学习任务及碎片时间,在不同的时间段完成学习任务所获得的学习效果因人而异。

然而,大部分学生对时间及学习任务缺乏合理规划,难以根据自身学习特点有效安排学习任务。不合理的时间安排可能会降低任务完成的效率和质量,最终导致学习效果有限、学习压力增大,严重时甚至会影响学生的心理健康。因此,目前迫切

到稿日期:2021-05-18 返修日期:2021-12-09

基金项目:国家自然科学基金(62172149,61632009);之江实验室开放课题(2019KE0AB02);湖南省自然科学基金(2021JJ30137)

This work was supported by the National Natural Science Foundation of China (62172149, 61632009), Open Project of Zhejiang Lab (2019KE0AB02) and Natural Science Foundation of Hunan Province of China(2021JJ30137).

通信作者:姜文君(jiangwenjun@hnu.edu.cn)

需要一种针对学生学习特点的个性化学习任务分配方案,用于帮助学生优化学习效率,收获最佳学习效果。

在真实的学习场景中,大量的学习任务和高度碎片化的时间给学习任务分配带来了巨大挑战。针对不同类型的学习任务,一方面,每个学生的学习能力和学习效率难以量化;另一方面,在不同时间段完成不同类型的学习任务也可能产生不同的学习效果。因此,如何量化学习效率以及探明时间因素对其的影响非常具有挑战性。此外,每位学生的学习任务和空闲时间片可能会因为一些情况发生动态变化,如何对这些变化及时进行动态处理也是个性化分配方案需要应对的难点。

现有的关于学习效率的研究^[1-3]大部分只关注学习行为对学习效率的影响,忽略了学生在不同时间段进行不同类型的学习活动时学习效率也会有差异。一部分工作^[4-5]研究了时间因素对学习效率的影响,但通常只提供了理论上的策略,未给出具体的任务安排表。部分研究者^[6]通过案例分析进行个案研究,不具备通用性。随着互联网的发展,也有部分研究者^[7]使用计算机手段解决学习任务管理问题,但忽略了个体间的差异性,无法提供个性化的任务管理方案。

本文研究了学生的学习任务分配问题,将计算领域的任务分配方法应用到学习领域,辅助学生合理安排学习任务,提升学生的学习效率,增强学生效率等级。具体来说,学习任务分配涉及两个对象,即学习任务和空闲时间片,将两者进行匹配时需要满足一定的约束条件,包括持续时长、开始时间、截止时间。考虑到实际学习中在一个时间段集中完成一个任务能够达到较好的学习效果,本文将任务-时间片匹配关系约定为一对一的方式,即一个任务分配到一个时间片里完成。这个约定并不影响方法的一般性,比如,可以将剩余时间作为新的独立时间片来模拟实际情况。本文提出了一种基于二分图的个性化学习任务分配方法,该方法通过数据分析来理解学生的学习特点,构建个性化的学习效率曲线,反映学生在不同时间段下的学习效率,将学生收获的最高总效用值(由学习效率、任务紧急度、效率等级和任务完成时间4个因素决定)作为分配目标,在动态二分图上设计了一种基于迁移学习的动态分配算法(Transfer-Learning-based Task Allocation Algorithm, TLTA),为学生大量的学习任务提供了一种通用可行的个性化分配方案。

以学生甲为例来阐述本文的基本思路。例1,假设学生甲在某一天内有4个时间片段是空闲的,他需要完成4个学习任务,每个学习任务都有发布时间和截止时间,每个时间片段都有开始时间和结束时间。学生甲需要考虑满足匹配的约束条件,以收获最高总效用值为目标,将这4个学习任务安排在各时间片段中。学习任务的详细信息如表1所列,时间片的详细信息如表2所列。使用二分图来构建该问题,如图1所示。如果任务 r_1 能分别在时间片 t_1 和 t_4 中完成,那么它们之间就存在边,对应边上的权重则通过分配目标来确定,通过寻求二分图的匹配方式来获得学习任务的分配方案。此外,随着学生甲状态的变化,未来还会有新的学习任务和时间片动态到来。

表1 学习任务详细信息

Table 1 Learning task details

Task ID	Detail	Type	Start	Dead line	Duration/min
r_1	reciting words	memory	7:00	22:00	60
r_2	reading comprehension	comprehension	8:00	14:00	50
r_3	Boolean calculation	regulation	9:00	13:00	20
r_4	matrix calculation	regulation	17:10	22:00	50

表2 空闲时间片详细信息

Table 2 Time slice details

Slice ID	Start	Deadline	Duration/min
t_1	7:05	8:05	60
t_2	12:20	13:10	50
t_3	12:30	13:00	30
t_4	18:00	19:00	60

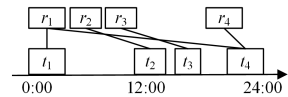


图1 任务-时间片二分图

Fig. 1 Task-time slice bipartite graph

随着网络技术的广泛应用,在线学习逐渐成熟并被广泛应用于各类学校,但由于学习内容的大量性和多样性,个体很难人为地根据自己的特征合理安排任务,目前也没有研究利用历史学习记录来量化学习效率与时间的关联性,刻画个体学习特征,为用户合理安排任务。经调研,本文首次提出用计算机手段来解决该问题。

本文方法能够优化个体的学习效率,达到事半功倍的效果。此外,该方法具有较强的通用性,可以被应用到各个领域的任务安排中,为学生、工人等提供更高的工作效率,具有很强的实用性。

本文的主要贡献如下:1)正式定义了个性化学习任务分配问题,并使用二分图方法将问题建模;2)通过分析学生的历史学习记录,量化了用户的学习行为,寻找学习效率与时间因素的关联性,计算出学生个性化的学习效率;3)以学生收获的最高总效用值为分配目标,确定了效用函数;4)设计了一种基于迁移学习的动态分配算法来解决个性化学习任务分配问题,并分析了其时间复杂度;5)在真实数据集上进行了大量的实验来验证方案的有效性和适用性。

本文第2节介绍了相关工作;第3节正式定义了个性化学习任务分配问题;第4节提出了相应的解决方法;第5节给出了实验结果及分析;最后总结全文。

2 相关工作

2.1 影响学习效率的相关因素

在教育学中有大量关于学习效率的研究,如Li等^[8]指出学习效率受到多种因素的影响,探索了外在、内在因素等对学习者的学习效果的影响。Shu^[9]提出了学习效率分析模型,该模型能够根据不同的情景来计算个体的学习效率。该研究指出不同学习的场景对个体学习效率的提高有一定的影响。文献^[4]认为学习时段是影响学习效率的重要因素,而文献^[10]则阐明了时间因素影响学习效率存在两种概念,即时间的

“量”和时间的“质”,前者指时间的延续性,是一种具有相等单位的量数,后者指个体在不同时间段的学习效率的变化。可以将时间的“质”引入生理学中人体生物钟的研究,生理学家认为个体在实际学习活动中各个时刻的学习效率是不同的,因此在不同的时间段最适合做的事情也有所区别^[11]。文献[6]指出,教师和学生如果能够了解并运用自身的生物钟规律,将有利于把握教与学的时机,提高教与学的效率。

与已有研究不同的是,本文工作并不是笼统地根据多数人的特征提供一种符合大概率人群的学习策略,而是根据个体的特征分析出最适合个体学习的时间和内容,计算出个性化的学习效率,更具实际意义。

2.2 时间管理策略

时间管理是一种重要的手段,可以帮助学习者提高学习效率和认知水平,因此有关时间管理的研究受到了各界学者的高度重视。目前相关的研究主要集中于教育学及行为心理学领域,通常采用问卷调查法、访谈法、案例分析法、实验法等对时间管理者进行分析研究。文献[11]指出大学生缺乏时间管理的能力,说明了时间管理表对大学生的重要性,指出了大学生在时间管理中存在的问题,并提出了时间管理四项图的策略以促进大学生提高自主管理能力。文献[6]通过对广东省深圳市宝安区实验学校初中部的145名学生进行抽样调查,研究了初中生的时间管理意识和时间管理行为,提出了6点优先工作制和要事优先的时间管理理念。文献[12]则对时间管理的现状进行了研究,分析了影响时间管理的众多因素,给出了一些培养自身时间管理能力的方法。

已有相关研究为学习者提供了大量时间管理的建议和策略,但都只停留于理论层面,具体时间安排表的质量仍取决于个体的时间管理能力,需要人为地把控和协调,不能直观地帮助学习者规划出一份科学合理的任务管理表。

2.3 任务分配问题

随着智能时代的到来,各种领域的匹配问题、分配问题、调度问题或安排问题等都可以利用计算机手段根据特定目标寻求最优的解决方案。目前,研究者们提出了多种方法来决策任务分配问题,并且尝试将其应用到各个领域。

解决任务分配问题最常见的方法是采用二分图构建模型,为其寻求最佳匹配。关于二分图,目前已有大量的研究,可以根据二分图的顶点是否变化将其分为静态二分图^[13]、单边动态二分图^[14]和双边动态二分图^[15-16],也可以根据匹配目标将其分为最大、最小匹配^[17-18]和最优匹配^[19-20]等。另外,还有研究者利用多目标优化^[21]、增强学习^[22-23]、遗传算法^[24]、迁移学习^[25-26]等多种方法来解决任务分配问题。

目前,研究者们将任务分配应用到了许多场景,如网约车平台最大化利益接单问题^[27];为飞机装配线制定多阶段工作站的工人分配,实现稀缺人力资源的有效安排问题^[28];外卖平台为改善食品配送体验寻求最优配送路线问题^[29];个体为提高工作效率解决拖延症问题^[7];基于任务分配和数据集副本的科学 workflow 数据布局问题^[30];基于机器人能力和群体协作相关性的任务分配问题^[31]等。但目前针对大学生的学习任务设计分配算法的研究尚不成熟。

3 问题描述

本节首先对相关概念进行了阐述,然后正式定义了个性化学习任务分配问题,并确定分配目标及分配约束条件。

3.1 概念定义

本节列出了文中所用到的符号及对应描述,如表3所列,介绍了学生用户、历史学习记录、空闲时间片、学习任务及学习效率的基本概念。

表3 符号说明

Table 3 Symbol description

Symbol	Description
S	学生集合
R	学习任务集
T	空闲时间片集
D	历史学习记录集
C	个性化学习效率集
$type_r$	学习任务类型
du, du_t	学习任务及时间片的时长
$start, end_r$	学习任务发布时间与截止时间
$start_t, deadline_t$	时间片开始时间与结束时间
$urgent_{rt}^u$	用户 u 的任务 r 紧迫度
$ahead_{rt}^u$	用户 u 的任务 r 与时间片 t 匹配对应的提前完成时间
$efficiency_{rt}^u$	用户 u 的任务 r 与时间片 t 匹配对应的学习效率
$level_{rt}^u$	用户 u 的任务 r 与时间片 t 匹配对应的效率等级
$Weight_{rt}^u$	用户 u 的任务 r 与时间片 t 匹配对应的效用值
$Total^u$	用户 u 的个性化总效用值
$Assignment^u$	用户 u 的个性化任务安排表

定义 1(学生用户) 每个学生用户 $s \in S$, 拥有 3 个数据集:学习任务集 R 、时间片集 T 和历史学习记录集 D 。本文通过分析学生的学习记录集 D , 为其每个学习任务寻找合适的学习时间片, 从而提供一个最佳任务安排表 $Assignment$ 来辅助学生完成所有的学习任务。

定义 2(历史学习记录) 每条历史学习记录 $d \in D$, 包含用户 ID 、任务类型、任务时间段、任务耗时、学习结果等信息, 记录了学生用户在网络平台上相关的历史学习行为。

定义 3(空闲时间片) 指学生除去正常作息、必要任务后的一些碎片化的空闲时间段。每个空闲时间片 $t \in T$, 包含时长、开始时间、结束时间等信息, 在分配过程中空闲时间片集 T 能根据学生取消、增加学习活动等情况动态变化。

定义 4(学习任务) 指除学校制定的课程外, 学生可自行安排的学习任务(如自学课程、课后作业等)。每个学习任务 $r \in R$, 包含任务类型、任务时长、任务发布时间、任务截止时间等信息, 分配过程中学习任务集 R 能根据学生状态动态变化。本文参考文献[28]中的研究, 根据学生的认知过程, 将学习任务分为以下 4 类。

(1)记忆型。该类任务需要用户再认、再现之前已习得的知识, 或需要对规则、公式、事实等进行记忆, 它与所隐含的数学知识无关联。

(2)规则型。该类任务需要用户利用规则、公式等来解决问题, 只需利用用户的些许认知能力, 不需要理解、解释其中深层次的数学知识。

(3)理解型。该类任务需要用户理解其中蕴含的意义,

需要利用一定的认知能力进行有意识的思考,以关联相关数学知识,做出正确推理和迁移学习。

(4)情景丰富型。该类任务是日常任务的模拟或再现,情境信息丰富且复杂,因此,需要用户具有复杂思维和发散性思维,去探索、理解以及在自我认知过程中进行自我调节与监控。

定义 5(个性化学习效率) 个体在实际学习活动中各个时刻的学习效率是不同的,在不同时间段最适合做的事情也有所区别。本文通过分析学生的学习记录集来理解学生在不同时刻不同任务上的学习特点,从而为学生构建较为细化的个性化学习效率。

3.2 个性化学习任务分配问题的定义

给定学生集合 S 和这些学生的历史学习记录 D , 每名學生都有需要完成的学习任务集 R 和空闲的时间片集 T , 在不同时间片完成任务获得的学习效率有所差别。通过分析学生的历史学习记录 D , 根据分配目标求得效用函数, 目的是找到 R 与 T 的匹配集 $Assignment \subseteq R \times T$, 使学生收获最高总效用。分配过程中任务集 R 、时间片集 T 会根据每名学生的状态动态变化, 未来的学习任务和时间片信息无法预知。

(1)分配目标为使学生收获最高总效用值,其计算公式如下:

$$Maxsum(Assignment) = \sum_{r \in R, t \in T} Weight_n$$

(2)约束条件如下:

- 1)任务必须分配给在其发布日期之后、截止日期之前的时间片;
- 2)任务不能分配给时长小于完成任务所需时间的的时间片;
- 3)已分配的任务(时间片)不再参与分配;
- 4)任务与时间片是一一对应的匹配关系。

4 解决方案

本文方法的整体框架如图 2 所示。A 部分包含了通过在线学习平台和利用调查问卷的方式收集到的所有学生数据集,将学生数据集划分为历史学习记录集、学习任务集和空闲时间片集;B 部分将对学生的历史学习记录集进行预处理后,再进行数据分析,从而获取学生的学习特征,计算个性化的学习效率;C 部分则对学习任务集和时间片集进行预处理后,通过控制学习任务和时间片从数据集中动态到来的顺序来构建动态二分图,再根据分配目标制定效用函数,从而确定二分图边上的权重;D 部分则结合 B 部分和 C 部分的工作来设计分配策略,提出了一种动态分配算法,以得到一个合理的个性化学习任务安排表。

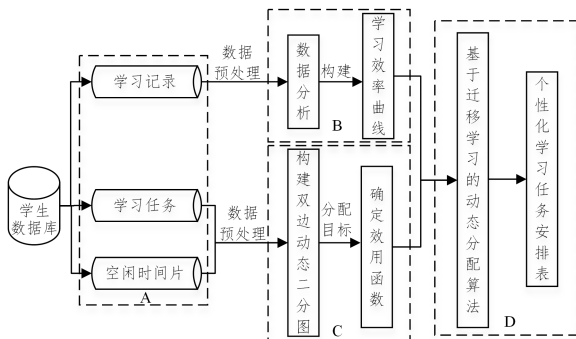


图 2 个性化学习任务分配整体框架图

Fig. 2 Overall framework of personalized learning task assignment

因此,本文的解决方案可以分为 3 个部分:

- (1)量化学习效率,通过数据分析获取学生的学习特征,计算个性化的学习效率;
- (2)构建动态二分图,根据分配目标确定效用函数;
- (3)设计分配策略,提出基于迁移学习的分配算法,将学习任务与时间片进行匹配。

4.1 量化学习效率

4.1.1 数据分析

生理学家认为人体在一天中有学习高效期,如果使用得当,则可以轻松自如地掌握、消化和巩固知识。个体都有自己独特的学习规律和习惯,把握好高效的学习时段往往能收获好的学习效果^[32]。这表明学习者在不同时间段有不同的学习效率,并且不同任务类型间的学习效率也存在差异。因此,需要根据个体的学习特征,即每个人一天中大脑最适合学习的时间段不同,来安排学习任务和学习时间。

本节通过对学生大量的历史学习记录进行分析,将学生的学习特征进行量化,找出学习效率与时间因素的关联,然后为学生的每类学习任务都计算出基于时间的个性化学习效率,以反映学生在不同时间段学习效率的变化。

首先,本文根据常识将一天划分为 4 个时间段,时间段说明如表 4 所列。

表 4 时间段说明

Table 4 Time period description

ID	Period	Span	Duration/h
1	morning	6:00-8:00	2
2	am	8:00-12:00	4
3	pm	12:00-18:00	6
4	night	18:00-24:00	6

其次,为了更好地对学生的学习行为进行分析,本文根据学习特征将学生划分为以下 4 种类型。

- (1)百灵鸟型:这类学生一到清晨大脑就异常活跃,学习高效期在早上。“一日之计在于晨”就是描述这类学生的。
- (2)猫头鹰型:这类学生一到晚上脑神经细胞就进入兴奋状态,学习高效期在晚上。
- (3)混合型:这类学生面对不同学习任务的学习高效期都不同,大多数学生属于这一类型。
- (4)稳定型:这类学生在不同时刻的学习效率都比较稳定,没有太大波动,因此对这类学生来说学习效率受时间因素的影响不大。

本文收集了某大学 4 个班级共 134 名学生在超星学习平台上的 11220 条学习记录,并提取了每个学生的日志信息,通过分析这些学习记录发现,学生在进行学习活动时确实存在一些问题。

首先,由图 3 所示的统计数据可知,72.59% 的学习任务普遍是在下午和晚上完成的,而早上完成的学习任务只占到了 1.71%。该结果也真实地反映了大学生的学习现状,一方面学生的课程一般上午较多,空闲时间主要分布在下午或晚上;另一方面也反映出部分学生早上起床困难,喜欢在下午或者晚上完成学习任务。

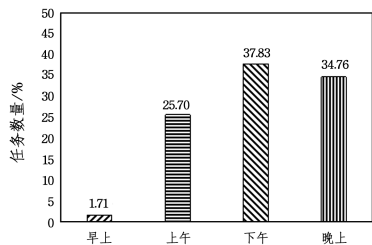


图 3 各时间段的任务完成数量

Fig. 3 Number of tasks completed in each time period

其次,对学生历史行为记录进行分析,按 4 个类别统计不同类型的学生数量,如图 4 所示。其中 55.22% 的学生属于混合型,其次是夜猫子型,百灵鸟型和稳定型的学生各自只占到了 5.22%。

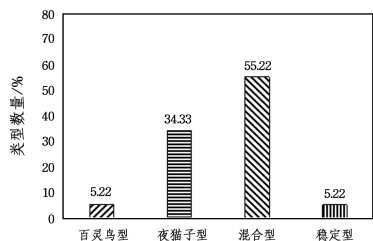
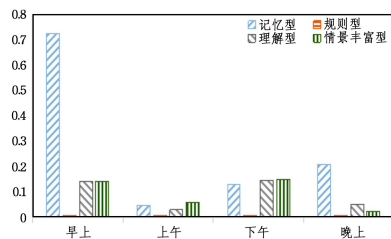


图 4 不同类型的学生数量

Fig. 4 Number of students in each category

这两项分析结果说明了每位学生都有不同的学习习惯,学生的学习高效期也受到学习任务类型和时间因素的影响,同时也反映了为不同类型的学生提供一个合理且具有个性化的任务安排表的必要性。

需要说明的是,将学生划分为 4 种类型的目的是从整体上把握学生学习策略的主要特征。在具体分配任务时,我们是根据每个学生的效率曲线特征进行分配的,即每个学生的效率曲线都是个性化的,即便是同一类型的学生,其效率曲线也不同。另外,本文将一天大致分为 4 个时段,实际上还可



(a) 百灵鸟型

根据数据情况进行更细粒度的划分,若有充分的数据,则可获取学生每小时的学习效率。

4.1.2 构建个性化的学习效率曲线

本节参考 Su 等^[2]提出的学习效率量化公式来定义学生不同时段、不同任务类型的学习效率。

$$\text{学习效率} = \frac{\text{学习的产出}}{\text{学习的投入}} \quad (1)$$

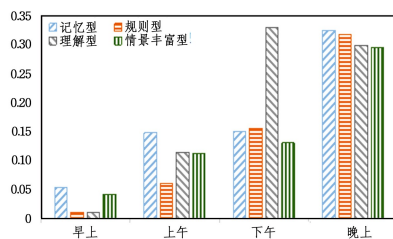
以式(1)为基础,本文将学习成绩作为学习的产出,将所花费的学习时间作为学习的投入。为研究学习效率和学习时段的相关性,本文定义了不同时段下的学习效率量化公式,可计算出学生在某一类任务 c 中的学习效率。

$$Eff_{p(1\sim 4)}^c = \frac{\sum_{i=1}^n (\frac{result_p^c}{time_p^c})}{n_p^c} \quad (2)$$

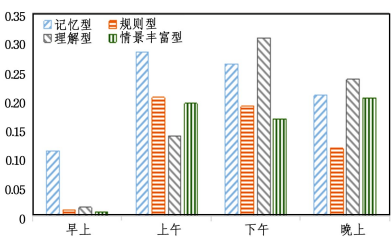
其中, Eff_p^c 表示在任务类型 c 中 p 时间段(早上、上午、下午、晚上)的学习效率, $result_p^c$ 表示在任务类型 c 中 p 时间段的学习成绩, $time_p^c$ 表示 p 时间段实际花费的学习时间, n_p^c 表示 p 时间段的学习任务量。以学生甲为例,假设他在上午花了 30 min 打卡英语单词取得的百分制成绩是 90 分,花了 25 min 默写一篇古诗词取得的百分制成绩是 80 分,那么学生甲上午在记忆型任务中的学习效率为:

$$EFF_{\text{记忆型}}^{\text{甲}} = (90/30 + 80/25)/2 = 3.1$$

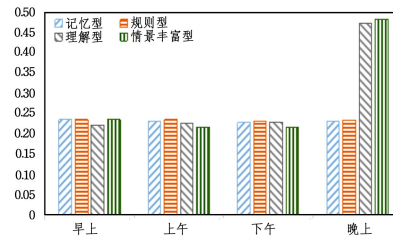
根据式(2)可以求出学生分别在早上、上午、下午、晚上这 4 个时间段完成每类学习任务所对应的学习效率。如图 5 所示,将 134 名学生根据学习特征划分为 4 种类型,观察每种类型学生在不同任务类型下的平均表现效果:百灵鸟型学生在早上完成各类任务时都能获得最高的学习效率;夜猫子型则相反,随着时间的推移,学习效率逐渐提高,在晚上完成各类任务的效率最高;而混合型的学生在不同时间段完成各类学习任务呈现不同的学习效果;稳定类型的学生在不同时刻的学习效率大多呈稳定状态。



(b) 夜猫子



(c) 混合型



(d) 稳定型

图 5 4 种类型学生的平均学习效率

Fig. 5 Average learning efficiency of four types of students

为了更直观地展示个性化学习效率曲线,本文从 4 种类型

中分别选择了一位学生进行展示。如图 6 所示,可以发现 4 位

学生的效率曲线都具有个性化且符合他们所属类型的特征。

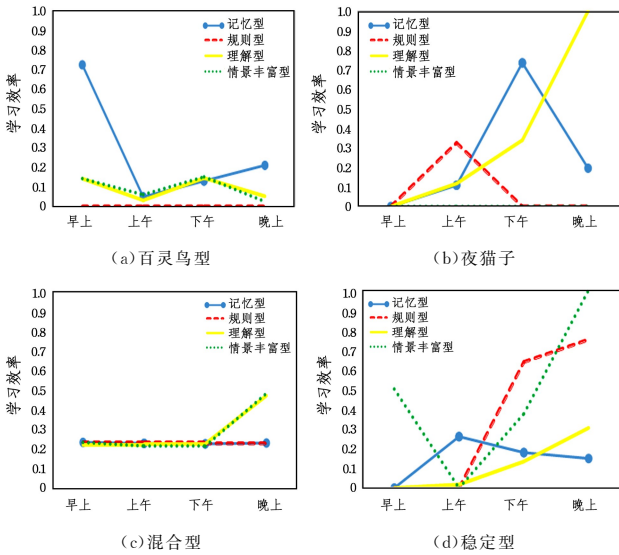


图 6 个性化学习效率曲线图

Fig. 6 Personalized learning efficiency curve

4.2 构建个性化动态二分图

4.2.1 使用二分图方法对学习任务分配问题建模

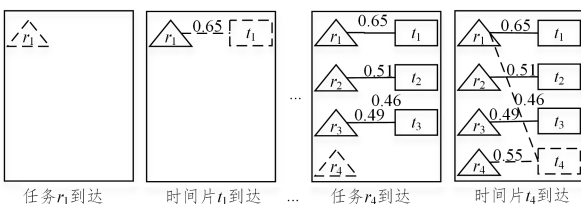
根据个性化学习任务分配问题的定义,本文使用二分图的方法来构建模型。以例 1 学生甲来示范,将他的学习任务集中陆续到来的任务作为二分图的左顶点,构建左顶点集 R' ;将时间片集中陆续到来的时间片作为二分图的右顶点,构建右顶点集 T' 。若学习任务与时间片满足约束条件,则将它们之间存在的边作为二分图边集 E ,对应的权重 W 由任务与时间片的效用函数计算决定,从而得到二分图 $G=(R', T', E)$ 。

考虑到在真实学习环境中,学生的学习任务和空闲时间片可能会因受到不确定因素的影响而发生变化。因此,在分配过程中,应允许任务与时间片根据学生的状态动态变化,二分图的节点可以动态到来,一旦一项新任务/新时间片到达,与其有关的边和对应权重也会确定。此外,当一个分配被确定后,分配结果不可更改。表 5 列出了例 1 中学生甲一天内的学习任务与时间片的动态到来顺序,根据该表可以构建出双边动态二分图。图 7 给出了针对表 5 中学习任务和时时间片部分节点动态到达时二分图的构建过程。

表 5 学习任务和时间片的到达顺序

Table 5 Learning task and time slice arrival order

time	7:00	7:05	8:00	9:00	12:20	12:30	17:10	18:00
order	r_1	t_1	r_2	r_3	t_2	t_3	r_4	t_4



注:虚线框表示当前到达的任务或时间片,与其有关的边也用虚线区分

图 7 动态二分图的构建过程

Fig. 7 Construction process of dynamic bipartite graph

4.2.2 确定效用函数

制定学习任务安排表的关键是确定任务优先级。本文参考了众多时间管理及任务管理的相关研究,如美国管理学家史蒂芬·科维提出的时间管理四项法^[33],根据任务紧急程度和重要程度来决定任务优先级;美国管理学家泰罗提出的科学管理方法^[34],经过对时间和动作进行科学地的研究来合理分配任务;苏联柳比歇夫提出的柳比歇夫时间管理法^[35],如实记录时间耗费情况,对记录进行分析,使人们能正确认识自己的时间利用状况,并养成管理时间的习惯。

本文将目前常用的时间管理方法与学习场景相结合,以学生获得最大总效用值为分配目标,将学习效率、任务紧急度、提前完成时间和效率等级作为分配目标的 4 个影响因素,从而确定效用函数,决定任务分配中的优先级。下文给出 4 个影响因素的计算方法。

(1)学习效率。首先根据用户 u 的学习任务的类型 c 确定对应的学习效率曲线 $F_{(c)}$,然后在学习效率曲线上找到每个时间片对应的区间;通过积分函数求出在该区间上所获得的学习效率,如式(3)所示:

$$efficiency_{r_n}^c = \sum_{p=(1,2,3,4)} \int_{x_p^-}^{x_p^+} F_p^c(x) dx \quad (3)$$

其中, $p=\{1,2,3,4\}$ 表示对应的(早上、上午、下午、晚上)时间段, $F_p^c(x)$ 表示用户 u 在任务类型 c 中 p 时间段的效率曲线函数。

(2)任务紧急度。用于衡量任务的轻重缓急,若该任务的发布时间逼近任务的截止时间,则表示该任务需要紧急完成,急迫度很高,否则表示该任务急迫度不高,计算方法如式(4)所示:

$$urgent_r^n = deadline_r - (start_r + du_r) \quad (4)$$

其中, $deadline_r$ 表示任务 r 的截止时间, $start_r$ 表示任务 r 的开始时间, du_r 表示任务 r 的持续时长。

(3)提前完成时间。将学习任务安排在某一时间片中去完成,通过将该任务的真实完成时间与该任务的截止时间进行对比即可知该任务提前完成的时间,计算式如式(5)所示:

$$ahead_r^n = deadline_r - (start_t + du_r) \quad (5)$$

式(5)是用任务 r 的截止时间减去时间片 t 的开始时间和任务 r 的持续时长之和,其差值为任务 r 提前完成的时间,其中 $start_t$ 表示时间片 t 的开始时间。

(4)学习效率等级。该等级的衡量由学生的效率图决定。若为任务安排的时间片在效率图上属于高效时间段,则认为该安排符合学生的偏好特征,学生能获得较高的效率等级;若为任务安排的时间片在效率曲线上属于低效时间段,则认为该安排违反了学生的偏好特征,学生获得较差的效率等级。用户效率等级的计算式如式(6)所示:

$$level_r^n = rank \left(\int_{x_p^-}^{x_p^+} F(x) dx \right)_{p=(1,2,3,4)} \quad (6)$$

$rank$ 函数将这 4 个时间段的效率值从大到小排序(输出结果用 1,2,3,4 表示,效用最大的是 4,最小的是 1)。 $start_t$ 可用于判断所属的 p 时间段,得到该时间段通过 $rank$ 函数排序后的效率等级。由于效率等级是由用户的效率曲线决定的,因此若给用户安排效率等级越高的任务,用户将有望更高效地完成,相应地也会提升用户的学习愉悦感。

将这 4 个影响因素利用线性比例变换法 $y_i = x_i / \max(x)$

进行归一化处理,确保值的范围都保持在 $[0,1]$ 之间。然后,选择一组权值 $\lambda = \{\gamma_1, \gamma_2, \gamma_3, \gamma_4\}$ 来决定这4个影响因素对效用函数的重要程度,从而计算效用函数,如式(7)所示:

$$\begin{aligned} Weight_n^n = & \gamma_1 \times efficiency_n^n + \gamma_2 \times urgent_n^n + \\ & \gamma_3 \times ahead_n^n + \gamma_4 \times level_n^n \end{aligned} \quad (7)$$

4.3 设计个性化分配策略

本节给出了个性化学习任务分配算法 TLTA,该算法使用二分图方法来构造模型,根据分配目标为学习任务匹配合适的时间片,每找到一对匹配,就将该匹配对 (r, t) ($r \in R, t \in T$)添加到任务安排表 *Assignment* 中,从而得到最佳任务安排表。

从学生的任务集和时间片集中随机选取节点,并将其加入二分图,以此模拟学习任务和时间片节点动态到来的过程。考虑到学生的学习活动具有日常性与重复性,很多学习任务非常相似,因此加入迁移学习,将学生已经完成的学习任务经验迁移给相似且未完成的学习任务,以节省寻找匹配的时间,加速分配过程。具体过程如算法1所示。

算法1 TLTA

输入:学生的任务集 *R*,空闲时间片集 *T*,效率曲线集 *C*

输出:最佳任务安排表 *Assignment*,总效用值 *Total*

1. 初始化:二分图 $G \leftarrow (R', T', E)$
2. $R' \leftarrow \emptyset, T' \leftarrow \emptyset, E \leftarrow \emptyset$
3. 源任务库 $SourceLibrary \leftarrow \emptyset$
4. for 当新的节点 *v* (学习任务或空闲时间片)到来 do
5. if $v \in R$ then
6. $R' \leftarrow R' \cup v$
7. $sourcetask \leftarrow Similar(SourceLibrary, v)$ //调用子算法1
8. if *sourcetask* 存在 then
9. $t \leftarrow Transfer(sourcetask, v)$ //调用子算法2
10. else
11. 节点 *v* 自学,计算与时间片的效用 $Weight(v, T')$
12. $t \leftarrow Greedy(v, T')$
13. if *t* 存在 then
14. $SourceLibrary \leftarrow SourceLibrary \cup v$
15. if *t* 存在 then
16. $Assignment \leftarrow Assignment \cup (v, t)$
17. $Total \leftarrow Total + Weight_{v,t}$
18. else
19. $T' \leftarrow T' \cup v$
20. $r \leftarrow Greedy(R', v)$
21. $Assignment \leftarrow Assignment \cup (r, v)$
22. $Total \leftarrow Total + Weight_{r,v}$
23. return *Assignment*

当新的学习任务到来时(算法1中第4-17行),根据情况对其进行分配。首先,将到来的学习任务加入到二分图左顶点集(算法1中第6行)。然后,通过调用子算法1对源任务库中的学习任务与到来的学习任务进行相似性判断(算法1中第7行),从而将任务分为两部分进行处理:若源任务库中存在相似任务(算法1中第8-9行),则调用子算法2将相似源任务的学习经验迁移给该学习任务,并利用此经验在已到来的时间片中找出满足条件且效用值最大的时间片;若不存在相似任务(算法1中第10-17行),则该任务自学,需要先依次计算其与每个时间片匹配所能获得的效用值,再使用贪婪策略找出效用值最大的时间片(算法1中第12行),并将

该任务加入源任务集中(算法1中第13-14行)。最后将返回的时间片与当前学习任务进行匹配,将该匹配对加入到 *Assignment* 中并累加效用值(算法1中第15-17行)。

当新的时间片到来时(算法1中第18-22行),首先,将该时间片加入到二分图右顶点集(算法1中第19行);然后,使用贪婪策略找出与该时间片匹配效用值最大的待分配学习任务(算法1中第20行);最后将该匹配对加入到 *Assignment* 中并累加效用值(算法1中第21-22行)。

衡量相似度是进行迁移学习的前提^[36],只有当两个任务具有一定的相似性时才能达到较好的效果。若在两个任务的相似度并不高时就进行迁移,则可能会造成反作用,导致匹配错误。子算法1通过一系列度量指标来衡量两个任务是否相似,以解决上述问题。

子算法1 Similar(*SourceLibrary*, *v*)

输入:源任务库 *SourceLibrary*,目标任务 *v*

输出:相似源任务 *sourcetask*

初始化: $mindistance \leftarrow \infty, sourcetask \leftarrow -1$

1. for $k \leftarrow 1$ to $|SourceLibrary|$ do
2. $source \leftarrow SourceLibrary[k]$
3. if $type_r^v = type_r^{source}$ then
4. $d1 \leftarrow normalization(|du_r^v - du_r^{source}|)$
5. $d2 \leftarrow normalization(|start_r^v - start_r^{source}|)$
6. $d3 \leftarrow normalization(|deadline_r^v - deadline_r^{source}|)$
7. $distance \leftarrow d1 + d2 + d3$
8. if $mindistance > distance$ then
9. $mindistance \leftarrow distance$
10. $sourcetask \leftarrow source$
11. return *sourcetask*

首先进行初始化(子算法1中第1行),*mindistance* 初始值是一个月内两个任务间差异的最大值,用 ∞ 表示;然后循环遍历整个源任务库(子算法1中第2-11行),确定衡量任务相似度的特征,在确保任务类型一致的前提下(子算法1中第4行),比较两个任务间的开始时间、截止时间及持续时长的近似度(子算法1中第5-8行);最后从源任务库中选择一个与之类型一致且差异最小的任务作为相似任务(子算法1中第9-11行),若能找到这样的任务,就存入 *sourcetask* 中,否则 *sourcetask* 的值为-1。

将源任务 *sourcetask* 的学习经验迁移给目标任务 *v* (文中的经验指对于某类任务最适合的时间片特征),然后目标任务 *v* 利用该学习经验在已到达的时间片中寻找效用值最大的时间片进行匹配,具体如子算法2所示。

子算法2 Transfer(*sourcetask*, *v*)

输入:具有学习经验的源任务 *sourcetask*,无学习经验的目标任务 *v*,

已到来时间片集 *T'*

输出:与目标任务 *v* 匹配的时间片 *t*

1. 初始化: $feature_1 \leftarrow 0, feature_2 \leftarrow 0, max \leftarrow 0$
2. 获取 *Assignment* 中与源任务匹配的时间片 *timeslice*
3. 提取 *timeslice* 的特征:
4. $feature_1 \leftarrow start_{timeslice}$ 的24小时制时间
5. $feature_2 \leftarrow start_{timeslice}$ 的24小时制时间
6. $Slice \leftarrow T'$ 中在 $[feature_1, feature_2]$ 区间上的时间片
7. for $k \leftarrow 1$ to $|Slice|$ do
8. $w \leftarrow Weight(v, Slice[k])$ // *v* 与 *Slice[k]* 的效用值

```

9. if  $w > \max$  then
10.    $\max \leftarrow w$ 
11.    $t \leftarrow \text{Slice}[k]$  //将效用值最大的时间片赋值给  $t$ 
12. return  $t$ 

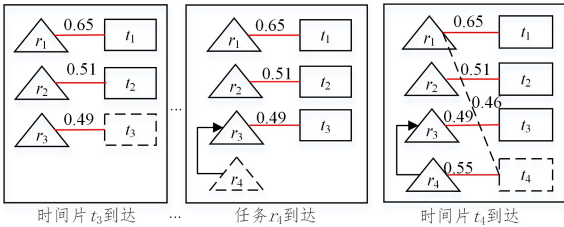
```

首先获取与源任务匹配的时间片(子算法 2 中第 2 行),通过提取该时间片的特征(子算法 2 中第 3-5 行)可知源任务在哪段时间区间能获得较高的学习效率,将已到来时间片集 T' 中满足该区间的时间片与目标任务计算效用值(子算法 2 中第 6-8 行),选择效用值最大的时间片进行匹配(子算法 1 中第 9-11 行)。此方法无须将每一个到达的时间片都与目标任务计算效用值,缩短了运算时间,加速了分配过程。

任务 $r \in R'$, 时间片 $t \in T'$, 则构建的源任务库 $SourceLibrary \in R'$ 。判断源任务库中是否存在相似学习任务的时间复杂度为 $O(|SourceLibrary|)$, 该复杂度取决于源任务库的大小, 学习任务集中各任务之间的相似度越高, 需添加到源任务库的任务数量就越少, 相应地, 时间复杂度也就越低。根据新到来的学习任务是否与源任务库中的任务相似可分为两种情况: 若不能找到相似任务, 则自学使用贪婪策略进行匹配, 时间复杂度为 $O(|T'|)$; 若能找到相似任务, 则使用迁移学习进行匹配, 时间复杂度为 $O(|Slice|)$ 。该复杂度取决于满足特征区间的时间片数量, 数量越少, 复杂度越低。由于 $Slice \subseteq T'$, 因此 $O(|Slice|) \leq O(|T'|)$, 说明使用迁移学习比使用贪婪策略更节省时间。

4.4 运行样例

将 TLTA 算法应用于例 1 来展示本文学习任务动态分配的过程, 如图 8 所示。根据表 5 所列的学习任务和到达顺序可知, 到达的节点分别是 $r_1, t_1, r_2, r_3, t_2, t_3, r_4, t_4$ 。



注: 当前到达的任务或时间片用虚线框区分, 若任务与时间片匹配成功则用红色实线, 若存在边但未匹配成功则用灰色虚线

图 8 TLTA 算法的运行示意图(电子版为彩色)

Fig. 8 Operation diagram of TLTA algorithm

初始时, 二分图左右顶点集和源任务库 $SourceLibrary$ 中都为空。

当 r_1, r_2, r_3 陆续到达时, 将它们都加入到二分图左顶点集, 依次调用子算法 1 来判断是否能在源任务库 $SourceLibrary$ 中寻找到与其类型一致且相似的任务。由于 r_1 是记忆型任务, r_2 是理解型任务, r_3 是规则型任务, 而当时源任务库 $SourceLibrary$ 中不存在这些类型的任务, r_1, r_2, r_3 都需要自学。分别计算 r_1, r_2, r_3 与二分图右顶点集中已到达时间片匹配的效用值, 然后为其选择效用值最大的时间片进行匹配, 将得到的匹配对 $(r_1, t_1), (r_2, t_2), (r_3, t_3)$ 加入到 $Assignment$ 中, 并将 r_1, r_2, r_3 加入到源任务库中。

当 r_4 到达时, 同样先将其加入到二分图左顶点集, 再调用子算法 1 来判断是否能在源任务库 $SourceLibrary$ 中寻找到与 r_4 类型一致且最相似的任务。结果发现 r_4 与源任务库

中的 r_3 非常相似, 因此调用子算法 2 将 r_3 的经验迁移给 r_4 。 r_4 利用该经验寻找到 t_4 进行匹配, 然后将匹配对 (r_4, t_4) 放入 $Assignment$ 中。最后, 得到任务安排表 $Assignment = \{(r_1, t_1), (r_2, t_2), (r_3, t_3), (r_4, t_4)\}$, 总的效用为 $Total = 0.65 + 0.51 + 0.49 + 0.55 = 2.2$ 。

5 实验分析

5.1 实验设置

实验环境如下: 处理器为 3.4GHz Inter(R) Core(TM) i5-7500, 内存为 8GB, 操作系统为 Windows 10, 编程语言为 C++。

本文收集了两类数据: 一类是历史学习记录, 具体为某大学超星学习平台上 2020 年 4 个班级共 134 名学生的课程记录, 共有 11 220 条历史学习行为记录, 包含了 35 种不同的任务内容; 另一类是通过调查问卷的方式收集到的这 134 名学生一个月内所有要完成的学习任务及空闲时间片的详细信息。

5.1.1 数据预处理

(1) 按阶段分配: 考虑到每名学生在在学习期间的学习任务非常大, 而且未来的学习任务与空闲时间片无法精确预计, 因此将任务与时间片分阶段划分, 按每月进行分配(上月未完成的任务延续到下一个月完成)。

(2) 时间片预处理: 对时间片的长度进行限制, 将短于 5 min 的时间片丢弃, 对超过 6 h 的时间片进行强制拆分。将 6 点到 24 点的时间片认定为正常学习时间。

(3) 学习任务预处理: 将时长大于 6 h 的学习任务拆分成小的学习任务。另外, 需要根据任务的内容划分任务所属类型, 在本文的数据集中, 将 35 种任务内容按照本文第 2 节定义 4 的认知过程划分为 4 种任务类型, 如表 6 所列。

表 6 数据集任务类型划分

Table 6 Data set task type division

Task type	Include task details
memory	线性结构记忆题、数据结构概述记忆题等
regulation	高数矩阵运算、数字逻辑-布尔运算等
comprehension	线性结构理解题、四级阅读理解训练等
scene	树的自测题应用、栈和队列应用层次等

5.1.2 参数设置

实验的参数设置如表 7 所列, 加粗的数据为参数默认取值。

表 7 实验数据参数

Table 7 Parameters of experimental data

Parameters	Value
Students type	百灵鸟型, 夜猫子型, 混合型, 稳定型
$ R $ 一月学习任务数量	50, 100, 200, 400, 600
$ T $ 一月时间片数量	50, 100, 200, 400, 600
	$\lambda_1 = \{0.25, 0.25, 0.25, 0.25\}$
	$\lambda_2 = \left\{ \alpha, \frac{1-\alpha}{3}, \frac{1-\alpha}{3}, \frac{1-\alpha}{3} \right\}$
各因素对效用函数的占比 $\alpha \in (0, 1)$	$\lambda_3 = \left\{ \frac{1-\alpha}{3}, \alpha, \frac{1-\alpha}{3}, \frac{1-\alpha}{3} \right\}$
	$\lambda_4 = \left\{ \frac{1-\alpha}{3}, \frac{1-\alpha}{3}, \alpha, \frac{1-\alpha}{3} \right\}$
	$\lambda = \left\{ \frac{1-\alpha}{3}, \frac{1-\alpha}{3}, \frac{1-\alpha}{3}, \alpha \right\}$

5.2 实验结果

5.2.1 参数敏感性测试

本组实验通过改变不同参数的值(未改变的参数使用表7中的默认值)来观察实验结果并进行分析。

(1)效用函数的影响因素变化

由第3节中的式(7)可知,效用函数有4个影响因素,分别是学习效率、任务急迫度、任务提前完成和学习效用等级。本实验通过改变这4个影响因素对效用函数的占比来进行测试。考虑到各因素对效用函数的不同影响程度,本文将其为5种情况,其中, $\lambda_1 = \{0.25, 0.25, 0.25, 0.25\}$ 表示4个影响因素对效用函数的影响程度均占25%,影响程度相同; $\lambda_2, \lambda_3, \lambda_4, \lambda_5$ 则分别表示某个影响因素的变化情况,具体由 α 决定, α 的取值范围为(0,1)(由于 α 取0或1时,其只依赖于一个因素,不符合本文的分配目标,因此取值不包含0和1)。当某因素对应位置上 α 的取值大于0.25时,该因素对效用函数的影响更大。

影响效用函数的参数值变化如图9所示。从图9(a)可以看出,由于 α 决定了 λ_2 学习效率的占比,因此随着 α 变小,它所获得的学习效率降低,而效率等级由效率求得,因此效率等级也随之下降;而随着 α 变小,其他位置的 $(1-\alpha)/3$ 会增大,任务急迫度降低,完成时间也会缩短。图9(b)一图9(d)给出了类似的效果。另外,通过观察 λ_1 在各指标上的表现效果可知,其在各影响因素占比相同时的表现一般。

由于本文主要考虑对学习效率的优化,因此,在后续实验中,将 λ_2 作为效用函数,将图9(a)中相交线附近的0.7作为后续实验中 α 的取值。在真实应用场景中,根据用户对任务完成的需求来决定效用函数及各因素的影响程度。

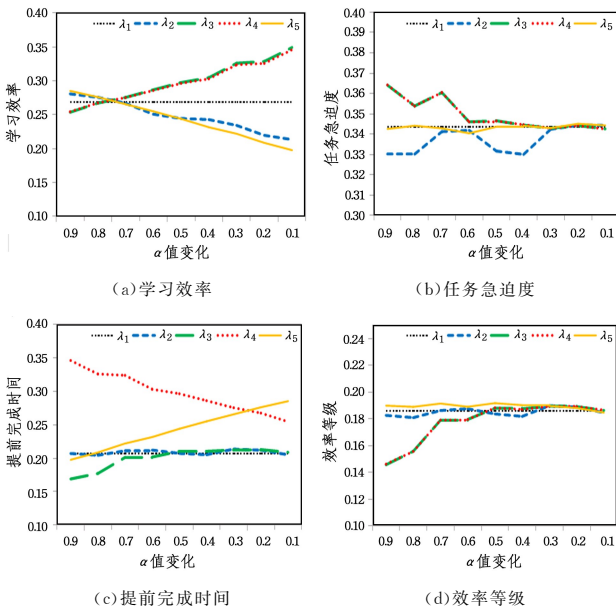


图9 影响效用函数的参数值变化

Fig. 9 Change of parameter value affecting utility function

(2)学习任务量的变化

本文以一个月为单位对学生的学习任务进行分配,因此将任务量控制在50到600之间。学习任务量的变化如图10所示,随着任务量逐渐增加,获得的总效用、学习效率、效率等级

等等都呈递增态势。本实验的时间片量为默认值600,这说明学生学习任务量的增多可以使空闲时间得到更好的利用,可以让学生收获更大效用。

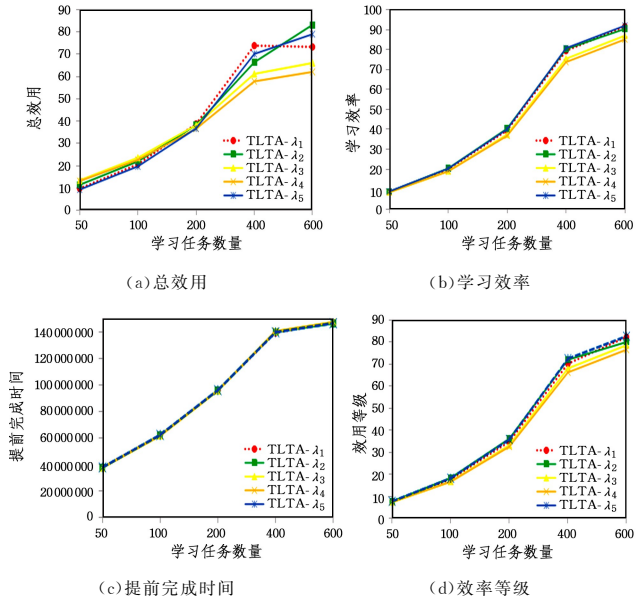


图10 学习任务量变化

Fig. 10 Changes in the amount of learning tasks

(3)时间片量的变化

同样,考虑以一个月为单位,将学生空闲时间片量也控制在50到600之间。实验结果如图11所示,随着时间片量的增加,获得的总效用、学习效率、效率等级等等都呈递增的态势。这说明随着学生空闲时间片的增加,学习任务能够更好地完成。从图11(a)一图11(d)这4幅图中的曲线趋势可以发现,当时间片的数量增加到一定程度时,总效用、学习效率、效率等级等都逐渐趋于平稳,原因是随着空闲时间量的增加,学生的学习任务都得到了很好的分配,再多的时间片也就无意义了。

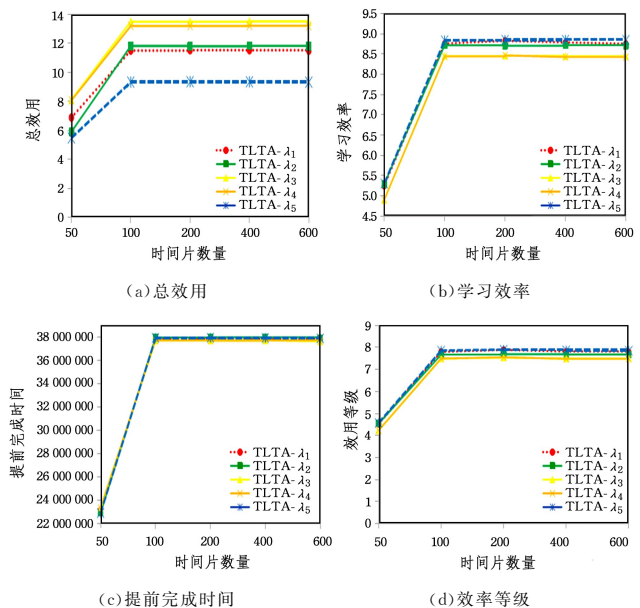


图11 时间片量变化

Fig. 11 Changes in the amount of time slices

5.2.2 对比实验

为了评估本文提出的为学生提供任务安排表的方法的适用性,本组实验选取了目前常用的较好和较新的算法与本文方法在不同指标上进行对比。

常用于二分图匹配的算法有 GREEDY 算法^[37]、FCFS 算法^[38]和 KM^[39]算法,目前基于二分图解决任务分配的较好算法是 Tong 等提出的 TGOA 算法^[29]。因此,本文选择了 GREEDY 算法、FCFS 算法、KM 算法和 TGOA 算法作为本组实验的对比算法。

(1)GREEDY 算法指在每一次匹配时都只顾眼前利益,使用贪婪策略选择当前权重最大的匹配方法;

(2)FCFS 算法是一种先到先服务的方法,优先为提前到达的学习任务和任务片进行匹配;

(3)KM 算法是一种求二分图最佳匹配(权值和最大的完备匹配)的普遍算法,其思想是把权值转化为可行顶标,再用匈牙利算法求出一组完备匹配,如果无法求出完备匹配,则修改可行顶标,直至找到完备匹配为止,这时的完备匹配为最佳匹配;

(4)TGOA 算法根据节点到达的顺序将其分为两个相等的组,并采取不同的策略对节点进行匹配,前部分采用贪婪算法,后部分使用匈牙利算法。

其中,只有 KM 算法是在学生所有学习任务和时间片都已知的静态场景下进行,其他算法则是在学习任务和时间片动态到达的场景下进行。

本实验将学生分为 4 种类型:夜猫子型(占比为 34.33%)、百灵鸟型(占比为 5.22%)、混合型(占比为 55.22%)、稳定型(占比为 5.22%),比较了 5 种算法面对不同类型学生时的表现效果。

本文的目标是帮助学生用户优化学习效率、增强效率等级、收获最佳学习效果,因此本实验根据分配目标自定义了 4 个评价指标,用于衡量所提方案的有效性及其适用性,分别是总效用、学习效率、效率等级和提前完成时间。另外,为了更好地呈现实验的对比效果,本文对各指标的运行结果都进行了均值化和归一化处理。

(1)总效用对比

本组实验对比了 5 种算法所提供的任务安排表在总效用指标上的性能。如图 12 所示,面对 4 种不同类型的学生,本文提出的基于迁移学习的动态分配算法 TLTA 都能取得最高的总效用。其中,对于混合型学生,TLTA 算法的总效用值比次好的 TGOA 算法和 GREEDY 算法分别高出 9.8% 和 17.92%,比 KM 算法和 FCFS 算法分别高出 52.09% 和 47.82%;对于夜猫子型,则分别提升了 10.29%,17.81%,45.2% 和 41.89%;对于百灵鸟型,分别提升了 16.07%,17.40%,77.2%,49.74%;对于稳定型,分别提升了 8%,17.82%,34.07%,26.7%,并且可以看到,随着时间片和任务数量的变化,TLTA 算法仍能保持最佳效果。FCFS 算法采用的是先到先服务的策略,分配时不考虑效用值;KM 算法则更注重匹配对数,因此这两种算法在 4 类学生中的表现都不佳。另外从图 12(d)可以看出,5 种算法在稳定型学生中的表现效果差异不太明显,这可能是由于稳定型学生的学习效率曲线相对平稳,在各个时间段下的学习效率都基本一致,因此

在不同的分配算法下具有稳定的表现。

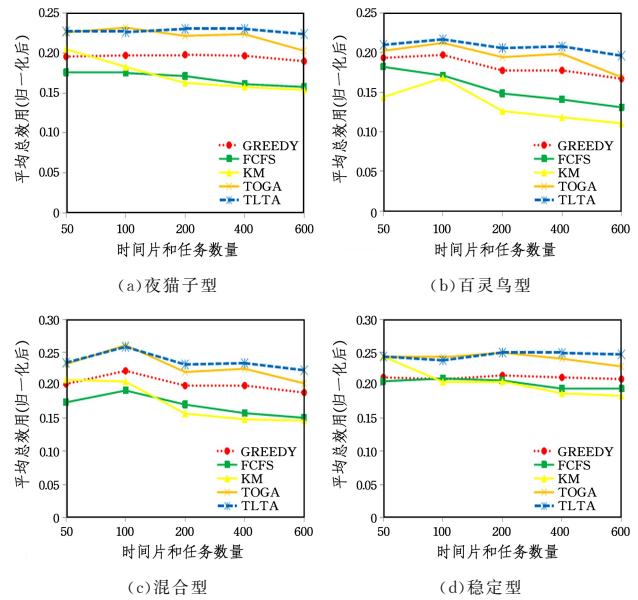


图 12 总效用对比结果

Fig. 12 Comparison results of total utility

(2)学习效率对比

本组实验对比了 5 种算法在学习效率指标上的性能。如图 13 所示,本文的 TLTA 算法随着任务量的增多表现变好。其中,对混合型学生,TLTA 算法的学习效率值比次好的 TGOA 算法和 GREEDY 算法分别高出 11.15% 和 25.11%,比 KM 算法和 FCFS 算法更是分别高出 54.65% 和 57.59%;对于夜猫子型,则分别提升了 11.22%,24.84%,46.58% 和 50.05%;对于百灵鸟型,则分别提升了 17.94%,24.86%,86.78%,61.49%;对于稳定型,分别提升了 8.14%,24.88%,31.83%,28.18%。实验结果表明,TLTA 比基线算法能收获更高的学习效率,原因是 TLTA 是根据学生的学习特征来安排学习任务,它能关注不同学习任务在不同时段的学习效率差异,因此能收获更高的学习效率。

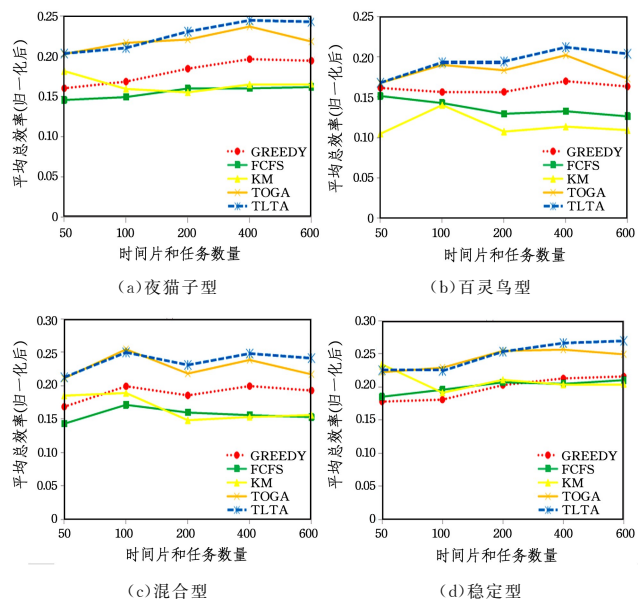


图 13 学习效率对比结果

Fig. 13 Comparison results of learning efficiency

(3)效率等级对比

本组实验对比了5种算法在效率等级指标上的性能。如图14所示,可以看出面对4种不同类型的学生,各类分配算法的结果曲线都有不同的波动,但整体来说,相比其他方法,本文TLTA算法具有较高的效率等级。其中,对于混合型学生,TLTA算法的学习效率等级值比次好的TGOA算法和GREEDY算法分别高出7.38%和12.81%,比KM算法和FCFS算法更是分别高出63.12%和66.15%;对于夜猫子型,则分别提升了12.14%,10.55%,53.7%和57.3%;对于百灵鸟型,分别提升了18.69%,22.57%,93.14%,65.23%;对于稳定型,分别提升了11.56%,41.75%,50.81%,49.9%。结果表明,在时间片和任务量较少时,各算法之间的差别不大;但随着时间片和任务量的增大,本文提出的TLTA算法的优势逐渐凸显,相反,KM算法由于更注重匹配对数,表现越来越差。以占比最高的混合型学生为例,在任务量和时间片量都为50时,与TGOA算法和GREEDY算法相比,TLTA算法最低只提升了0.65,但当任务量和时间片量增加到600时,与TGOA算法和GREEDY算法相比,TLTA算法最低提升了7.38%,与注重匹配对数的KM算法相比,学习效率等级值从提高46.04%变成提高了66.15%,表明TLTA能够适用于大任务量的情况。

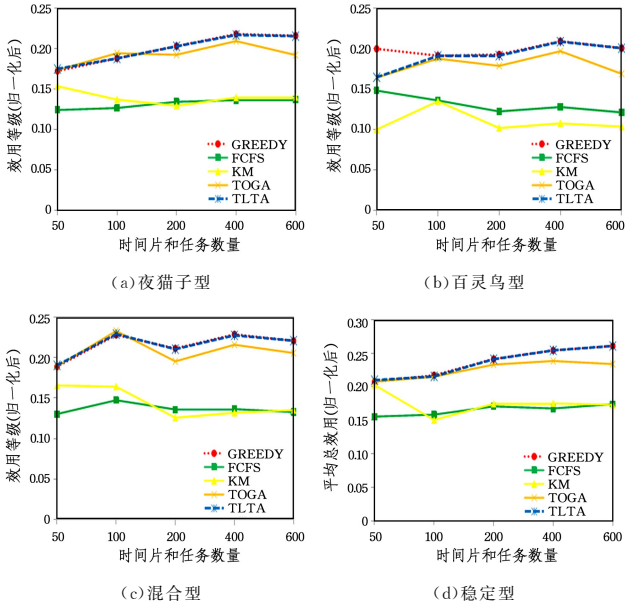


图14 效率等级对比结果

Fig. 14 Comparison results of efficiency rank

(4)提前完成时间对比

本组实验对比了5种算法在提前完成时间指标上的表现。如图15所示,对于混合型学生,TLTA的学习效率值比次好的GREEDY和TGOA分别高出0.2%和7.57%,比FCFS和KM分别高出9.03%和74.12%;对于夜猫子型,分别提升了0.27%,7.25%,8.12%,73.21%;对于百灵鸟型,分别提升了0.39%,8.02%,11.39%,78.79%;对于稳定型,则分别提升了0.42%,8.39%,11.32%,78.21%。实验结果表明,TLTA比基线算法能更早完成任务,因为TLTA在

分配任务时考虑了提前完成时间这个因素。

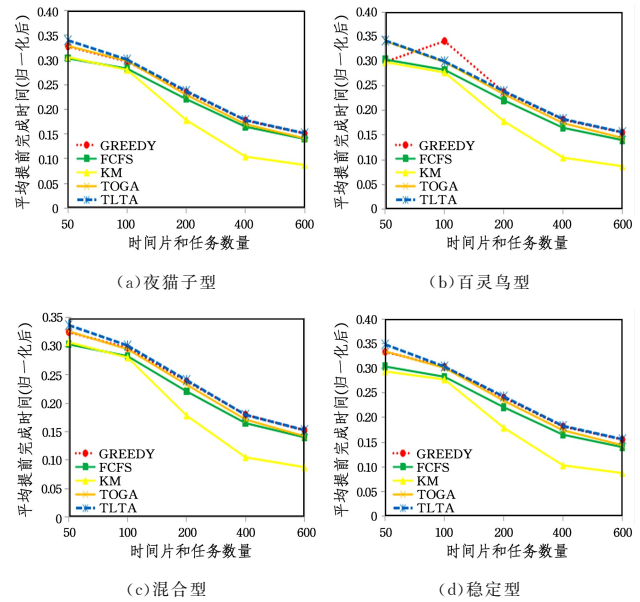


图15 提前完成时间对比结果

Fig. 15 Comparison results of completion time in advance

5.3 实验总结

由5.2节可知,相比基线算法,本文基于迁移学习的任务分配算法TLTA为不同类型的学生所提供的任务安排表在总效用、学习效率、效率等级和提前完成时间指标上都有不同程度的提高。以占比最多的混合型为例,TLTA的总效用比次好的TGOA提升了9.8%,比最差的KM提升了52.09%;学习效率比次好的TGOA提升了11.15%,比最差的FCFS提升了57.59%;效率等级比次好的TGOA算法提升了7.38%,比最差的FCFS算法提升了66.15%;提前完成时间指标比次好的FCFS算法提升了0.2%,比最差的KM提升了74.12%。表8列出了混合型学生在4个指标上的实验结果,表9则列出了TLTA相比各基线方法的提升百分比。

表8 4个指标的对比(混合型)

Table 8 Comparison of four metrics (mixed type) (单位:%)

metric baseline	GREEDY	FCFS	KM	TGOA	TLTA
total	0.19	0.15	0.15	0.20	0.22
efficiency	0.19	0.15	0.16	0.22	0.24
level	0.20	0.13	0.14	0.21	0.22
Ahead	0.15	0.14	0.09	0.14	0.15

表9 TLTA相对基线方法的提升百分比(混合型)

Table 9 Improvement percentage of TLTA relative to baseline (mixed type) (单位:%)

metric baseline	GREEDY	FCFS	KM	TGOA
total	17.92	47.82	52.09	9.80
efficiency	25.11	57.59	54.65	11.15
level	12.81	66.15	63.12	7.38
Ahead	0.20	9.03	74.12	7.57

综上所述,实验验证了本文方法的有效性和实用性,表明

其能提供一种相对高效的任務管理方案,实现预期的分配目标;提供一种个性化任务安排表,用于辅助学生合理安排学习任务,帮助学生优化学习效率,提高效率等级,达到最佳的学习效果。

结束语 本文首次提出并解决了个性化学习任务分配问题,针对学生的学习特点提供了个性化的任务分配方案,用以帮助学生优化学习效率、收获最佳学习效果。针对不同类型的任务,如何量化效率以及探明时间因素对其的影响非常具有挑战性;任务和时问片的动态变化,也是个性化分配方案的难点。因此,本文首先通过分析学生的历史学习记录来理解学生的学习特点,计算不同时段个性化的学习效率,使用二分图方法来构建动态学习任务分配场景,将学生收获的最高总效用作为分配目标;然后设计了基于迁移学习的动态分配算法 TLTA 来制定具体的任务安排表;最后,收集大量学生真实的学习数据进行实验,验证了本文方案的有效性及其适用性。本文方法可用于多领域的任务分配问题,具有较好的通用性。

本文主要考虑了一对一的匹配模式,未来将尝试对多种匹配关系(一对多、多对一、多对多)进行分配。现阶段只考虑了单个任务分配的最优情况,全局优化有待改进。另外,还将考虑完成多个任务时数量和质量的关系。考虑不同基础的学生成绩提升与效率间的关系,更准确地刻画学习效率,也是将来研究的重要方向。

参 考 文 献

- [1] GENG X, XU Y, CHEN L, et al. Learning Analytics of the Relationships among Learning Behaviors, Learning Performance, and Motivation[C]// Proceedings of IEEE ICALT 2020. IEEE, 2020;161-163.
- [2] SU Y, LIU Q W, LIU Q, et al. Exercise-enhanced sequential modeling for student performance prediction[C]// Thirty-Second AAA Conference on Artificial Intelligence(AAAI). 2018; 2435-2443.
- [3] CHEN J Y, FANG H, LIN X. Personal Learning Recommendation Based on Online Learning Behavior Analysis[J]. Computer Science, 2018, 45(S2):432-436, 462.
- [4] TIAN H, LAI S, WU F. Does Time Play a Role? Prediction of Learning Performance with Time-use Habits in Online Assignments[C]// 2019 International Joint Conference on Information, Media and Engineering (IJCIME). IEEE, 2019;473-477.
- [5] LI H J, ZHANG Z, ZHANG P W. Personalized Learning Resource Recommendation Method Based on Three-dimensional Feature Cooperative Domination[J]. Computer Science, 2019, 46(S1):471-477.
- [6] ZHANG C J. Practical research on effective time management strategy for junior high school students; a case study of junior high school students in Bao'an Experimental School of Shenzhen City, Guangdong Province[J]. Education, 2020(7):44.
- [7] WANG L, TONG Y, HU C, et al. Procrastination-aware Scheduling: Bipartite Graph Perspective[C]// 2019 IEEE 35th International Conference on Data Engineering (ICDE). IEEE, 2019; 1650-1653.
- [8] LI F. Research on the impact of fragmentation on students' online learning effect and continuous learning intention [D]. Beijing: Beijing University of Posts and Telecommunications, 2019.
- [9] SHU S Y. Research on learning efficiency analysis and association mining of curriculum knowledge points[D]. Shanghai: East China Normal University, 2016.
- [10] HU Z G, DAI S F. Using the circadian rhythm of human body to grasp the opportunity of Education[J]. Journal of Harbin University, 2001, 22(3):82-88.
- [11] BI K K. Research on the cultivation of College Students' time management ability[J]. Comparative Study on Cultural Innovation, 2018, 2(29):22-23.
- [12] LIU L. Research status and application of time management ability[J]. Education Modernization, 2018, 5(51):255-256.
- [13] HOU U L, MAMOULIS N, MOURATIDIS K. A Fair Assignment Algorithm for Multiple Preference Queries[J]. Proceedings of the Vldb Endowment, 2009, 2(1):1054-1065.
- [14] CHEN Z, SUN A. Anomaly Detection on Dynamic Bipartite Graph with Burstiness[C]// 2020 IEEE International Conference on Data Mining (ICDM). IEEE, 2020, 966-971.
- [15] TAN W J, LIU Y, CHEN R. User Allocation Approach in Dynamic Mobile Edge Computing [J]. Computer Science, 2021, 48(1):58-64.
- [16] WANG Y, TONG Y, LONG C, et al. Adaptive Dynamic Bipartite Graph Matching: A Reinforcement Learning Approach [C]// 2019 IEEE 35th International Conference on Data Engineering (ICDE). IEEE, 2019;1478-1489.
- [17] GAO X, LIU R, KAUSHIK A. Hierarchical Multi-Agent Optimization for Resource Allocation in Cloud Computing [J]. IEEE Transactions on Parallel and Distributed Systems, 2021, 32(3): 692-707.
- [18] XU Y, WEI S, WANG Y. Privacy preserving online matching on ridesharing platforms[J]. Neurocomputing, 2020, 406:371-377.
- [19] MHAISEN N, AWAD A, MOHAMED A, et al. Optimal User-Edge Assignment in Hierarchical Federated Learning based on Statistical Properties and Network Topology Constraints [J]. IEEE Transactions on Network Science and Engineering, 2021, 9(1):55-56.
- [20] TRIVEDI U. An Optimized Aho-Corasick Multi-Pattern Matching Algorithm for Fast Pattern Matching [C]// 2020 IEEE 17th India Council International Conference (INDICON). IEEE, 2020;1-5.
- [21] SHU J, LIANG C Y, XU J. Multi objective task allocation model of cloud service system based on trust [J]. Computer Research and Development, 2018, 55 (6):1167-1179.
- [22] LIU H, LIU S, ZHENG K. A reinforcement learning-based resource allocation scheme for cloud robotics [J]. IEEE Access, 2018, 6:17215-17222.
- [23] YE H, LI Y, JUANG F. Deep reinforcement learning based re-

- source allocation for V2communications[J]. *IEEE Transactions on Vehicular Technology*, 2019, 68(4): 3163-3173.
- [24] SOLEIMANPOUR-MOG H A M, NEZAMABADI-POUR H. Discrete Genetic Algorithm for Solving Task Allocation of Multi-robot Systems[C]//2020 4th Conference on Swarm Intelligence and Evolutionary Computation (CSIEC). 2020:83-100.
- [25] CHEN Q, ZHENG Z, HU C, et al. Data-driven task allocation for multi-task transfer learning on the edge[C]//2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS). IEEE, 2019: 1040-1050.
- [26] CHEN Q, ZHENG Z, HU C, et al. On-Edge Multi-Task Transfer Learning: Model and Practice With Data-Driven Task Allocation[J]. *IEEE Transactions on Parallel and Distributed Systems*, 2020, 31(6): 1357-1371.
- [27] XU Y, TONG Y, SHI Y, et al. An efficient insertion operator in dynamic ridesharing services[C]//2019 IEEE 35th International Conference on Data Engineering (ICDE). IEEE, 2019: 1022-1033.
- [28] FANG P C, YANG J J, LIAO Q M, et al. Flexible worker allocation in aircraft final assembly line using multi-objective evolutionary algorithms[J]. *IEEE Transactions on Industrial Informatics*, 2021, 17(11): 7468-7478.
- [29] TONG Y, SHE J, DING B, et al. Online mobile micro-task allocation in spatial crowdsourcing[C]//2016 IEEE 32nd International Conference on Data Engineering (ICDE). IEEE, 2016: 49-60.
- [30] SHANG L, LIU X P. Scientific Workflow Dataset Layout Based on Task Assignment and Dataset Replicas[J]. *Computer Engineering*, 2020, 46(5): 122-130, 138.
- [31] LI H, FANG B F. Emotional Robot Collaborative Task Assignment Auction Algorithm Based on Positive Group Affective Tone [J]. *Computer Science*, 2020, 47(4): 175-183.
- [32] SANG M Y Y. How to improve time utilization[M]. Beijing: Science Popularization Press, 1986: 12-20.
- [33] MERRILL A R, MERRILL R R. First things First[J]. *Executive Excellence*, 1995, 20(16): 23-44.
- [34] FREDERICK WINSLOW T. The principles of scientific management[M]. NuVision Publications, 1911: 5-10.
- [35] ALEXANDER A L. This strange life[M]. Foreign Literature Press, 1979: 113-122.
- [36] FENG Y T, TENG X F, GUO Y T. Web Violence Image Recognition Based on Transfer Learning[J]. *Journal of Chongqing Technology and Business University(Natural Science Edition)*, 2021, 38(3): 42-49.
- [37] TING H F, XIANG X. Near optimal algorithms for online maximum edge-weighted b-matching and two-sided vertex-weighted b-matching[J]. *Theoretical Computer Science*, 2015, 607: 247-256.
- [38] WEISS G. Directed FCFS infinite bipartite matching[J]. *Queueing Systems*, 2020, 96: 387-418.
- [39] GAO Z, CHEN D, SUN P, et al. KM-based efficient algorithms for optimal packet scheduling problem in cellular/information integrated networks[J]. *Ad Hoc Networks*, 2018, 77(8): 84-94.



TAN Zhen-qiong, born in 1997, post-graduate, is a student member of China Computer Federation. Her main research interests include data mining, task allocation, intelligent education and learning optimization.



JIANG Wen-jun, born in 1982, Ph. D., professor, is a senior member of China Computer Federation. Her main research interests include social network analysis, user behavior analysis and opinion mining, intelligent education and learning optimization.

(责任编辑:喻黎)