

面向数据特征的人机物融合服务分派方法^{*}

袁敏, 陈卓, 徐冰青



(南京师范大学 计算机与电子信息学院/人工智能学院, 江苏 南京 210023)

通讯作者: 袁敏, E-mail: myuan@nynu.edu.cn

摘要: 随着工业互联网的不断发展,大数据和人工智能促成了人机物全面互联.用户使用服务时产生的任务数据量正呈指数级增长,在为线上用户推荐服务满足个性化需求的同时,对于需要通过人机物交互完成的服务,如何整合线上和线下资源,并分派合适的人快速、有效地完成任务,也已成为一个挑战性问题.为了保证服务分派的准确性,提出了一种综合考虑人机物各方面数据特征的跨域融合服务分派方法,分别对用户评价的情感倾向性和业务数据的相似性进行分析,然后加入对业务执行有影响的物理世界的属性特征,以获得更合理的分派.最后,以一个互联网在线诊疗平台的医患分派为例,结果表明,文中提出的分派方法具有较高的准确性,可以获得更好的用户体验.

关键词: 跨域融合;智能服务;服务分派;用户偏好;情感倾向分析

中图法分类号: TP311

中文引用格式: 袁敏,陈卓,徐冰青.面向数据特征的人机物融合服务分派方法.软件学报,2021,32(11):3404-3422. <http://www.jos.org.cn/1000-9825/6090.htm>

英文引用格式: Yuan M, Chen Z, Xu BQ. Human-cyber-physical services dispatch approach for data characteristics. Ruan Jian Xue Bao/Journal of Software, 2021,32(11):3404-3422 (in Chinese). <http://www.jos.org.cn/1000-9825/6090.htm>

Human-cyber-physical Services Dispatch Approach for Data Characteristics

YUAN Min, CHEN Zhuo, XU Bing-Qing

(School of Computer and Electronic Information/ School of Artificial Intelligence, Nanjing Normal University, Nanjing 210023, China)

Abstract: With the continuous development of the industrial Internet, big data and artificial intelligence contribute to the comprehensive interconnection in human-cyber-physical system. The amount of task data generated by users using the service is growing exponentially. While recommending services for online users to meet personalized needs, and for services that need to be completed through human-cyber-physical interaction, it has become a challenging problem how to integrate the various offline and online resources to dispatch the right person to complete the task quickly and effectively. In order to ensure the accuracy of services dispatch, this study proposes a cross-domain collaborative service dispatch method that takes into account the data characteristics of all these factors in human-cyber-physical system. In order to get a more reasonable dispatch, the sentiment characteristics of user evaluation and the similarity of business data are analyzed respectively, and then the attributes inherent in the real world are added of which have an impact on business processes. Finally, taking the doctor-patient assignment of an online diagnosis and treatment platform on the Internet as an example, the results show that the method proposed in this study has high accuracy and can improve the efficiency of task execution.

Key words: cross-domain integration; intelligent service; service dispatch; user preference; sentiment analysis

工业互联网的快速发展,深度支撑并改造了人类用户的各种社会关系.人的作用和影响不断增强,形成一种人机物一体融合的综合发展态势^[1].软件正在成为未来社会运行的基础设施,其应用已经成为世界经济社会发

* 基金项目: 国家自然科学基金(41771411); 江苏省教育科学“十三五”规划(C-b/2016/01/24); 江苏省研究生科研与实践创新计划(SJCX19_0201)

Foundation item: National Natural Science Foundation of China (41771411); “13th Five-Year Plan” of Education Science in Jiangsu Province (C-b/2016/01/24); Postgraduate Research & Practice Innovation Program of Jiangsu Province (SJCX19_0201)

收稿时间: 2020-01-02; 修改时间: 2020-03-09; 采用时间: 2020-05-18

展的热点,驱动着信息空间与物理世界的深度融合^[2]。人机物融合交互产生的业务数据变得越来越复杂多样,服务成为了桥接这些数据的连接者。在为线上用户推荐服务满足个性化需求的同时,对于需要通过人机物交互完成的服务来说,如何整合线上和线下资源、分派合适的人快速有效地完成任务,也已成为一个挑战性问题。服务推荐与分派常常共存于同一个业务场景,例如,互联网在线服务平台一般都包含服务推荐和分派两个阶段。用户在平台提交需求后,系统大都会结合平台历史数据向用户推荐评分较高服务方案,但往往评分高低并不能真实反映以人为本的价值观,即人与信息系统中对应的角色存在不一致性。为了提高业务的执行效率,信息系统一般只在服务推荐阶段考虑用户需求,在服务分派阶段则减少用户干预,用户往往处于被动接受状态,因而常常会出现诊疗平台推荐的治疗方案虽是合理的,但分派的医生不是患者喜欢的;网约车平台推荐的出行方案虽是满意的,但分派的司机没有考虑乘客的意愿,甚至是存在危险的情况。

人机物融合环境下的人是信息化社会的主要载体,其在自然界与人类社会的能力越来越强,对信息技术的依赖也越来越大。机主要是指软件的基础平台支撑,计算机开始从以工具为主要形态演变为与机器形态无关的平台概念,然后上升为计算思维的概念,变为人类除理论与实验之外,认识世界的第三种手段。物主要是指面向物理世界的信息化空间,是物理世界的信息化体现,作为人类生存和信息化平台的物质支撑,物的因素也至关重要。纵观人机物研究的发展,通常以计算机为载体,人或物为主体,将信息空间作为承载平台,而忽略了信息空间与人和物的一致性。各种信息系统主要优化目标集中在提升信息空间的容量和安全性,而人的属性和物理世界状态复杂多变,人在信息空间的映射往往是不全面的,简单的映射不能处理物理世界的不稳定因素。

近年来,许多研究通过协同过滤算法进行服务推荐,以缓解信息过载问题。推荐算法通过建立用户与项目之间的二元关系,利用相似性关系挖掘每个用户潜在感兴趣的对象^[3]。典型的协同过滤算法有两类:基于用户和基于项目^[4]。尽管这些研究取得了很好的效果,但是随着用户和服务数量的不断增加,协同过滤算法在为用户推荐服务项目时,一方面为用户推荐的服务质量可能与真实情况存在偏差;另一方面,用户和业务员参与的服务实行过程中,可能存在人为因素的影响。根据上述分析,当前服务推荐存在的问题归纳如下。

- (1) 用户评分不能完全反映用户的情感倾向。根据用户对项目的评分进行排序后产生推荐结果,但是用户的打分不能准确体现用户的真实情感倾向^[5];另外,用户打分中可能包含不真实的分数,使项目的预测分与真实的体验相差较大。
- (2) 项目与用户之间的关联度评价不够准确。由于业务员之间的个体差异,不同的业务员提供的服务也具有差别,用户提出的需求同样也存在着差异。因此,用户需求与项目特征的匹配程度,决定了任务能否较好地完成。传统服务推荐方法中,通过对相似用户或者相似项目关联分析来预测目标用户和目标项目的关联度,然而相似用户和相似项目的关联具有偶然性,不能准确表现用户需求和项目特征的匹配程度。另外,新加入的用户因缺少历史记录而存在冷启动的问题^[6],不能为新用户准确推荐项目。
- (3) 对服务执行过程中现实属性的影响考虑不足。传统推荐方法缺少对服务项目执行过程的考虑,也就是没有考虑物理世界的各种属性对任务执行的影响。例如,由于自身等级和所在平台不同,业务员所能调用的资源也就不同,这些属性会对服务效率和结果造成一定的影响。

为了解决上述问题,本文提出了一种综合考虑人机物各方面数据特征的跨域协作服务分派方法,在为服务请求者或消费者分配合适的服务资源时,可以把服务分派场景看作是服务推荐的一种特殊情况,它更强调用户与业务员在服务执行过程中的参与,着重考虑人为因素对任务完成情况产生的影响。首先考虑人的认知空间属性特征,将用户的情感倾向性分析和评分结合;其次挖掘信息空间的属性特征,对平台上的历史数据和用户提交的任务特征进行关联性分析;最后分析物理世界的属性特征对服务分派合理性的影响,构建映射评分机制。总体来说,本文有以下4点贡献。

- (1) 针对问题(1),引入评价的情感倾向性分析。相对于历史评分数据,评论是用户能够更具体表达自己喜好的一种方式,能够更准确地表现出用户对于项目的真实情感倾向^[7]。因此,本文以评论为对象,将评论转化为情感类别,并映射为情感得分,与评级得分加权计和后,作为用户对服务项目的情感倾向分值,与直接采用评级评分的方法相比会更加合理。

- (2) 针对问题(2),直接对用户需求和服务项目特征进行分析.基于内容的推荐方法能很好地应对用户冷启动的问题^[8],本文在这种方法的基础上,从平台的历史数据对项目提供者的历史任务进行抽取生成摘要文件,可以更简要地表现项目的特征,以此为基础,直接对用户需求和项目特征进行匹配分析.相较于通过寻找相似用户和相似项目的间接预测,本文方法预测用户需求和项目特征的匹配度更加准确.
- (3) 针对问题(3),建立对物理世界属性特征的评判方法.对影响到服务项目执行的物理世界客观属性建立分级映射得分机制,通过层次分析法确定各属性权重,并与满意度值加权求和,作为服务项目的综合得分,以综合得分为依据对项目排序,可以有效地解决新用户的冷启动问题,并提升服务执行过程中的用户体验.
- (4) 在所提出方法的基础上,利用收集到的在线诊疗平台数据集完成相关实验.实验验证了从人机物这 3 个视角,综合考虑服务匹配度(matching degree)、服务满意度(satisfaction degree)以及服务能力成熟度(maturity level),能够提高服务分派的合理性以及在人机物融合的环境下人为因素在服务执行过程中对任务完成情况产生的影响.实验结果表明:与其他服务推荐方法比较,本文方法具有较高的准确性,可以获得更好的用户体验.

本文第 1 节对人机物融合和服务分派的相关工作进行综述,分析目前此类研究中的现状和不足.第 2 节介绍人机物融合环境下的服务分派的概念视图和优化目标.第 3 节从认知空间、信息空间和现实空间的属性特征这 3 个不同的角度分别阐述面向数据特征的多域协作的具体的服务分派方法.第 4 节以病患分诊场景为例,展示相关的数据并对本文方法的各个步骤进行详细举例说明.第 5 节给出实验评估,设计完整的实验方案,进行实验结果对比与分析.第 6 节对本文的工作进行总结,分析目前的不足和未来可能展开的工作.

1 相关工作

1.1 人机物融合发展过程

- (1) 计算机的出现和发展,使纸媒体向信息化转变.

计算机出现之前,人类社会活动通常以纸为载体.信息技术的发展催生了信息化时代的产物,如办公自动化(office automation,简称 OA)和企业资源企划(enterprise resource planning,简称 ERP).OA 通常指使用集成在办公环境中的计算机和通信系统来支持管理资源.最早的办公自动化从 WPS 和 MSOffice 等单机版的办公应用软件开始,实现了从手工办公到计算机办公的转变,从而形成了 OA 的概念.随后,以工作流为中心,在文件型 OA 的基础上增加了公文流转、流程审批、文档管理、会议管理、资产管理等实用功能.工作流技术在办公自动化中被广泛应用,对于工作流的非形式化语言定义有很多,虽然各有不同,但基本都认为工作流是经营过程的一个计算机实现^[9-11].文献[12]阐述了工作流实施的 3 个基本步骤:映射、建模和管理.文献[13]对工作流进行了形式化描述,提出一个三维工作流模型,它包含 3 个子模型:组织模型、数据模型和过程模型,分别从不同的侧面描述工作流的各种属性.随着企业和用户需求的不断提高,研究人员意识到,办公实际上是一个管理过程.电子商务时代带来的企业事务处理对象瞬息万变,这就要求作为企业的办公自动化系统能够提供足够的灵活应变和开放交互能力.20 世纪 90 年代,美国 Gartner Group 公司提出了 ERP 的概念,ERP 是针对物资资源管理、人力资源管理、财务资源管理和信息资源管理集成一体化的企业管理软件.Telumbde 提出了一个基于名义组技术和层次分析法(analytical hierarchy process,简称 AHP)的流程框架来评估 ERP 项目^[14].Colmenares 提出了一个基于 AHP 的决策支持系统原型,以协助企业资源规划评估进程,主要为中小型企业服务^[15].Wortmann 等人将自然语言处理(NLP)和执行信息系统(executive information system,简称 EIS)数据库结合,利用企业系统数据库中大量有价值的文本来丰富公司活动和流程,利用 EIS 数据库的结构或架构中固有的高质量附加信息来改进其性能^[16].信息化系统对提高工作效率和生活质量有着不可忽视的作用,但其主要是将人类社会和物理世界映射到信息空间,通过信息化实现高效管理和精准决策.

- (2) 人机交互让信息化系统中的人与机更好地交换信息.

人机交互是指人与计算机之间使用某种对话语言,以一定的交互方式完成确定任务的人与计算机之间的

信息交换的过程,主要包括人到计算机和计算机到人的信息交换两部分.文献[17]中认为,普适计算使得信息空间与人们生活的物理世界相融合.与此相应,人机接口也将随之扩展到人们生活工作的整个三维物理世界.在普适计算的模式下,用户和环境的物理因素将成为交互的重要因素.计算机视觉的发展,促进了人机交互的发展.文献[18]将眼动跟踪方法应用于面向大屏幕的目标选择,进而避免鼠标操作方式在空间上的远距离移动;同时提出一种融合眼动跟踪与手势的多通道交互方法,通过眼动跟踪选择目标,利用手势进行选择确认.文献[19]设计出一种增强显示模式,通过突出的视觉减少飞行员的文字认知负载,驾驶人员能够迅速获得周围环境和载具状态等信息,这种增强人机交互模式有效增强了驾驶员的应急反应.

(3) 物联网实现了计算、通信与物理系统的一体化设计.

信息物理系统(cyber-physical systems,简称 CPS)是一个综合计算、网络和物理环境的多维复杂系统,它建立在环境感知的基础上,融合了计算进程和物理进程,可以感知以及改变物理世界,大部分研究工作都围绕如何对信息物理空间进行建模与分析.Walton 等人利用偶图的反应系统对环境中的物理及网络进行建模,刻画了室内物体和主体在空间中的移动,分析可能出现的各种情况,并对 CPS 的动态演化进行描述^[20].Tan 等人提出了 Lattice-based 事件模型,它由事件类型、外部和内部属性组成,再通过这些属性定义一组事件组合规则,用于指导面向事件的 CPS 软件建模^[21].Jensen 等人提出了基于模型的 CPS 设计方法,对 CPS 模型进行了全面的分析^[22].Tsigkanos 等人对自适应系统进行建模分析,通过识别可能由拓扑更改引起的违反安全要求的行为,并选择一组可防止此类违反的安全控制措施来设计可感知拓扑的自适应安全系统,主要考虑的是用 CPS 对物理拓扑进行建模分析,维持运行时拓扑的实时表现,表示在物理结构发生变化时的更新.该方法说明:CPS 适合于对一个封闭的环境进行分析,考虑当环境发生变化时可能引起的其他的变化,达到及时更新^[23].信息物理系统轻视了人对信息空间和物理世界的影响,使其面临情景感知和安全问题的巨大挑战.

综上所述,人们生活中经历了从纸媒体到信息化时代的变革,再到人机交互与信息物理系统的发展,这些变革都是人机物三元空间部分融合的例证.因此,人机物的高度融合是未来发展的主要趋势,以达成机从计算能力到计算思维的转化、物从个体空间向人类社会和自然社会的深入、人从固定角色向社会人和自然人的转变.人机物在融合环境下对彼此的作用和依赖越来越大,更加密不可分.

1.2 服务推荐与分派

在 20 世纪 90 年代中期,推荐系统才被作为一个独立的研究领域.推荐系统的起源可以追溯到其他的领域,例如认知科学、近似性理论、信息检索、预测理论、管理科学、市场营销建模等.文献[24]提出了一种众包中的面向人工智能任务(human intelligence task,简称 HIT)的推荐方法,向用户提供最合适的 HIT.因为在众包平台上,请求者可以发布各种项目让工人完成.在由一系列 HIT 组成的项目中,只有通过信任评估的工人才能申请 HIT.这种方法是基于工人与请求者交易的过去性能来计算请求者和工人之间的直接信任(DT).然后,基于信任强度(SOT)提出一种新的信任模型,该模型测量两个未与任何普通员工交易的请求者之间的信任.这种方法是从用户的角度来思考服务推荐方法.也有一些工作从用户和项目两方面来考虑,推荐的性能和准确率常常高于常规的基于用户或项目的方法.文献[25]中将用户和云服务提供者之间的交易关系抽象为斯塔克尔伯格模型,定义为一种价格领导的竞争关系,用户和服务提供者是竞争的关系,用户会根据提供者的价格和评级来决定自己的需求量,服务提供者也会根据用户的评级来决定自己的价格和对应的容量.从博弈的角度,还详细分析了若服务请求者想要获得优质服务,而服务提供者想要卖出自己的服务时两者的收益、优质服务和劣质服务的成本以及劣质服务所面临的惩罚所处的状态.在为服务请求者或消费者分配合适的服务资源时,这些服务项目的执行常常需要人的参与,需要通过人机物交互完成,分派合适的人快速有效地完成任务.服务推荐一般聚焦如何将服务推送给潜在的用户,常常忽视了接受推荐用户的真实体验和任务执行过程,即忽略了信息空间中构造的角色与对应的人和物之间的一致性,例如用户对接单的司机或问诊医生的真实评价等.

2 服务分派中的人机物三元空间

服务分派过程如同一个顺时针循环的数据加工流程,如图 1 所示.首先,用户提交若干任务至平台,平台将这

些任务分派给各机构中不同部门的业务员,业务员执行任务并提供服务.在考虑某项任务的合适人选时,本文从人机物融合的不同视角出发.

- (1) 人类社会作为信息化社会的主要载体,文中关注认知空间的特征属性,例如用户的情感倾向,表达用户对服务的满意度.
- (2) 信息空间是指软件基础平台和数据,这里表现为信息空间的特征属性,例如,平台在分派任务时考虑用户提交任务与业务员服务内容的匹配度.
- (3) 物理世界是指客观世界中固有的信息化属性,文中考虑现实空间的特征属性,例如,反映业务员服务能力成熟度的相关属性,如所属机构位置、等级、业务员职称及专长等.

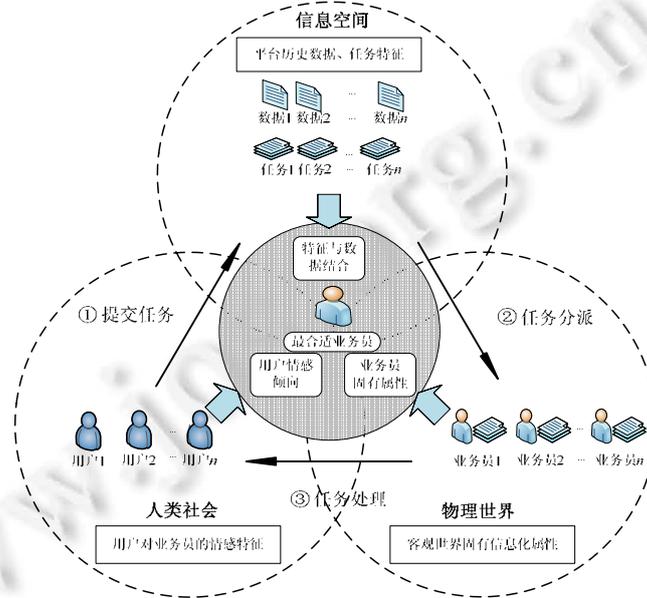


Fig.1 A conceptual view of the services dispatch in human-cyber-physical system

图1 人机物融合环境中服务分派的概念视图

在用户情感倾向性分析中引入用户评价和评分,常见的评分表现形式为五级打分制,用户评价方式存在多元化,通常以客服电话录音的语音信息、评论文本和图形符号等为载体,其中,以文字进行评价的平台占大多数.各种形式的评价相对于量化的评分,可以在更细微的尺度上体现用户的情感,且都可以通过情感分析进行辨别.在用户提交的任务与业务员的匹配度分析中引入业务员的历史任务数据,由业务员历史任务生成概要描述文件,并与任务特征进行相似度计算.在分析物理世界中可能影响任务执行的特征属性时,本文引入业务员的能力等级和所在机构的等级两个较为普遍的特征属性.文中用到的符号及其含义见表1.

服务分派问题的优化目标可以概括为公式(1)和公式(2).

$$\left. \begin{aligned} \text{Max} : & \text{Mean} \left(\sum_i^{|U|} su_i \right) \\ \text{s.t.} : & 0 \leq su_i \leq 5, i \in [1, |U|] \end{aligned} \right\} \quad (1)$$

$$\left. \begin{aligned} \text{Max} : & \text{Mean} \left(\sum_i^{|U|} \text{similar}(v_i^{\wedge}, v_i) \right) \\ \text{s.t.} : & 0 \leq \text{similar}(v_i^{\wedge}, v_i) \leq 1, i \in [1, |U|] \end{aligned} \right\} \quad (2)$$

其中, su_i 表示用户对业务员 v_i 的综合满意度评分,数值范围在 0~5 之间; $\text{similar}(v_i^{\wedge}, v_i)$ 表示分配的业务员 v_i 与平台中最理想的业务员 v_i^{\wedge} 之间的相似度,该值越大,差异越小,数值范围在 0~1 之间.

Table 1 List of notations

表 1 注释表

V	业务员集合
v_{ij}	待分配给用户 $U[i]$ 的业务员集合 v_i 中第 j 个业务员
d	某分支部门代码
$branch_j$	业务员 $V[j]$ 的所属部门代码
$profile_j$	业务员 $V[j]$ 的概要描述文件
$institution_j$	业务员 $V[j]$ 所在的机构等级
si_j	业务员 $V[j]$ 所在的机构等级得分
$level_j$	业务员 $V[j]$ 自身的等级
sl_j	业务员 $V[j]$ 的自身的等级得分
$assess_j$	其他用户对业务员 $V[j]$ 的历史评价
sa_j	对业务员 $V[j]$ 评价的情感得分
ri	其他用户对业务员 $V[j]$ 评级
sr_j	其他用户对业务员 $V[j]$ 评级得分
su_j	其他用户对业务员 $V[j]$ 的综合满意度得分
U	用户集合 $U, U[i]$ 是第 i 个用户
$depict_i$	用户 $U[i]$ 提交的任务描述
s_{ij}	$depict_i$ 和 $profile_j$ 的相似度
$L[i][j]$	业务员 j 对于用户 i 的最终得分

在服务分派的过程中,为了提高用户的体验,一方面希望执行任务的业务员的综合评分越高越好;另一方面,对于任务描述而言,要求由最擅长该类任务的业务员来完成,即业务员的专长与任务特征相似度越大越好。

3 面向数据特征的跨域融合服务分派

3.1 认知空间的服务用户满意度

本节从人类社会角度出发,关注认知空间的属性特征,着重分析用户对于业务员服务的满意度.常规评分中的分数是用户给出的评级,为了更全面地表达用户对此次服务的满意度,文中通过挖掘用户评价信息,对服务评价进行情感倾向性分析,并与评级分数结合,实现用户情感倾向预测.因此采用情感分析方法计算评价得分,其中,词典规则的方法是当前情感倾向性分析中较常用的方法.Cheng 等人提出了一种基于规则的中文微博文本的情感分类方法,进行分析微博的情感倾向性^[26].Gao 等人在研究中文微博文本中,根据其具有的特点,提出在语义规则下分析微博热点^[27].Park 等人基于种子词词典和 3 个在线词典,选出所有的同义词存储进同义词词典,以此来提高同义词词典的可靠性;同时,递归地收集同义词和反义词来扩展同义词词典.在不使用人力资源的情况下,有效地扩展了词典,增强了词典的可用性,提高了情感分类的准确性^[28].Rao 等人使用三剪枝策略建立了对于社会情绪检测的准确的情感词典,该方法生成的词典是与语言无关的、细粒度的^[29].从技术角度来看,情感词典的构建还可以分为以下 4 种方法:基于语义和规则的方法^[30,31]、基于图传播的方法、基于词对齐模型的方法以及基于表示学习的方法^[32,33].目前的语音情感分析和图形符号的情感分析也取得了较好的成果,以目前比较成熟的文本情感倾向分析为例,分为以下几个关键步骤.

- (1) 分词:对文本评价进行分词处理.分词就是将连续的字序列按照一定的规范重新组合成词序列的过程.因为循环神经网络训练无法直接处理整段的文字,所以需要进行分词的处理.在分词之后,去掉一些停用词,防止停用词影响情感倾向分类的准确度.对于 $assess=\{assess_1, assess_2, \dots, assess_n\}$ 这样的句子数组,以其中的一个句子分词为例,得到表达式(3).

$$assess_i = w_i = \{w_{i1}, w_{i2}, \dots, w_{ij}\} \quad (3)$$

其中, w_i 表示对每个句子的分词后的词数组, w_{ij} 表示其中的一个分词.

- (2) 词向量化.经过文本分词后,文本评论数据已经变成切分好的词片段,这些词片段需要经过向量化后,才能放入神经网络中训练.词向量化提供了一种数学化的方法,把自然语言这种符号信息转化为向量形式的数字信息.这样,就把自然语言理解的问题转化为机器学习的问题.本文通过 word2vec 方法进

行词量化,一个句子经过词向量化后,表示为公式(4).

$$assess_i = \vec{w}_i = \{\vec{w}_{i1}, \vec{w}_{i2}, \dots, \vec{w}_{ik}\} \quad (4)$$

其中, \vec{w}_i 是对 w_i 向量化的后得到的, \vec{w}_{ij} 对应 w_{ij} 的向量化.

- (3) 分数映射.对每一条评论,SnowNLP 的输出为该文本的情感倾向为正面的概率.经过多次实验后,将概率与得分建立出映射表,见表 2.其中, p 为概率区间, $p \in [0, 1]$, $1 \leq sa \leq 5$.采用直接映射的方法,经过统计划分出 $1/m$ 个概率区间,以确保输出落在每个区间内的机率相等.

Table 2 Probability and score correspondence table

表 2 概率与分数对应表

正面概率 p	p_1	p_2	...	p_m
评价得分 sa	sa_1	sa_2	...	sa_m

一般来说,大部分评级系统共有 5 级,粒度通常较大.表 3 所示的是一种常见的评级系统,在此列出其对应的评级得分,计算出用户对业务员的评价对应得分后,与用户对业务员的评级打分按公式(5)加权计算出综合得分,作为用户对业务员的满意度预测值,如公式(5)所示.

$$su_i = \alpha \cdot sa_i + (1 - \alpha) \cdot sr_i \quad (5)$$

Table 3 Rating and score correspondence table

表 3 评级与分数对应表

用户评级 r	很满意	满意	一般	不满意	很不满意
评级得分 sr	5	4	3	2	1

算法 1 展示了计算用户对业务员的综合满意度的伪代码.

算法 1. 用户满意度算法.

输入:评论数组 $assess$,权值 α ,业务员集合 V .

输出:综合满意度列表 su .

1. **Begin**
2. **For** $i \leftarrow 1$ to $V.length$
3. **For** $j \leftarrow 1$ to $assess_i.length$ **do**
4. $w_{ij} \leftarrow$ 对 $assess_{ij}$ 进行分词; //切分词
5. $\vec{w}_{ij} \leftarrow$ 向量化 w_{ij} ; //词向量化
6. $p_{ij} \leftarrow \vec{w}_{ij}$; //词向量为正面的概率
7. $sa_{ij} \leftarrow p_{ij}$; //概率映射为评价得分
8. $sr_{ij} \leftarrow r_{ij}$; //评级映射为评级得分
9. **End for**
10. **End for**
11. **For** $i \leftarrow 1$ to $V.length$ **do**
12. $sa_i \leftarrow mean(\sum_j^{|assess_i|} sa_{ij})$;
13. $sr_i \leftarrow mean(\sum_j^{|assess_i|} sr_{ij})$;
14. $su_i = \alpha \cdot sa_i + (1 - \alpha) \cdot sr_i$;
15. **End for**
16. **End**

3.2 信息空间的服务内容匹配度

从信息空间的角度出发,关注任务与服务内容的匹配度,着重分析任务与业务员服务内容的匹配度.对于所

有用户提交的任务请求,本文的目标是,选择最适合执行该任务的业务员.通过引入业务员概要描述文件,并与任务特征进行相似度计算,以推荐更适合执行相应任务的业务员.

文献[8]中通过获取对项目的描述来建立概要描述文件,本文则通过获取项目的历史纪录生成概要描述文件.概要描述文件是对业务员执行过的历史任务的概述,表示了该业务员处理过哪些任务以及经常处理哪些任务,是业务员业务能力的概述.在获取到服务分派原始数据后,根据历史订单生成每个业务员的概要描述文件.通过对任务和概要描述文件进行相关性分析,可以判断任务和服务的匹配度.匹配度越高,说明该业务员提供的服务越契合提出该任务的用户,更容易满足用户需求.

此外,本文通过分类来缩小任务可分配的人员范围.随着统计学习方法的迅速发展、互联网在线文本数量的增长和机器学习学科的兴起,一套解决大规模分类问题的经典方法逐渐形成.这个阶段的主要工作是人工特征工程加浅层分类模型,整个分类问题就拆分成了特征工程和分类器两部分.

- 特征工程主要有手工挑选和机器学习等方法.手工挑选包括数据预处理、特征量化和特征筛选这3步:数据预处理需要根据现实情况进行,可能会涉及缺失值处理、标点符号、停用词和中文分词等等;特征量化的主要思想是在特征空间中将不同特征作为坐标轴对文件进行向量化,将向量化后的文件作为分类器的输入;而特征筛选就是选择适合作为坐标轴的特征,一般是区分度明显的特征,常用的有文档频率、信息增益等.
- 分类器的原理有很多种,包含朴素贝叶斯^[34]、SVM^[35]和KNN^[36]等.随着ELMo^[37]、BERT^[38]等模型的发布,一般都采用大规模语料上进行无监督预训练的深层模型,通过在下游任务数据上进行微调,就可达到很好的效果.

本文实验采用了百度的开源深度学习平台 PaddlePaddle 开发的预训练模型管理工具 PaddleHub,可以借助预训练模型更便捷地开展迁移学习工作.这里选择百度发布的知识增强的预训练模型 ERNIE^[39],相较于 BERT 学习原始语言信号,ERNIE 直接对先验语义知识单元进行建模,增强了模型语义表示能力.

如算法 2 所示,根据任务描述进行分类,在适合的范围内筛选,再计算该任务与分支内的所有业务员概要描述文件的相似度,根据相似度排序,然后选择 k 位业务员作为备选,为下一步算法的输入.

算法 2. 内容匹配度排序算法.

输入:用户集合 U ,业务员集合 V ,任务描述 $depict$,分支表 $branch$.

输出:相似度最高的 k 位业务员集合.

1. **Begin**
2. **For** $i \leftarrow 1$ to $V.length$ **do**
3. $profile_i \leftarrow V[i].depict$; //根据历史任务生成配置文件
4. **End for**
5. **For** $i \leftarrow 1$ to $U.length$ **do**
6. $v_i \leftarrow \emptyset$; //备选集初始为空
7. $d \leftarrow classify(depict_i)$; //根据任务描述分类
8. **For** $j \leftarrow 1$ to $V.length$ **do**
9. **if** $branch_j = d$
10. $v_i \leftarrow v_i + V[j]$; //该分支所有业务员加入备选集
11. $s_{ij} \leftarrow similar(depict_i, profile_j)$; //相似度计算
12. **End if**
13. **End for**
14. $Sort(v_i, s_i)$; //按相似度降序排序
15. $v_i \leftarrow Top(v_i, k)$; //留下相似度最高的 k 位业务员
16. **End for**
17. **End**

数据量过多时,相似度计算量很大,因此在调用相似度计算中加入了差错处理,并考虑服务器每秒的响应请求数,也即是最大吞吐能力,在每条相似度计算完成后,间隔 0.5s 再输入数据,满足大部分服务器的 QPS(query per second)每秒查询率限制.

3.3 现实空间的服务能力成熟度

从物理世界的角度出发,关注现实空间的属性特征,分析业务员的服务成熟度.通过分析固有属性特征对于服务执行的影响,为其建立分级映射机制,以推荐服务成熟度更高的业务员.例如,业务员的等级和其所处机构的等级是用户自选业务员时会考虑到的两个特征,等级越高,对于任务的执行可能越有利.故将特征等级与分数通过公式(6)建立映射,使各属性等级对应的分数都处于[1,5]之间,与用户评级等其他方面的分数保持区间一致:

$$score_j = 5 \cdot \frac{f_j}{feature} \quad (6)$$

其中, $feature$ 表示特征的级别总数, f_j 表示业务员 j 关于该特征的等级, $score_j$ 表示业务员 j 在该特征上的得分.由于数据来源的限制,这里只考虑机构等级和业务员自身等级两个相关特征.

将属性特征得分与用户满意度得分结合,选取最适合执行该项任务的业务员.如算法 3 所示,在算法 2 中得出的备选业务员集合中,根据业务员等级计算等级得分,由机构等级计算机构得分,再与用户评价得分等各项得分加权相加得出最终得分,选出得分最高的业务员.

算法 3. 最终得分算法.

输入:用户满意度算法计算出的评分表 su , 内容匹配度排序算法得到的备选业务员集合 v .

输出:最终被分配的业务员集合 v .

1. **Begin**
2. **For** $i \leftarrow 1$ to $U.length$
3. **For** $j \leftarrow 1$ to $v_i.length$ **do**
4. $si_j \leftarrow institution_j$;
5. $sl_j \leftarrow level_j$;
6. ... //依次计算各项得分
7. $L[i][j] \leftarrow \alpha \cdot su_j + \beta \cdot si_j + \gamma \cdot sl_j + \dots$; //计算最终得分
8. **End for**
9. $Sort(v_i, L[i][j])$; //按最终分降序排列
10. $v_i \leftarrow Top(v_i, 1)$; //只留下最终分最高的 1 位业务员
11. **End for**
12. **return** v ;
13. **End**

对于权重的设置,本文采用层次分析法.层次分析的思路是:将复杂的系统分解成若干个层次,其层次系统的形状与树形结构相似,同一层次的元素作为标准,对下一层的某些元素起支配作用,同时又受上一层次的元素支配.通过两两比较的方式确定层次中各指标的重要度排序,然后进行综合判断.依据这种模型,将第 1 层目标确定为业务员最终得分,第 2 层准则层确定为属性选取的判断方法,第 3 层指标层确定为业务员所在机构的等级、业务员等级和用户评价分等多个特征属性.即通过层次分析法从不同判断准则的角度选取重要指标,每一个准则对应的指标是相同的.可从 3 个角度来判断各指标的重要性:指标在衡量业务员最终得分中发挥的作用、指标的计算依据以及涉及因素、获取难易程度或者获取的精度.根据构成最终得分的属性数量及各属性的重要程度建立判断基准表.

每个属性在业务员的最终评分中反映的是不同的侧面,发挥的作用也是不同的.准则层的目标是判断各属性在评价业务员最终得分中可发挥的作用,通过定性分析法比较各属性的作用,从而得到判断矩阵.分析研究上

述 n 个指标对业务员最终评价的贡献程度后依次排序,根据表 4 建立判断矩阵 M ,由公式(7)计算矩阵 M 各行元素之积的 n 次方根,最后通过公式(8)计算得出的各属性的权重:

$$A_i = \sqrt[n]{\prod_{j=1}^n M[i][j]} \quad (7)$$

$$weight_i = \frac{A_i}{\sum_{j=1}^n A_j} \quad (8)$$

业务员最终得分的计算如算法 3 所示,计算出每个用户提交的任务对应的备选业务员最终得分后,选择分值最高的业务员作为最终结果.

Table 4 Indicators for analytic hierarchy process

表 4 层次分析的指标

基准值	含义
n	因素 a 比 b 的重要等级高 $n-1$ 级
$n-1$	因素 a 比 b 的重要等级高 $n-2$ 级
...	...
1	因素 a 和 b 的重要等级相同
...	...
$1/(n-1)$	因素 a 比 b 的重要等级低 $n-2$ 级
$1/n$	因素 a 比 b 的重要等级低 $n-1$ 级

4 算法举例

下面以病患分诊场景为例,给出具体的服务分派流程.目前,分诊平台在用户选择科室后随机分配医生,将患者平均分给各个医生,或者患者凭自己的判断选择医生.本文从人机物融合的角度出发,考虑各方面的数据特征,帮助患者选择最合适的医生.文中将医生所属医院名称匿名化,医院名称中第 2 个、第 3 个字用星号隐去,这里给出部分医生的数据(见表 5).

Table 5 Data of some doctors

表 5 部分医生数据

医生 ID	医生所属医院	医生等级	医生职称	医生专长
3088	中***医院	三级甲等	主治医师	耳鼻咽喉科常见疾病
7857	江***属医院	三级甲等	主任医师	喉癌 CO ₂ 激光微创手术,喉乳头瘤精细化治疗等
3136	广***自治区人民医院	三级甲等	副主任医师	支气管炎,支气管扩张症,各种肺炎等, 呼吸系统感染性疾病诊疗等
3086	广***自治区 人民医院	三级甲等	副主任医师	擅长耳内镜,微创耳外科,耳鼻喉科, 侧颅底外科,头颈肿瘤,外周性眩晕等领域
7898	南***大学珠江医院	三级甲等	主任医师	擅长慢性或脓性中耳炎及胆脂瘤导致的鼓膜穿孔等
3119	江***民医院	三级甲等	副主任医师	耳鼻咽喉科常见疾病
4039	河***医院	三级甲等	主任医师	中医体质辩证治疗过敏性鼻炎,小儿腺体肥大, 眩晕等耳鼻喉科疾病有独特经验
7855	江***民医院	三级甲等	主治医师	咽喉疾病和头颈肿瘤的诊治,主要包括慢性咽炎...
1495	烟***顶医院	三级甲等	副主任医师	擅长各类美容整形手术及皮肤溃疡,疤痕, 皮肤恶性肿瘤的综合治疗
7859	复***附属眼耳鼻喉科医院	三级甲等	副主任医师	擅长延后肿瘤及咽喉嗓音疾病的治疗
4052	深***岗区耳鼻喉医院	三级甲等	主任医师	鼻科(窦外科,鼻眼相关外科和鼻颅底相关外科)腔 良性和恶性肿瘤疾病的诊治
3091	广***大学附属第二医院	三级甲等	副主任医师	擅长喉癌,面部皮肤癌等多种头颈部恶性肿瘤, 放射治疗,靶向治疗,药物治疗等...
3018	遵***院附属医院	三级甲等	副主任医师	先天性耳畸

依据文中的服务分派方法,病患分诊过程包括 3 个步骤:用户满意度计算、匹配度计算和最终得分计算.

4.1 服务满意度计算

提取所有医生的文本评价和评级.表 6 列出了用户对部分医生的各项评分,通过对所有医生的所有文本评价进行情感分析,得出其文本评价分 sa ,将其与该医生的评级得分 sr 按公式(5)加权计算满意度得分.由于文本评价更能体现患者对这次诊疗的真实情感,因此对参数 α 的设置采用 CRITIC 法、乘法集成法和主观评判相结合的方式.CRITIC 依据不同属性的信息量对各自的权值进行客观赋值,最终 α 取值约为 0.8,则综合满意度则可由 $su_{ij}=0.8 \cdot sa_{ij}+0.2 \cdot sr_{ij}$ 计算得到.

Table 6 Comprehensive score table

表 6 综合分数表

医生 ID	平均文本评价得分	平均评级得分	综合满意度得分
3088	3.946 237	4.301 075	4.017 204
7857	4.090 909	4.545 455	4.181 818
3136	4.088 889	5.000 000	4.271 111
3086	4.034 483	5.000 000	4.227 586
7898	4.311 111	5.000 000	4.448 889
3119	4.068 182	5.000 000	4.254 545
4039	3.500 000	4.000 000	3.600 000
7855	4.409 091	5.000 000	4.527 273
1495	4.142 857	5.000 000	4.314 286
7859	4.000 000	4.444 444	4.088 889
4052	4.000 000	4.777 778	4.155 556
3091	4.112 500	5.000 000	4.290 000
3018	4.255 814	5.000 000	4.404 651

4.2 服务匹配度计算

通过文本提取的方式生成文字摘要这种特殊的概述描述文件,将每个医生的若干问诊记录进行如下处理:只保留第一次问答,去掉“历史问诊 n ”、“患者信息:”和“问诊记录:”等这些无用信息.SnowNLP 依据 TextRank 算法提取文本摘要.我们根据病情描述的一般长度,对每个医生的文件用 SnowNLP 提取 12 句话,中间用逗号隔开,作为该医生历史问诊的摘要.由于后续相似度计算的限制,摘要截取前 250 字,对绝大多数摘要没有影响.由于摘要和任务的描述均为文字文本,故采用文本相似度计算.文本相似度就是对比两个输入文本之间的共性和差异,共性越大,差异越小,则相似度就越大.近年来,关于相似度计算的研究有很多,包括基于词向量、基于字符的、基于概率统计以及基于词嵌入模型的.目前的计算方法主要依赖于深度学习.本文的文本相似性分析方法利用词向量技术解决关键词匹配失败的问题,可以有效得到同义词、近义词之间的相似度,泛化能力好.同时,基于深度神经网络,对从单词语义到短文本语义的组合过程进行建模,表达效果也更好.

病情分科的处理与摘要提取基本相同,分类器的输入为文本文档,每行一条病情,分科结果为科室代码,添加在病情的同一行,使用分隔符隔开.相似度计算时,从上述文件中逐行提取病情及科室代码,在数据库中调取该科室所有医生的摘要,分别与病情描述成对输入相似度计算模型,提出病情的患者 ID 与医生 ID 作为键对值,用于捕获计算结果.例如,患者王某提交了关于病情的一段文字描述如下.

女,32 岁:您好,医生,我最近半年多来总是早晨起床从喉咙里有浓痰咳出,有时候是一小块咖啡色,有时候就是浓黄痰,也不是天天有,最近几天又有了,晨起咳完以后白天又没有了,我能察觉出来是喉咙里咳出来的,不是肺部也没有其他什么不适,一直有单侧扁桃体肿大.请问我这是慢性咽喉炎么?可以吃点什么药么?

先将病情根据其描述分类,即判断该病情适合的科室,随后对病情与科室中的医生诊疗病例记录的摘要进行相似度计算和排序.王某的病情被分类为耳鼻咽喉科,则该科的医生全部进入备选集,经过对这些医生诊疗摘要和病情描述的相似度计算,得到相似度排序表(见表 7).在该表的基础上,选出相似度最高的 k 位医生(这里假设 $k=10$).

医生最终得分由医生用户满意度分、医生职称和医院等级计算得出.此例中,医院等级共 10 级,医生职称等级共有 5 级,故根据公式(6)可以分别计算出对应得分.通过对各属性在医生最终得分中发挥的作用,根据文中所述方法,建立判断矩阵 M ,见表 8.

Table 7 Similarity ranking table (ear nose throat branch)**表 7** 相似度排序表(耳鼻咽喉科)

医生 ID	相似度
3088	0.902 617
7857	0.898 001
3136	0.889 262
3086	0.885 038
7898	0.882 464
3119	0.876 795
4039	0.875 016
7855	0.872 750
1495	0.871 829
7859	0.871 360
4052	0.870 101
3091	0.869 392
3018	0.865 689

Table 8 Judgment matrix of criterion layer**表 8** 准则层的判断矩阵

	准则		
	<i>su</i>	<i>sl</i>	<i>si</i>
医生的满意度得分 <i>su</i>	1	2	3
医生的职称得分 <i>sl</i>	1/2	1	2
医院的等级得分 <i>si</i>	1/3	1/2	1

由公式(7)和公式(8)计算得出医生的用户满意度得分、医生职称得分和医院等级得分的权重分别为 0.54, 0.3 和 0.16. 权重赋值后,前文筛选出的 10 位医生各项分值见表 9. 其中,医生 7898、7857 和 3136 的最终得分最高,故在综合考虑到人机物环境下的各方面因素后,最终推荐给患者王某的最合适的医生为 7898、7857 和 3136.

Table 9 Various score table**表 9** 各项得分表

医生 ID	综合满意度得分	职称等级得分	所属医院等级得分	最终得分
3088	4.017 204	3	5	3.869 290
7857	4.181 818	5	5	4.558 182
3136	4.271 111	4	5	4.306 400
3086	4.227 586	4	5	4.282 896
7898	4.448 888	5	5	4.702 400
3119	4.254 545	4	5	4.297 454
4039	3.600 000	5	5	4.244 000
7855	4.314 286	3	5	4.029 714
1495	4.314 286	4	5	4.268 000
7859	4.088 888	4	5	4.208 000

本文方法旨在为患者推荐合适的医生而不是最好的医生. 随机选择 100 个病例,对与每个病历匹配的候选医生的职称以及医院等级分布做了统计. 结果表明:与每个病历匹配的候选医生中,主任医师、副主任医师和主治医师出现的频率较高,且主要分布在三级甲等医院,二级甲等医院也有出现. 最终推荐的 3 位医生中,虽然主任医师占了大多数,但副主任医师以及主治医师也有出现. 因为该就诊平台上的医生 90%来自于三甲医院,因此在算法开始之前,非三甲医院的医生入选推荐队列的机率就比较小;另一方面是由于各属性的权值设置的影响,使得医院等级的权重略大于实际情况. 因此在不同的服务平台上,可能需要根据实际情况修正权值.

5 实验评估

本次实验采用百度的飞桨(PaddlePaddle)平台进行. 它是集深度学习核心框架. 工具组件和服务平台为一体的开源深度学习平台. 实验准备阶段,利用 PyCharm 完成了数据的爬取和处理;实验前期,在基于飞桨的一站式 AI 开发平台 AI Studio 上完成了科室分类器模型的训练以及对新患者病情描述的科室预测;后期在计算新病情与医生诊疗病例记录摘要的相似度以及实验评估阶段时计算的工作量庞大,实验过程中采用了并行计算同时

处理多批数据,提高了计算的速度和效率.

5.1 数据集概述

本实验使用的数据集为某在线医生诊疗平台的历史诊疗数据,使用爬虫从该平台上获取了共 11 个科室的 18 000 条原始数据,包含患者病情描述、医生 ID、医生专长、用户评论、医生所在医院以及医生职称.将原始数据的每个科室分别选取各 800、600、400、200 条病情作为训练集训练分类器.通过比较 4 个数据集训练出模型的分科室准确率,最终选择各科室 600 条病情的数据集训练分类器.按训练集比例,每科室 177 条共 1 947 条问诊记录作为测试集,将测试集的医生作为标准分派结果进行标注,用于评估本文方法的分派效果.

5.2 评价指标

本文旨在根据用户提交的任务描述,分配擅长执行该任务的业务员.在本文的病患分诊实验中,需要评价分派的医生和测试集标注的医生的差异.文中数据集由爬虫获取,没有类似数据的属性.实验中采用他们的诊疗病历记录摘要的相似性和专长相似性来刻画,这里的相似度计算与前文的病情和诊疗摘要相似度计算类似.实验选取的 3 个评价指标是:

- (1) 平均用户满意度 $Mean\left(\sum_i^{|U|} su_i\right)$;
- (2) 平均摘要相似度 $Mean\left(\sum_i^{|U|} asim(v_i^{\wedge}, v_i)\right)$;
- (3) 平均专长相似度 $Mean\left(\sum_i^{|U|} ssim(v_i^{\wedge}, v_i)\right)$.

其中, v_i^{\wedge} 表示各种方法为第 i 个病情推荐的医生, v_i 表示测试集标注的医生, $asim$ 是两个医生间的摘要相似度, $ssim$ 是两个医生之间的专长相似度.

5.3 实验集设置

通过对用户文本评价的情感分析和用户评级结合,预测用户的满意值,将物理世界固有属性经过分级,并与预测的用户满意值相结合,通过层次分析法确定权重,得到综合分数.通过历史任务的记录生成医生诊疗病历记录的摘要,对摘要和任务进行相似度计算并排序,选取前 k 位医生再根据综合分数排序,生成推荐列表.因此需要确定 k 值以进行后续实验,选取结果较好时的 k 值继续实验.实验步骤如下.

- (1) 设置 k 值从 10 以步长为 5~50 不断变化,问诊患者数目 N 分别设置为测试集的 20%,40%,60%,80%和 100%,使后续实验的 k 值能在用户综合评价和任务业务员匹配度上表现较好.
- (2) 依据历史患者的文本评价和评级计算得出医生的综合得分,预测患者对医生的满意度.
- (3) 寻找与病患匹配度高的医生.依据病患与医生诊疗摘要的相似度排序,选择相似度较高的 k 位医生.
- (4) 将步骤(3)中得到的医生的职称和所在医院的等级映射为现实空间的特征得分,并与步骤(2)中得到的综合得分加权相加,最终得分是对医生的多个角度综合评估,以此排序最终生成容量为 3 的推荐列表.
- (5) 为了验证患者数目发生变化时本文方法是否有相同的效果,将患者数目设置为从测试集病情数目的 10%以步长为 10%增加到 100%(同时,为了确保与训练集的病情记录分布一致,这里指各科室病情数目的 10%).

5.4 对比方法

为了评价文中方法的有效性,通过对比本文方法的 3 个阶段,并选取两种现有的主流服务推荐方法作为基准进行比较.为了保护患者隐私,本文实验在收集在线诊疗数据时已将患者个人信息排除,文中不再详细对比类似文献[4]中兼顾考虑用户和项目两方面的服务推荐方法.一方面,这些工作因为对人机物交互带来的各方面人为因素的影响考虑不足,会降低服务分派的准确率和用户体验;另一方面,也说明了本文方法属于第三方评测方式,独立于服务请求者和提供者双方,具有更好的适用性和公平性.因此,文中选择基于用户和基于项目的两种服务推荐方法作为基准方法进行对比.

- (1) 考虑服务匹配度的方法(C-M):选择与用户需求服务最相似的一组服务提供者.根据分类器得出新病

- 情的合理科室,计算出新病情与科室中医生诊疗病例记录摘要的相似度,得到相似度较高的 3 位医生.
- (2) 考虑服务匹配度和满意度的方法(CS-MS):选择用户倾向性最高的一组服务提供者.根据分类器得出新病情的合理科室,计算出病情与该科室中所有医生诊疗摘要的相似度,得到相似度较高的 k 位医生,从这 k 位医生中选择满意度最高的 3 位医生.
 - (3) 基于项目相似性的方法(I-S)^[40].即将相似的一组项目的用户打分作为标准,为其他用户推荐预测分高的项目.在病患就诊场景下,相似的一组项目可以理解为同科室的医生提供的服务,根据历史患者对医生的评级作为新病患对医生的评级预测,推荐预测值较优的 3 位医生.
 - (4) 基于用户相似性的方法(U-S)^[41,42].即以与目标用户相似的一组历史用户的打分为依据,根据与目标用户的相似度作修正打分的权值.计算历史患者的病情与新病患的相似度,并根据历史患者的评级计算预测分.预测分=病患相似度*评级得分,最终择优选择 3 位医生.
 - (5) 本文方法:考虑服务匹配度、满意度和成熟度(CSP-MSM).根据分类器得出新病情的合理科室,计算出病情与该科室中所有医生诊疗摘要的相似度,得到相似度较高的 k 位医生.在这 k 位医生中,结合医院等级、医生职称和满意度得分,找出最合适的 3 位医生.

5.5 实验结果与分析

图 2~图 4 分别展示了在 k 值变化时的平均满意度得分、平均专长相似度和平均摘要相似度.图中的 5 条曲线来自于病患数目 N 分别取为测试集的 20%,40%,60%,80%和 100%.结果表明:尽管病患数目 N 不同,但 k 值低于 35 时的各指标各不相同且存在波动, k 值在 35 后的各项指标都趋于平稳.

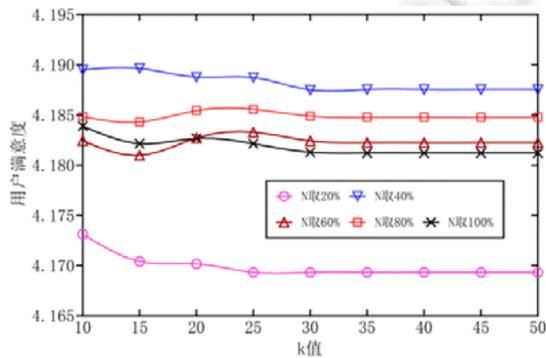


Fig.2 Average satisfaction scores at different k values

图 2 不同 k 值时的平均满意度得分

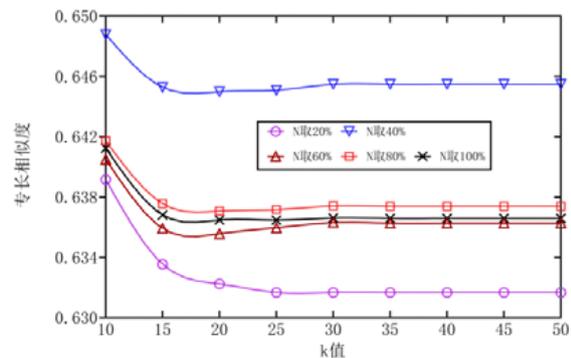


Fig.3 Mean expertise similarity at different k values

图 3 不同 k 值时的平均专长相似度

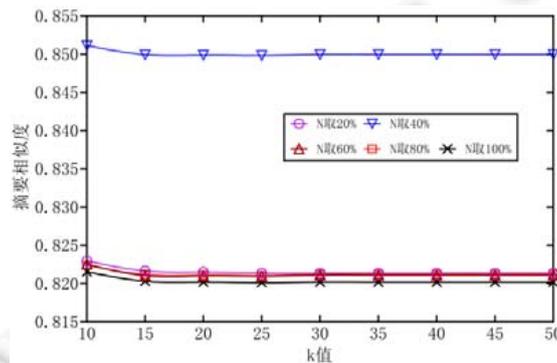


Fig.4 Mean abstract similarity at different k values

图 4 不同 k 值时的平均摘要相似度

表 10 展示了 k 值变化时,3 项指标平均值的变化.可以看出: k 值大于 35 后,在足够小的数量级上,各项指标没有明显变化.因此,若 k 值设置过大,不会改善本文方法的推荐效果,且会降低分配效率.故在后续实验中,取参数 k 值为 35 更新备选集.

Table 10 Mean value of each index under different k values

表 10 不同 k 值下各指标的平均值

k 值	平均用户满意度	平均专长相似度	平均摘要相似度
10	4.182 763 27	0.642 288 37	0.828 119 08
15	4.181 500 17	0.637 828 82	0.826 834 93
20	4.181 962 19	0.637 280 43	0.826 754 28
25	4.181 818 46	0.637 276 15	0.826 679 46
30	4.181 086 45	0.637 507 49	0.826 771 34
35	4.181 024 25	0.637 489 30	0.826 765 87
40	4.181 024 25	0.637 489 30	0.826 765 87
45	4.181 024 25	0.637 489 30	0.826 765 87
50	4.181 024 25	0.637 489 30	0.826 765 87

同时,为了验证本文方法 3 个主要模块的作用,按照重要程度,分别将本文方法拆分为服务匹配度计算模块、服务满意度预测模块和服务成熟度分析模块.在核心的匹配度模块上,逐步加入服务满意度预测和服务成熟度分析,即考虑服务匹配度的方法(C-M)、考虑服务匹配度和满意度的方法(CM-MS)以及本文考虑服务匹配度、满意度和成熟度的方法(CSP-MSM),进行纵向对比实验.

图 5 展示了在病人数目 N 变化时,各个方法选出的医生的平均用户满意度得分.其中,基于项目相似性的方法表现最优,平均用户满意度得分基本维持在 4.9 左右;考虑服务匹配度和满意度的方法则略低,平均用户满意度得分基本维持在 4.5 左右;基于用户相似性的方法和考虑服务匹配度的方法表现最差,是由于这两种方法忽略了用户评价,缺少对用户情感倾向的预测,本文方法的平均用户满意度得分居中.

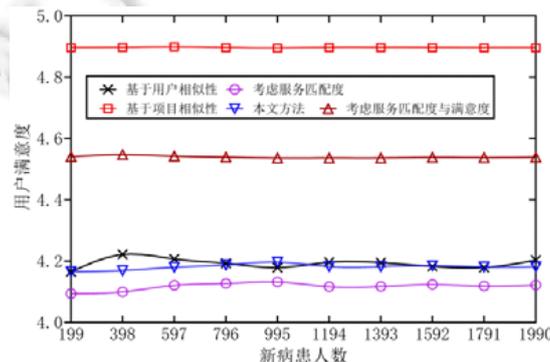


Fig.5 Average satisfaction scores for different numbers of patients

图 5 患病人数不同时的平均满意度得分

表 11 展示了不同方法在各项指标上的平均值,本文方推荐出的医生的评分平均值为 4.2 左右,是可以接受的,说明本文方法推荐的医生基本符合患者的情感倾向,能够获得较好的用户满意度.

图 6 展示了在病人数目 N 变化时,各个方法选出的医生与测试集标注医生的相似度平均值.客观上来说,医生的专长是患者在选择医生时首先考虑的,而诊疗摘要是对医生历史问诊记录的概括,因此在以专长相似度为指标对比时,摘要相似性高的医生,他们的专长相似性也高.本文方法在考虑服务匹配度的基础上同时考虑用户情感倾向,与只考虑服务匹配度的方法在纵坐标上相差很小,曲线基本重合,且明显优于考虑服务匹配度和满意度的方法.

图 7 展示了在病人数目 N 变化时,各个方法选出的医生与测试集标注医生的平均摘要相似度.因为考虑服务匹配度的方法只考虑病情与诊疗摘要的相似度,故该方法在此项指标上表现最好,维持在 82%以上.服务匹配度和满意度的方法和本文方法相差不大,略低于本文方法 1%左右.图 5~图 7 的结果验证了本文方法的各个模块

具有较好的效果,能够在多个角度达到预期的功能.

Table 11 Mean of each indicator of different methods

表 11 不同方法各个指标的平均值

用户满意度	基于用户相似性(U-S)	4.192 07
	基于项目相似性(I-S)	4.896 67
	考虑服务匹配度(C-M)	4.117 21
	考虑服务匹配度和满意度(CM-MS)	4.539 77
	本文方法(CSP-MSM)	4.181 09
专长相似度	基于用户相似性(U-S)	0.589 44
	基于项目相似性(I-S)	0.573 83
	考虑服务匹配度(C-M)	0.647 83
	考虑服务匹配度和满意度(CM-MS)	0.605 31
	本文方法(CSP-MSM)	0.637 59
摘要相似度	基于用户相似性(U-S)	0.789 69
	基于项目相似性(I-S)	0.783 12
	考虑服务匹配度(C-M)	0.827 91
	考虑服务匹配度和满意度(CM-MS)	0.815 81
	本文方法(CSP-MSM)	0.823 82

图 6 和图 7 中,患者数目在测试集的 10%到 100%变化时,基于用户相似性的方法和基于项目相似性的方法在平均专长相似度和平均摘要相似度两项指标上表现较差,仅有 57%~59%和 78%~81%,均低于本文方法的 63%~65%和 82%~85%,说明本文方法在患者数目变化时推荐的医生能够与不同病情的患者更加匹配.图 5 中,基于项目相似性的方法在预测用户情感倾向时效果优异,但基于用户相似性的方法表现最差.

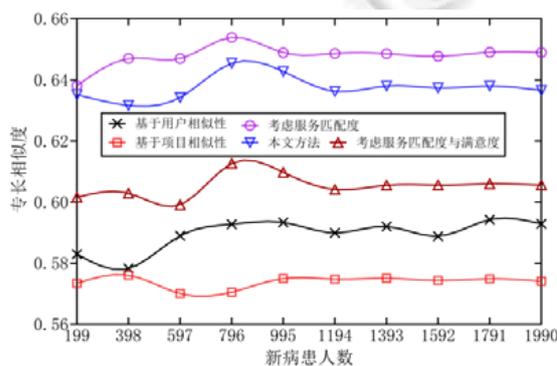


Fig.6 Mean expertise similarity for different numbers of patients

图 6 患病人数不同时的平均专长相似度

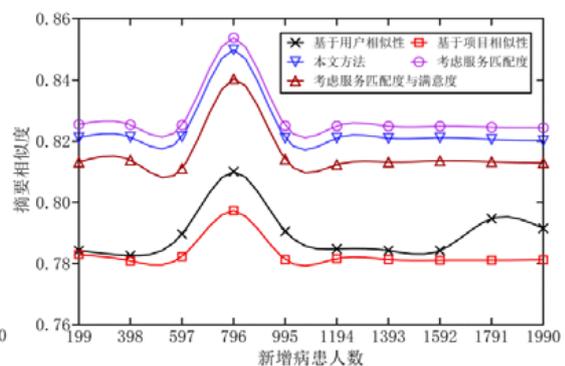


Fig.7 Mean abstract similarity for different numbers of patients

图 7 患病人数不同时的平均摘要相似度

综合 3 个评价指标来说,本文方法推荐的医生首先与患者的病情更加匹配,能够适应患者不同病情的需求,且该医生的用户满意度也比较优异,说明比较符合用户的情感倾向,能够获得较好的满意度.另外,本文方法也考虑了其他有利于患者的相关因素,相对于其他方法,可以提升用户体验和诊疗效果.

6 总结与展望

本文从人机物三元一体的多个角度出发,提出了面向数据特征的多域协作的服务分派方法.从人类社会角度出发,关注认知空间的属性特征,着重分析用户对于业务员服务的满意度.常规评分中的分数是用户给出的评级,为了更全面地表达用户对于此次服务的满意度,文中通过挖掘用户评价信息,对服务评价进行情感倾向性分析,并与评级分数结合实现情感倾向预测.从信息空间的角度出发,关注任务与服务内容的匹配度,着重分析任务与业务员服务内容的匹配度.通过引入业务员概要描述文件,并和任务特征进行相似度计算,来推荐更适合执

行相应任务的业务员.从物理世界的角度出发,关注现实空间的属性特征,分析业务员的服务成熟度,通过分析固有属性特征对于服务执行的影响,为其建立分级映射机制,以推荐服务成熟度更高的业务员.将人机物各方面数据特征相结合,最终得到更合理的服务分派结果,优化了任务分配和用户体验.本文在传统服务推荐的基础上,考虑人机物融合环境下特殊推荐情况,深入分析了用户请求服务后的执行过程中的各方面因素,判断这些因素是否对服务执行具有影响,以及影响的大小,并将这些因素融入到服务分派的选择过程中,优化了传统推荐模式下的效率问题,降低了安全隐患,使得分派之后,服务拥有更为良性的反馈.与其他服务推荐方法比较,本文方法注重人机物的三元环境的各个角度,而不仅仅是只从服务执行效率或是用户评价等个别方面推荐服务,因此具有更高的可信度,降低了因人的参与而导致的真实性差异.

由于患者位置信息属于个人隐私数据,本文没有加入地理位置的相关属性.此外,本文方法没有考虑用户对于服务时间的要求,以及业务员在单位时间内服务的响应次数.可以结合地理信息数据集进一步改进本文方法,更多考虑时间和执行效率的影响.未来的工作可以结合一些经典的机器学习分析方法,挖掘目标任务的业务员分派倾向.例如,可以在深度学习的基础上,采用投票机制与本文使用的深度学习方法相结合,可以进一步优化用户体验.

致谢 感谢审稿专家提出的宝贵意见和建议,同时感谢江苏省未来网络科研基金项目提供的网络实验环境,以及百度、优速和谷歌支持的教育部产学研合作协同育人项目提供的在线计算资源.特别感谢课题组的孔瀚民、陈光忠、沈天和张梁等同学为数据处理和实验工作付出的努力,以及周嘉浩和蒋雨霏同学主持的大学生创新创业训练计划项目完成了论文的部分工作.

References:

- [1] Ma XX, Liu XZ, Xie B, Yu P, Zhang T, Bu L, Li XD. Software development methods: review and outlook. *Ruan Jian Xue Bao/Journal of Software*, 2019,30(1):3–21 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/5650.htm> [doi: 10.13328/j.cnki.jos.005650]
- [2] Wang HM, Mao XG, Ding B, Shen J, Luo L, Ren Y. New insights into system software. *Ruan Jian Xue Bao/Journal of Software*, 2019,30(1):22–32 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/5648.htm> [doi: 10.13328/j.cnki.jos.005648]
- [3] Huang ZH, Zhang JW, Tian CQ, Sun SL, Xiang Y. Survey on learning-to-rank based recommendation algorithms. *Ruan Jian Xue Bao/Journal of Software*, 2016,27(3):691–713 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/4948.htm> [doi: 10.13328/j.cnki.jos.004948]
- [4] Zou G, Jiang M, Niu S, Wu H, Pang S, Gan Y. QoS-aware Web service recommendation with reinforced collaborative filtering. In: *Proc. of the 16th Int'l Conf. on Service-oriented Computing*. Cham: Springer Int'l Publishing, 2018. 430–445. [doi: 10.1007/978-3-030-03596-9_31]
- [5] Pessiot JF, Truong TV, Usunier N, Amini MR, Gallinari P. Learning to rank for collaborative filtering. In: *Proc. of the 9th Int'l Conf. on Enterprise Information Systems*. New York: ACM, 2007. 145–151.
- [6] Zhang W, Wang JY. A collective Bayesian Poisson factorization model for cold-start local event recommendation. In: *Proc. of the 21th ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining*. New York: ACM, 2015. 1455–1464. [doi: 10.1145/2783258.2783336]
- [7] Li L, Zhu G, Xie Q, Su C, Yang ZL. Recommendation approach by simultaneous learning latent features and preferences guidance. *Ruan Jian Xue Bao/Journal of Software*, 2019,30(11):3382–3396 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/5542.htm> [doi: 10.13328/j.cnki.jos.005542]
- [8] Pazzani MJ, Billsus D. Content-based recommendation systems. In: *Proc. of the Adaptive Web: Methods and Strategies of Web Personalization*. Berlin, Heidelberg: Springer-Verlag, 2007. 325–341.
- [9] Aalst WMP. Three good reasons for using a Petri-net-based workflow management system. In: *Proc. of the Information and Process Integration in Enterprises: Rethinking Documents*. Boston: Springer-Verlag, 1998. 161–182.
- [10] Alonso G, Agrawal D, Abbadi AE, Mohan C. Functionality and limitations of current workflow management systems. *IEEE Expert*, 1997,12(5):1–25.
- [11] Mohan C. Recent trends in workflow management products, standards and research. In: *Proc. of the Workflow Management Systems and Interoperability*. Berlin, Heidelberg: Springer-Verlag, 1998. 396–409.

- [12] Luo HB, Fan YS, Wu C. Overview of workflow technology. *Ruan Jian Xue Bao/Journal of Software*, 2000,11(7):899–907 (in Chinese with English abstract). http://www.jos.org.cn/jos/article/abstract/20000706?st=article_issue
- [13] Li HC, Shi ML. Workflow models and their formal descriptions. *Chinese Journal of Computers*, 2003,26(11):1456–1463 (in Chinese with English abstract).
- [14] Teltumbde A. A framework for evaluating ERP projects. *Int'l Journal of Production Research*, 2000,38(17):4507–4520. [doi: 10.1080/00207540050205262]
- [15] Colmenares LE. A prototype decision support system for ERP evaluation in small and medium enterprises. In: *Proc. of the Int'l Conf. on Information Resources Management Association*. Hersher: Idea Group, 2006. 1015–1018.
- [16] Wortmann JC, Ittoo A. Incorporating text in enterprise information systems. *Int'l Federation of Automatic Control Proc. Volumes*, 2013,46(9):590–595.
- [17] Xu GY, Tao LM, Shi YC, Zhang X. Human computer interaction for ubiquitous/pervasive computing mode. *Chinese Journal of Computers*, 2007,30(7):1041–1053 (in Chinese with English abstract).
- [18] Cheng SW, Zhu AJ, Fan J. Eye tracking and gesture based interaction for target selection on large display. *Ruan Jian Xue Bao/Journal of Software*, 2018,29:75–85 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/18019.htm>
- [19] Li WC, Horn A, Sun Z, Zhang JY, Braithwaite G. Augmented visualization cues on primary flight display facilitating pilot's monitoring performance. *Int'l Journal of Human-computer Studies*, 2020,135:102377–102386. [doi: 10.1016/j.ijhcs.2019.102377]
- [20] Walton LA, Worboys M. A qualitative bigraph model for indoor space. In: *Proc. of the Int'l Conf. on Geographic Information Science*. Berlin, Heidelberg: Springer-Verlag, 2012. 226–240.
- [21] Tan Y, Vuran MC, Goddard S, Yu Y, Song M, Ren SP. A concept lattice-based event model for cyber-physical systems. In: *Proc. of the 1st ACM/IEEE Int'l Conf. on Cyber-physical Systems*. New York: ACM, 2010. 50–60.
- [22] Jensen JC, Chang DH, Lee EA. A model-based design methodology for cyber-physical systems. In: *Proc. of the 7th Int'l Wireless Communications and Mobile Computing Conf. Piscataway: IEEE*, 2011. 1666–1671.
- [23] Tsiganos C, Pasquale L, Ghezzi C, Nuseibeh B. On the interplay between cyber and physical spaces for adaptive security. *IEEE Trans. on Dependable and Secure Computing*, 2018,15(3):466–480. [doi: 10.1109/tdsc.2016.2599880]
- [24] Ye B, Wang Y. CrowdRec: Trust-aware worker recommendation in crowdsourcing environments. In: *Proc. of the 2016 IEEE Int'l Conf. on Web Services*. 2016. 1–8. [doi: 10.1109/ICWS.2016.10]
- [25] Taghavi M, Bentahar J, Otrok H, Wahab OA, Mourad A. On the effects of user ratings on the profitability of cloud services. In: *Proc. of the IEEE 24th Int'l Conf. on Web Services*. Piscataway: IEEE, 2017. 1–8. [doi: 10.1109/icws.2017.8]
- [26] Cheng JJ, Zhang X, Fan PY, Li P, Wang H. A rule-based approach to sentiment classification of Chinese microblogging texts. *Advanced Materials Research Volumes*, 2013,765-767:1441–1445. [doi: 10.4028/www.scientific.net/AMR.765-767.1441]
- [27] Gao Y, Zhao WQ, Hou XK. Sentiment analysis based on micro blog hot topics. *Advanced Materials Research*, 2013,753-755:3056–3059. [doi: 10.4028/www.scientific.net/AMR.753-755.3056]
- [28] Park S, Kim Y. Building thesaurus lexicon using dictionary-based approach for sentiment classification. In: *Proc. of the IEEE 14th Int'l Conf. on Software Engineering Research, Management and Applications (SERA)*. Piscataway: IEEE, 2016. 39–44.
- [29] Rao YH, Lei JS, Liu WY, Li Q, Chen ML. Building emotional dictionary for sentiment analysis of online news. *World Wide Web*, 2014,17(4):723–742. [doi: 10.1007/s11280-013-0221-9]
- [30] Hatzivassiloglou V, McKeown KR. Predicting the semantic orientation of adjectives. In: *Proc. of the 35th Annual Meeting of the Association for Computational Linguistics and 8th Conf. of the European Chapter of the Association for Computational Linguistics*. Stroudsburg: Association for Computational Linguistics, 1997. 174–181.
- [31] Turney PD. Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. In: *Proc. of the 40th Annual Meeting of the Association for Computational Linguistics (ACL 2002)*. Stroudsburg: Association for Computational Linguistics, 2002. 417–424.
- [32] Mikolov T, Sutskever I, Chen K, Corrado G, Dean J. Distributed representations of words and phrases and their compositionality. In: *Proc. of the 26th Int'l Conf. on Neural Information Processing Systems*. Red Hook: Curran Associates, 2013. 3111–3119.
- [33] Le Q, Mikolov T. Distributed representations of sentences and documents. In: *Proc. of the 31st Int'l Conf. on Machine Learning*. 2014. II-1188–II-1196.
- [34] Du C, Wang ZH, Jiang JJ, Sun YG. Bayesian classifier algorithm based on emerging pattern for data stream. *Ruan Jian Xue Bao/Journal of Software*, 2017,28(11):2891–2904 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/5350.htm> [doi: 10.13328/j.cnki.jos.005350]
- [35] Chen Q, Zhang L, Jiang J, Huang XY. Review analysis method based on support vector machine and latent Dirichlet allocation. *Ruan Jian Xue Bao/Journal of Software*, 2019,30(5):1547–1560 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/5731.htm> [doi: 10.13328/j.cnki.jos.005731]

- [36] Deng ZY, Zhu XS, Cheng DB, Zong M, Zhang SC. Efficient KNN classification algorithm for big data. *Neurocomputing*, 2016,195: 143–148. [doi: 10.1016/j.neucom.2015.08.112]
- [37] Peters ME, Neumann M, Iyyer M, Gardner M, Clark C, Lee K, Zettlemoyer L. Deep contextualized word representations. In: Proc. of the 2018 Conf. of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Stroudsburg: Association for Computational Linguistics, 2018. 2227–2237.
- [38] Devlin J, Chang MW, Lee K, Toutanova K. BERT: Pre-training of deep bidirectional transformers for language understanding. In: Proc. of the 2019 Conf. of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Stroudsburg: Association for Computational Linguistics, 2019. 4171–4186.
- [39] Sun Y, Wang SH, Li YK, Feng SK, Chen XY, Zhang H, Tian X, Zhu DX, Tian H, Wu H. ERNIE: Enhanced Representation through Knowledge Integration. arXiv preprint arXiv:1904.09223, 2019.
- [40] Sarwar B, Karypis G, Konstan J, Riedl J. Item-based collaborative filtering recommendation algorithms. In: Proc. of the 10th Int'l Conf. on World Wide Web. New York: ACM, 2001. 285–295.
- [41] Arora G, Kumar A, Devre GS, Ghumare PA. Movie recommendation system based on user's similarity. *Int'l Journal of Computer Science and Mobile Computing*, 2014,3(4):765–770.
- [42] Ekstrand MD, Riedl JT, Konstan JA. Collaborative filtering recommender systems. *Foundations and Trends in Human-Computer Interaction*, 2011,4(2):81–173. [doi: 10.1561/1100000009]

附中文参考文献:

- [1] 马晓星,刘讚哲,谢冰,余萍,张天,卜磊,李宣东. 软件开发方法发展回顾与展望. *软件学报*,2019,30(1):3–21. <http://www.jos.org.cn/1000-9825/5650.htm> [doi: 10.13328/j.cnki.jos.005650]
- [2] 王怀民,毛晓光,丁博,沈洁,罗磊,任怡. 系统软件新洞察. *软件学报*,2019,30(1):22–32. <http://www.jos.org.cn/1000-9825/5648.htm> [doi: 10.13328/j.cnki.jos.005648]
- [3] 黄震华,张佳雯,田春岐,孙圣力,向阳. 基于排序学习的推荐算法研究综述. *软件学报*,2016,27(3):691–713. <http://www.jos.org.cn/1000-9825/4948.htm> [doi: 10.13328/j.cnki.jos.004948]
- [7] 李琳,朱阁,解庆,苏畅,杨征路. 一种潜在特征同步学习和偏好引导的推荐方法. *软件学报*,2019,30(11):3382–3396. <http://www.jos.org.cn/1000-9825/5542.htm> [doi: 10.13328/j.cnki.jos.005542]
- [12] 罗海滨,范玉顺,吴澄. 工作流技术综述. *软件学报*,2000,11(7):899–907. http://www.jos.org.cn/jos/article/abstract/20000706?st=article_issue
- [13] 李红臣,史美林. 工作流模型及其形式化描述. *计算机学报*,2003,26(11):1456–1463.
- [17] 徐光祐,陶霖密,史元春,张翔. 普适计算模式下的人机交互. *计算机学报*,2007,30(7):1041–1053.
- [18] 程时伟,朱安杰,范菁. 面向大屏幕目标选择的眼动跟踪与手势交互. *软件学报*,2018,29:75–85. <http://www.jos.org.cn/1000-9825/18019.htm>
- [34] 杜超,王志海,江晶晶,孙艳歌. 基于显露模式的数据流贝叶斯分类算法. *软件学报*,2017,28(11):2891–2904. <http://www.jos.org.cn/1000-9825/5350.htm> [doi: 10.13328/j.cnki.jos.005350]
- [35] 陈琪,张莉,蒋竟,黄新越. 一种基于支持向量机和主题模型的评论分析方法. *软件学报*,2019,30(5):1547–1560. <http://www.jos.org.cn/1000-9825/5731.htm> [doi: 10.13328/j.cnki.jos.005731]



袁敏(1977—),男,博士,副教授,CCF 专业会员,主要研究领域为服务计算,智能系统,软件工程,大数据分析,形式化验证.



徐冰青(1999—),女,学士,主要研究领域为大数据分析,边缘计算.



陈卓(1996—),男,硕士生,主要研究领域为大数据分析,边缘计算.