

区块链在个性化推荐系统中的应用研究综述

许小颖,陈熙,陈源,谢永靖
华南理工大学工商管理学院,广东 广州 510640

摘要

区块链作为一种新兴技术,以其去中心化、难以篡改、匿名性和可追溯性等特点,为个性化推荐系统的改进提供了一种崭新的思路。为此,首先对近年来推荐系统面临的主要问题和区块链技术带来的机遇进行归纳总结,然后采用文献分析方法,从时间分布、文献类型、研究问题和评估指标4个层面,对推荐系统中区块链技术的应用研究进行分析和总结。分析结果表明:区块链对于解决推荐系统的数据安全和隐私保护、数据共享、数据可信和推荐透明度问题有重要意义;已有研究主要集中于解决推荐系统中用户的数据安全和隐私保护问题,而在跨平台数据共享、数据激励机制设计和系统可扩展性等方面的研究仍有待进一步突破。

关键词

区块链;推荐系统;隐私安全;透明可信

中图分类号:C94

文献标志码:A

doi: 10.11959/j.issn.2096-0271.2022031

A review of blockchain applications in personalized recommender systems

XU Xiaoying, CHEN Xi, CHEN Yuan, XIE Yongjing

School of Business Administration, South China University of Technology, Guangzhou 510640, China

Abstract

Blockchain, as an emerging technology, provides a brand-new idea for the improvement of personalized recommender systems with its characteristics of decentralization, tamper-proof, anonymity and traceability. Therefore, the main problems faced by the recommender systems in recent years and the opportunities brought by blockchain technology were summarized firstly. Then, literature analysis was adopted to analyze and summarize the research on the application of blockchain technology in recommender systems from four aspects: time distribution, literature types, research questions and evaluation indicators. The results show that the blockchain is of great significance for solving the problems of data security and privacy protection, data sharing, data trustworthiness and recommendation transparency of recommender systems. Existing studies mainly focus on solving the problem of data security and privacy protection in recommender systems, while further breakthroughs are needed in cross-platform data sharing, design of incentive mechanisms and system scalability.

Key words

blockchain, recommender systems, privacy and security, transparency and reliability

0 引言

互联网的飞速发展,给用户带来了极大的便利。然而,在网络信息量快速增长的同时,也出现了严重的“信息过载”问题,导致用户难以从海量的数据中获取真正有用的信息,降低了信息的利用效率。推荐系统是解决该问题的有效手段,它由一系列智能推荐算法组成,通过信息过滤的方式为每个用户推荐可能感兴趣的内容^[1]。推荐系统已经在许多领域发挥着不可或缺的作用,特别是在电子商务领域,个性化推荐加快了用户的决策过程^[2],不仅提高了用户的满意度,也增加了商品的销量^[3-4]。

尽管推荐系统在研究上和应用上已经取得了一定的进展,但是依然面临着许多问题。在数据层面上,现有中心化的数据收集和存储方式使得用户的隐私保护和数据安全受到严重威胁。此外,推荐系统的有效性严重依赖于用户数据的丰富程度和准确程度^[5]。但受缺乏激励、缺少信任等不利因素的影响,用户提供和共享数据的意愿不高,数据的真实性也难以保证。在系统实现层面上,传统推荐系统的黑箱操作模式使得推荐流程的透明度较低,用户对推荐逻辑难以理解,严重影响了用户对推荐结果的信任程度。

近年来,区块链技术的迅速发展为解决上述问题提供了崭新的思路。区块链是比特币等数字加密货币的核心技术,以分布式记账为基础,具有去中心化、难以篡改、可追溯、可编程、公开透明和集体维护

等特征。由于区块链的独特优势,越来越多的学者开始分别从数据层面和系统实现层面研究区块链技术对推荐系统的优化和改进,有望从根本上应对数据安全和隐私保护、数据共享、数据可信和推荐透明度等方面的挑战,从而解决推荐系统存在的安全隐患和信任危机问题。

然而,目前国内外还缺乏对区块链在推荐系统中应用的研究综述。因此,本文将对近年来推荐系统面临的主要问题和区块链技术带来的机遇进行归纳总结,并通过对比已有文献的梳理,明晰区块链在推荐系统中应用的研究现状,提出未来的研究重点和方向。

1 推荐系统的研究现状和挑战

1.1 推荐系统简介

推荐系统可以被定义为一种应用,其核心是推荐算法。推荐系统的主要目的是在特定的场景下,探索用户的偏好,并从海量的数据中找到用户可能感兴趣的信息,为用户生成推荐列表。作为解决信息过载的一种有效手段,推荐系统受到产业界和学术界的广泛关注。

基于推荐原理的差异,传统的推荐算法可以分为3种主要的类型^[6],分别是协同过滤的推荐算法、基于内容的推荐算法和混合推荐算法。除了这些传统的推荐算法,在推荐系统的发展过程中也演变出许多效果显著的改进算法,比如基于社交的推荐算法^[7]、跨领域的推荐算法^[8]等。在

近几年的文献中,深度学习以其强大的表达能力,逐渐成为推荐领域采用的主流算法;对抗学习、强化学习、嵌入技术和注意力机制等技术也受到广泛关注,这些技术的应用显著提高了推荐算法的整体性能。

1.2 现有推荐系统的挑战

推荐系统在快速发展的同时,也面临着许多挑战,具体表现在数据和系统实现两个层面。在数据层面上,传统的推荐系统被部署在中心化服务器上,数据安全和隐私保护、数据共享、数据可信问题一直是制约推荐系统发展的主要障碍;在系统实现层面上,传统推荐系统呈现给用户的是一个无法观察和理解的黑箱,推荐逻辑不透明导致的用户信任程度低成为不可忽视的重要问题。

1.2.1 数据安全和隐私保护问题

推荐引擎一般架构在存储、计算资源充足的中心化服务器上,推荐数据中心化平台进行收集和存储,这些数据往往包含用户的隐私信息,例如姓名、位置和偏好等。这种集中式的存储方式使得用户产生的数据被平台控制,平台可在未经用户同意的情况下将数据转卖给第三方^[9],导致用户隐私的泄露。针对数据安全和隐私保护问题,一些企业提出了隐私保护政策来减少用户的风险感知,提高透明度;然而,这些政策并不能从技术层面保证用户的匿名性和数据的安全性、可控性^[10]。也有一些学者提出了保护用户隐私的数据存储和挖掘方法,包括完全同态加密^[11]和安全多方计算^[12]。但当数据量增大时,推荐计算的复杂度显著增加。

此外,这种存储在集中服务器上的客户数据也可能遭到黑客的恶意攻击,用户隐私

暴露风险大,在数据保护和安全方面非常脆弱。例如2018年,黑客利用系统漏洞窃取了脸书(Facebook)近5 000万用户的个人信息。如何从根本上解决数据安全和隐私保护问题,已经成为目前大多数推荐系统面临的难题之一。

1.2.2 数据共享问题

在推荐领域中,许多流行的推荐方法(例如协同过滤算法)往往面临着用户评分数据的稀疏性和冷启动问题。针对推荐数据稀疏问题,现有的解决途径主要集中在充分利用已有的数据上。一些学者提出采用数据降维技术,将高维稀疏数据映射到低维空间,从而更有效地利用有限的用户数据。例如Wang X H等人^[13]采用卷积矩阵分解处理推荐系统中的高度稀疏数据,以捕获更深层次的关系。此外,还有一些研究试图充分挖掘用户评分以外的其他信息作为补充。例如Reshma R等人^[14]从社交网络中获取用户间的信任关系和用户个人资料的相似性,结合评分数据进行推荐。这些方法都是在数据稀疏的前提下,尽可能多地从有限的现有数据中提取有用信息,对数据稀疏问题有一定的缓和作用。

然而,这些方法并没有从数据源的角度出发来提高数据供给的数量。数据共享是缓解数据稀疏性问题的有效途径,包括单个用户数据的分享和不同平台间数据的共享。为了促进数据供给,现有电商平台通常采用粗粒度的激励方式,即根据用户反馈的数据量,给予一定的平台积分奖励,这虽然在一定程度上增加了反馈的数据量,但也为推荐系统引入了大量无效数据,降低了推荐的质量^[15]。因此,需要根据数据实际价值和使用效果对用户进行精细化的激励。然而,在现有中心化平台的条件下,推荐系统的运行均为黑箱化,透明性和公

平性无法保证,精细化的激励机制也难以实施。同时,由于平台之间缺乏信任,无法保障数据使用的安全透明,跨平台跨领域的数据共享难以真正实现^[16]。

1.2.3 数据可信问题

除了需要确保数据的丰富性,真实有效的数据源也是推荐系统进行精准推荐的基础。然而,由于用户与平台之间缺乏信任、商家与平台合谋等情况的存在,推荐系统数据源的真实性难以保证。例如,在一些电商平台,许多商家为了提高自身的信誉,会雇佣大量网络写手为自己撰写虚假好评,甚至删除和修改用于推荐的用户数据。另外,现有电商平台大多采用粗粒度的激励方式,用户为了获取更多激励,会出现随意填写反馈的现象,这给推荐系统带来了大量的无效数据,平台难以对用户所反馈数据的质量进行有效的验证和控制。

针对数据可信问题,现有研究主要集中在如何抵御推荐攻击上。通常所说的推荐攻击指针对协同过滤推荐系统的攻击,在这种攻击中,恶意用户通过注入大量虚假信息来改变推荐结果,这可能会降低用户对推荐系统的客观性和准确度的信任^[17]。针对这些攻击,现有研究聚焦于两种应对方式。一种是检测推荐攻击,然后消除或降低这些攻击对推荐的影响^[18],例如Tong C等人^[19]提出一种基于卷积神经网络的方法来挖掘用户评分行为的深层特征,有效地检测推荐攻击。另一种是设计更加健壮的计算^[20],例如Yuan F等人^[21]提出一种基于神经网络推荐模型的通用对抗训练框架,通过加入人工噪声数据进行协同训练,提高模型的鲁棒性。目前,关于推荐系统的鲁棒性研究还比较少,而攻击策略却层出不穷。

1.2.4 推荐透明度问题

推荐透明指让用户了解推荐系统给出特定推荐结果的原因,这将直接影响用户对推荐系统的满意程度和对推荐结果的信任程度^[22],而让用户信任并采纳推荐结果,才是推荐系统发挥作用的关键所在。目前,大多数研究通过推荐解释的方式来提高推荐系统的透明度。例如,亚马逊在采用协同过滤推荐时,通过“买过本商品的用户还买了……”向用户提供推荐解释^[23]。然而,现有方法只是由中心化平台告知用户某个商品被推荐的原因,推荐过程仍然是个黑箱流程,推荐的透明度远远不够。因此,提高推荐透明度仍然是目前推荐系统面临的主要挑战之一。

综上所述,目前以数据驱动的推荐系统仍然面临着诸多问题,亟须一种创新性的解决方案从根本上打破当前困局。近年来,区块链技术的迅速发展为解决上述问题提供了崭新的思路。区块链具有去中心化、难以篡改、可追溯、可编程、公开透明和集体维护等特征,为解决推荐系统中的数据安全和隐私保护、数据共享、数据可信和推荐透明度问题提供了技术支撑。

2 区块链技术的优势

2.1 区块链技术及其特征

早期的区块链作为比特币系统中的底层支撑技术,是一个含有块链式数据结构且难以篡改伪造的分布式账本。而区块链发展至今,其本身的内涵与技术已发生了很多变化。总体来说,区块链技术是一项综合了加密算法、共识机制、P2P网络、智能合约等计算机技术的全新的基础架构和

技术体系。

区块链一般具有以下几个技术特征。

- 去中心化。区块链建立在P2P网络的基础上,参与网络的各个节点分散对等地参与协作,共同维护数据,达成共识。网络内部并不需要中心化的管理机构,当某个节点出现故障或自由加入和退出网络时,整个系统都能维持正常运行,单一机构对系统的影响大大降低。

- 公开透明。去中心化使得每个节点都能平等地发送和接收区块链网络中的消息,除了各方被加密的私有信息,区块链中的完整数据都能被任何节点查看和记录,整个区块链系统的运作对于节点来说都是公开透明的。

- 数据可信。数据的安全性主要体现在两方面:难以篡改伪造和可追溯。难以篡改伪造指由于共识机制的作用,数据一旦通过验证并写入区块链后就难以篡改,而伪造的数据难以使得各方节点达成共识,从而不被承认;可追溯指区块链存储数据的块链式结构有利于针对某一状态查出与其相关的所有操作历史。难以篡改伪造和可追溯这两点,再加上公开透明的特性,使得区块链上的数据是真实可信的。

- 匿名性。区块链中采用了非对称加密算法来保护用户节点的身份信息,用户对外公布的只是加密后的哈希地址,而其他用户无法从该地址中读取任何有效信息,这为用户提供了隐私安全保障。

2.2 区块链技术在推荐系统中的应用价值

区块链以其去中心化、加密匿名和数据的防伪溯源等特性,使得脱离中心化管理成为一种可能,在解决现有推荐系统面临的问题上具有重要的应用价值。

在数据安全和隐私保护问题上,首

先,区块链的去中心化消除了第三方控制的需 要,可以有效降低中心化数据节点存在的安全风险;其次,区块链内置的数字签名和加密技术方案,可与认证、访问控制等安全对策相结合,提高数据安全性,保护用户隐私^[24];最后,在区块链网络中,交易的整个过程被记录在区块链上,具有难以篡改和可追溯性,为数据所有者带来更大的透明度,保护数据不被滥用。

在数据共享问题上,区块链去中心化的结构使得数据交易不需要中介机构,结合加密技术和智能合约可构建全新的信任模式,保障交易安全快速完成;其次,由于其公开透明和难以篡改等特性,结合智能合约和代币机制,可以有效解决多方协同中激励计算与利益分配的信任问题。

在数据可信问题上,区块链采用基于共识的数据更新机制,可以在一定程度上消除不可信数据的影响,并结合代币激励机制来确保诚实用户的利益;此外,链上数据具有难以篡改和可追溯性,避免数据被恶意修改的风险。

在推荐透明度问题上,对用户数据的使用全过程进行跟踪和客观记录,使得数据具有难以篡改和可追溯的特点,从而提高了推荐系统内部的透明性和可信性;同时,结合智能合约,将区块链与推荐系统进行深度融合,可以提高整个推荐算法的透明度。

区块链通过可信任的点对点传输、难以篡改的数据存储和可追溯的数据查看,解决了多方协同的信任问题,让用户更加信任推荐系统及其应用平台,并愿意提供更加真实有效的数据,从而建立起平台的信任生态,促进推荐系统的良性发展。目前已有不少研究关注到了区块链技术带来的机遇,区块链在信誉系统^[24-26]、用户评价^[24]和电子商务^[10]中的应用受到了越来越多的关注,这也进一步推动了区块链与推荐系统的结合。

3 区块链在推荐系统中的应用现状

通过筛选国内外主要检索源中符合主题 的文献,对文献的时间分布、类型分布 等方面进行分类统计,并从主要研究问题 和评估指标两个层面对区块链在推荐系统 中的应用情况进行深入分析,明晰区块链 在推荐系统中的应用现状。

3.1 文献检索方法

在文献具体筛选方法层面,本文分别 以(“recommend” or “recommendation” or “recommender” or “recommender systems”) and “blockchain”、(“推 荐” or “推荐系统”) and “区块链”为检索 词,在Web of Science、EI Engineering Village 2、Google Scholar、中国知网、万 方数据等数据库中进行搜索;通过阅读所 选择的文献摘要,剔除研究重点不是区块 链在推荐系统上的应用的文章,并在文献 阅读过程中根据参考文献不断扩展文献范 围。最后发现相关文献40篇。

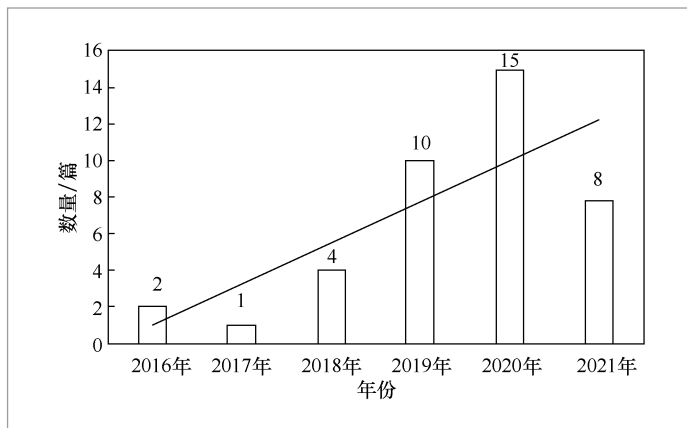


图1 文献的时间分布

3.2 文献分类统计

3.2.1 时间分布

如图1所示,从时间分布上看,最早的 与区块链推荐系统相关的文献出现在 2016年,区块链在推荐系统上的应用近年 才开始发展,是一个比较新的领域。近年 来,文献数量逐年增加,这也从侧面反映 了随着区块链技术的快速发展,区块链 推荐系统这一研究领域越来越受到学者们 的关注。

3.2.2 文献类型

如图2(a)所示,从文献类型分布 上看,大部分文献类型都是期刊论文 (60%),会议论文、专利和图书分别占 20%、18%和2%。从论文发表的期刊来 看,主要是信息系统和电子信息技术等技 术领域的期刊。从期刊的级别来看,有不 少论文发表在重要国际期刊和中文核心期 刊上,如*IEEE Transactions on Engineering Management*、*IEEE Transactions on Network Science and Engineering*和《西 安电子科技大学学报》等,这也从侧面反映 出区块链与推荐系统的结合已成为推荐 系统领域的研究前沿和热点。图2(b)是 从目前国内外文献分布的视角对筛选的结 果进行分类。从图2(b)可知,国内外文 献的占比分别为44%和56%,说明目前 国外对该领域的研究热度高于国内。图 2(c)根据研究方法对筛选结果进行分 类。从图2(c)可知,定性研究数量相 对较少,大多数研究是针对研究场景和 所需解决的问题设计一个系统框架,并 用实际的数据进行实验分析,通过定量 方法验证系统的有效性。因此,相比于 定性的纯理论论证和推导,基于实验和 实证的定量研究更适合此研究方向。

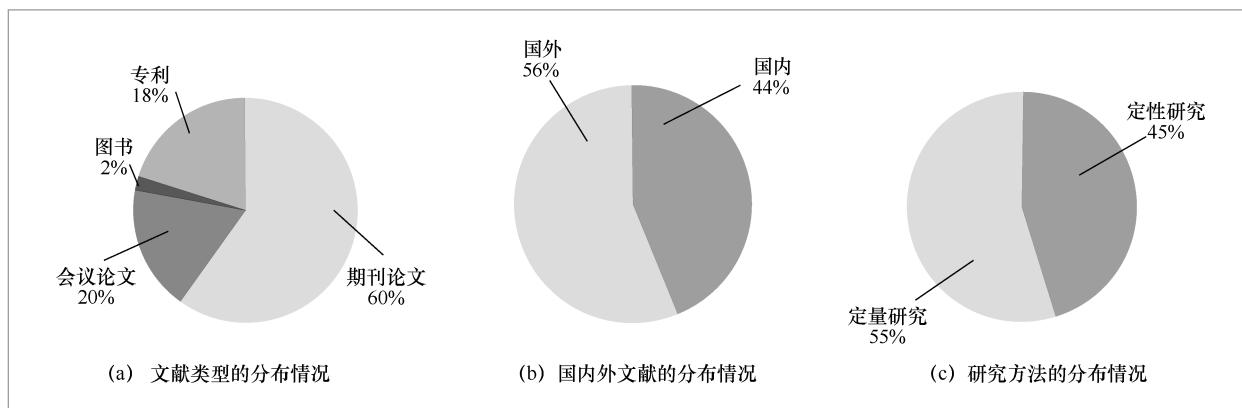


图2 文献类型

3.3 主要研究问题

从研究问题上看,区块链在推荐系统上的应用主要是为了解决4个问题,分别是推荐系统中的数据安全和隐私保护问题、数据共享问题、数据可信问题和推荐透明度问题。基于区块链的推荐系统研究总览见表1。关于研究解决数据安全和隐私保

护问题的文章占比最大,这说明保护用户隐私和数据安全被认为是区块链应用于推荐系统所带来的最显著优势。

3.3.1 数据安全和隐私保护

在传统的中心化推荐系统应用平台中,平台需要收集大量的用户敏感数据,这种集中式的数据存储方法容易受到恶意

表1 基于区块链的推荐系统研究总览

主要解决问题	代表文献	区块链的应用方式
数据安全和隐私保护	文献[10,27-35]	<ul style="list-style-type: none"> • 区块链使用户的数据信息存储在多个节点上,避免恶意攻击; • 用户拥有数据的完全控制权,可对敏感数据加密上链,仅授权方可解密,并且随时可终止授权; • 区块链上的用户通过账户地址参与交易,无法与真实身份对应; • 每笔交易信息都记录在区块链上,保护数据不被滥用
数据共享	文献[36-39]	<ul style="list-style-type: none"> • 区块链上的各节点可自由交易,无须中介机构,通过加密技术对交易细节加密,结合智能合约保证交易高效安全完成; • 通过区块链上的智能合约预设规则,对共享数据的用户进行奖励分配
数据可信	文献[30,40-47]	<ul style="list-style-type: none"> • 通过区块链共识机制对用户提交的数据进行验证和记录; • 由于共识机制的作用,数据一旦通过验证并写入区块链后就难以篡改; • 通过区块链上的智能合约预设规则,对提供有效数据的用户和矿工给予代币激励
推荐透明度	文献[27,48-50]	<ul style="list-style-type: none"> • 用户、物品、用户评分等都存储在区块链上,计算物品分数的算法存储在智能合约中,用户可选择特定算法来计算物品分数; • 区块链与推荐系统深度融合,构建难以篡改的模型,跟踪推荐过程

攻击,并且可在未经用户授权的情况下,转卖或利用用户数据进行分析,存在泄露用户信息的风险。区块链对用户隐私和数据安全的保护主要体现在3个方面:一是区块链的去中心化数据存储可以避免恶意攻击,消除第三方的控制;二是在公有链中,每个用户拥有自己的账户地址,很难把账户地址与真实世界中的人的身份对应起来;三是通过各种加密技术对链上的数据进行加密,仅授权方可以解密。

区块链作为一个去中心化的数据库,可以有效避免集中式数据中心存在的分布式拒绝服务(distributed denial of service, DDoS)攻击和单点故障等安全风险,并且消除了第三方的控制,将数据控制权归还给用户,降低了中央数据存储导致的隐私泄露的风险。例如Lin L J等人^[33]提出一种基于区块链技术的隐私保护推荐机制,利用区块链固有的去中心化特性,用户能够将原始数据保存在本地,降低了中心数据存储导致的隐私泄露风险。该机制还结合了星际文件系统和差分隐私机制,降低了计算和通信成本,可为推荐过程提供隐私保护。

在上述文献中,部分研究采用公有链搭建推荐系统框架,利用账户地址的匿名性保护用户的隐私安全。例如杨立等人^[29]提出了一个基于区块链的保险产品推荐模型,所有参与者都通过模型中的账号参与计算,并且每次交易都使用一个新的账号。Lisi A等人^[48]提出一个基于公有链的推荐系统框架,用户可以选择一个不透露任何真实信息的虚拟账号进行注册。在去中心化的应用场景下,任何人都很难将虚拟账号与用户真实身份之间的匹配关系对应起来,这在一定程度上保证了用户隐私安全。然而,完全匿名系统虽然可以保护用户隐私,但这意味着允许未经身份验证的用户创建虚假账号和信息,为欺诈行为留下了可能性。

区块链与加密技术的结合,使用户对

自己的数据有一定的控制权,可选择性地展现个人数据给指定的参与方。例如杨立等人^[29]采用非对称加密的方式,将用户的个人信息(家庭住址、健康状况)和保险公司的退保理赔信息等敏感数据加密上链,从而保证仅被授权的平台可以获取相关信息并进行推荐计算,而其他平台或用户无法看到。赵子军等人^[28]利用分布式双门陷公钥加密系统(distributed two-trapdoor public-key cryptosystem, DT-PKC),对车辆的敏感数据(如目的地等)进行加密上链,并通过设计一系列安全协议进行推荐计算,以保护用户的隐私安全。Frey R等人^[10]提出将客户的个人资料和公司获得的购物历史、信用卡记录数据进行加密上链,当用户给予访问权限时,公司才有权执行推荐任务。区块链能够将用户信息进行有效加密,并做到精细化的数据授权,任何不相关的用户都无法从加密数据中读取到有用信息,这为用户隐私提供了更深层次的保护。

此外,通过区块链与安全多方计算的结合,可在保护隐私的情况下完成相应的推荐计算任务。例如Frey R等人^[10]利用基于区块链的安全多方计算的优势,提出了一个保护隐私的推荐系统。但实际上在区块链上完成整个推荐计算过程是不可行的,因为通过区块链上的智能合约来完成推荐计算任务需要消耗很大的算力,并且通常不能进行实时推荐。

因此,利用区块链技术固有的去中心化、加密和匿名性等特点,可以有效解决推荐系统的隐私和安全性问题。当然,这需要对权限设置、推荐算法和系统框架等进行进一步的优化,在充分发挥区块链的优点的同时,保证系统的可行性和合理性。

3.3.2 数据共享

推荐系统的数据共享问题主要包括两

个方面：一是平台间进行数据共享时存在的信任和隐私问题；二是由于缺乏有效激励，用户不愿主动分享更多高质量的数据。这些问题影响了推荐系统数据源的丰富度和真实性，降低了推荐精度。

针对推荐系统中不同平台间的数据共享问题，已有研究对基于区块链技术的解决方案进行了探讨。例如Yan B W等人^[37]提出通过在不同云平台之间建立联盟链，实现云平台之间数据的高效安全共享。通过采用非对称加密算法对共享数据进行加密，联盟链上的节点可以验证共享数据的完整性和可靠性，但没有密钥就无法获得任何信息；联盟链中的所有节点都不能随意篡改共享数据，这是因为一旦篡改就会被检测出来。陈亚辉等人^[36]提出一个以区块链技术为基础的金融大数据共享与流通的大体框架，打破跨界机构之间的数据孤岛，以实现面向金融用户的精准可信推荐。区块链的去中心化使得数据交易不需要第三方机构，利用区块链上的加密技术和智能合约，可以对交易细节进行加密，从而确保共享数据的隐私安全，保证交易和支付的高效完成。此外，平台间数据共享的交易一旦经过验证并添加至区块链上，就很难被修改或抹除，任何一笔交易都有完整的记录，这可有效解决不同平台间数据共享的信任问题，进一步促进平台间的数据共享。

此外，针对单个用户的数据分享问题，Li X L等人^[30]提出数据请求者在部署某项推荐服务的智能合约时，可在区块链中存入交易费用，通过区块链的交易记录和智能合约的规则设定，对数据提供者和矿工进行激励，以鼓励用户分享更多的数据。Lisi A等人^[38]提出一种基于区块链技术的去中心化评分框架，通过对用户评分进行奖励来激励用户提供评分数据，以支持推荐系统。由此可见，区块链可对推荐过程中用户反馈数据的使用全过程进行客观、

透明和难以篡改的跟踪和记录，这为实施精细化激励机制创造了条件。借助智能合约，可对数据的实际价值进行准确、公平的量化，从而对用户进行精准激励，促进高质量数据的供给。区块链为数据价值的准确量化和精细化激励机制的顺利实施提供了必要条件。然而，目前针对这方面的研究仍相对缺乏。

3.3.3 数据可信

在传统的中心化推荐系统中，用户数据由平台进行收集和维护，出于追求商业利益的目的，平台可能对用于推荐的数据进行操纵和篡改。另外，由于缺少有效的激励机制和对隐私保护的不足，用户也可能提供虚假的信息。针对推荐系统的数据可信问题，Li X L等人^[30]基于区块链设计了一个联合矩阵分解模型，采用基于实用拜占庭容错（practical Byzantine fault tolerance, PBFT）的共识机制对用户提供的数据进行验证，并根据用户数据贡献水平进行合理的奖励分配，以鼓励用户提供真实数据。Wang S等人^[44]通过区块链驱动的智能合约来实现众包的逻辑，构建用于推荐的知识图谱，图谱中的每个节点和边都不能被轻易操纵和篡改，保证了用于推荐的数据的透明度和可审核性。

基于共识的数据更新机制，区块链上的每笔交易由若干节点共同进行验证，在一定程度上保证了上链数据的真实可信；同时，激励机制的引入也可以鼓励用户提供更多真实有用的数据。另外，每个用户生成的交易由矿工验证，整个过程被记录在区块链上，具有难以篡改和可追溯的特性，任何人都无法随意篡改用户的反馈数据，这也降低了用于推荐计算的数据被恶意操纵的风险。然而，现有的研究主要还是聚焦在鼓励用

户提供有效反馈数据上,在如何利用区块链的分布式共识打击虚假评论等方面仍缺少相应的讨论。

3.3.4 推荐透明度

提高推荐透明度是提高用户对推荐结果信任度的重要途径,然而现有的中心化推荐系统均为黑箱,推荐透明度远远不足。在推荐透明度的问题上,Lisi A等人^[48]提出一个基于区块链的推荐系统构建方法,为用户提供透明的评分收集和物品分数计算过程。该方法将用户、物品、用户评分等都存储在区块链上,将计算物品分数的算法存储在智能合约中,用户可选择特定算法来计算物品分数。由于评分数据、计算函数都存储在区块链上,难以篡改且公开透明,用户可以查看用于推荐计算的数据和方法,验证物品分数的计算过程,从而提高推荐的透明度。然而,由于智能合约的计算能力有限,该方法中仅使用了简单的计算函数,这可能会导致较低的推荐精度。

随着推荐系统在日常生活和关键业务中的重要性不断提高,提高用户对数据、模型、训练过程和结果的信任度越来越重要。对于推荐系统的透明度问题,区块链能够帮助研究人员在不同粒度级别上对推荐过程进行跟踪和记录。这种清晰的追踪过程不仅可以提高推荐数据的可信度,还可以让所有参与的用户了解推荐系统做出决定的过程,提高对推荐系统模型、训练过程和结果的信任度。然而,将区块链与推荐系统结合时,需要对推荐算法和推荐流程进行进一步的设计和优化,在提高推荐透明度的同时,满足系统的精度和效率要求。此外,即使提高了推荐的透明度,对于非专业用户来说,理解推荐逻辑还是有一定难度的,因此,除

了使用区块链底层保证推荐流程的透明性,仍需在界面设计上保证与用户的良好沟通,降低用户的使用门槛。

3.4 评估指标

文献采用的评估指标见表2。从评估指标上看,采用实验方法的文献中,大多对区块链推荐系统的准确度进行分析。其中主要通过和现有方法的比较,来证明提出的推荐系统可以取得相似或更高的准确度;也有文献通过问卷调查和案例研究来分析推荐结果的合理性。如Wang S等人^[44]通过一个给企业员工分配任务的案例实验,分别从自我验证(与被推荐员工的历史工作是否相关)和外部验证(推荐结果被采纳的比例)两个方面分析推荐的有效性,并通过问卷调查的方式来收集企业领导对系统有效性的反馈。

此外,引入区块链作为底层技术,一方面旨在提高推荐系统的安全性能,另一方面也给推荐系统的可行性和稳定性带来了新的挑战。因此,为了确保系统的有效性,除了对推荐系统的准确度进行评估,还必须对区块链系统进行测试。表2中大部分文献对区块链系统的性能进行了分析,包括安全性能和效率性能。安全性能主要是在区块链框架下,通过定性方法对推荐系统的隐私性、透明性和可验证性等进行分析。效率性能指系统的运行效率和可扩展性,对效率性能的分析主要通过定量分析的方法对系统运行时的燃料消耗、计算开销、通信开销、存储可扩展性和吞吐量等进行分析。这说明了在区块链推荐系统中,推荐系统的推荐性能、区块链框架的安全和效率性能受到广泛关注,对相关指标的分析是证明系统有效性的关键。因此,在提出区块链推荐系统框架后,对二者的分析是非常有必要的。

表2 文献采用的评估指标

作者	推荐性能	安全性能	效率性能
杨立等人 ^[29]	推荐转化率	隐私性、透明性、可验证性、可信性 (定性)	存储可扩展性(定量)
赵子军等人 ^[28]	—	安全性(定性)	通信开销、计算开销(定量)
Casino F等人 ^[31]	预测准确度	隐私性	局部敏感哈希(locality sensitive hashing, LSH)方法的效率(定量)
Li X L等人 ^[30]	预测准确度	隐私性、安全性(定性)	—
Bosri R等人 ^[32]	—	隐私性、完整性、可验证性(定性)	计算开销(推荐生成时间, 定量)
Lisi A等人 ^[48]	—	—	吞吐量、燃料消耗、可扩展性(定量)
Wang S等人 ^[44]	推荐有效性(案例研究、 问卷调查)	—	—
Yan B W等人 ^[37]	预测准确度	隐私性、难以篡改(定性)	燃料消耗
董学文等人 ^[40]	预测准确度	—	吞吐量(定量)
Yeh T Y等人 ^[49]	预测准确度	—	燃料消耗
Lisi A等人 ^[38]	—	安全性(定性)	燃料消耗、吞吐量、计算延迟时间(定量)
Sridevi S等人 ^[46]	预测准确度	—	—
Yan B W等人 ^[39]	预测准确度	难以篡改、完整性、避免单点故障 (定性)	内存消耗、吞吐量、计算延迟时间(定量)
Lin L J等人 ^[33]	预测准确度	隐私性	—
Cai W H等人 ^[45]	预测准确度	—	—

4 总结与未来研究建议

本文通过结构化的文献回顾,对近年来推荐系统存在的主要问题和区块链技术带来的机遇进行归纳总结,并采用文献分析方法,从时间分布、文献类型、研究问题和评估指标4个层面,对推荐系统中区块链技术的应用研究进行分析和总结。分析结果表明,区块链推荐系统的研究热度逐年增加,区块链技术对于解决推荐系统的用户隐私和数据安全、数据共享、数据可信和推荐透明度问题有重要意义。然而,已有文献主要集中于解决推荐系统中的用户隐私和数据安全问题,在其他方面的研究相对缺乏。

通过对已有文献的梳理和总结,笔者发现了一些重要的问题有待未来进一步深

入地研究。在此,本文针对性地提出以下几个观点和建议。

第一,重视跨平台数据共享的研究。目前研究中针对跨平台数据共享问题的研究相对较少,制约跨平台数据共享的主要原因来源于平台间缺少信任:一方面,数据分享存在信息泄露、损害平台用户隐私的风险;另一方面,数据价值存在不确定性,难以确保数据交换的公平性。借助区块链技术解决不同平台间的信任问题,需要对推荐模型的多方协同改进模式进行合理的设计。同时,通过设计合理的激励机制来保证各平台参与模型协同改进的公平性,提高各平台参与的积极性。

第二,结合区块链提出精细化的数据激励机制。推荐系统的数据稀疏问题是该研究领域的一个经典难题。其中,缺少有效的激励机制是导致用户分享数据的积极性不高的主要原因。但在传统的中心化

推荐系统中,由于缺少可信的执行环境,精细化的激励机制难以实施。区块链技术可对用户所反馈数据的使用全过程进行客观的跟踪和记录,使得透明性和公平性得到保证,为实施精细化激励机制创造了条件。依据平台上部署的智能合约,用户获得的收益将由其提供的反馈数据的实际价值和长期使用效果直接决定,从而解决了用户与平台间的信任问题。结合区块链上的加密货币,对用户反馈数据的精细化激励可以有效刺激交易市场,在解决数据稀疏问题的同时,实现各参与方的共赢。

第三,重视区块链框架的可扩展性研究。目前大多数研究是在简化的网络环境下进行少量实验,对区块链框架的可扩展性研究不足。现有的区块链框架只能处理少量的链上交易,适用于小规模、对吞吐量要求较低的应用场景。然而,推荐系统被广泛应用于各大领域,特别是电商平台。这些应用场景可能需要大规模地处理交易数据,对系统的延迟和吞吐量有一定要求。此外,许多业务应用场景对系统性能有动态需求,例如可能在某个时间请求大量资源。解决可扩展性是区块链被广泛采用的重要前提。已有研究针对区块链框架的可扩展性提出了一些解决方案,如分片技术、侧链技术等。在推荐系统与区块链的结合中,也应重视区块链框架的可扩展性研究。

第四,推荐算法与智能合约的相互融合。考虑到智能合约的可编程特性,可将推荐算法嵌入智能合约中。推荐系统领域的算法研究已较为成熟,如协同过滤推荐等算法已被成功应用于众多电商平台,但这些算法都只能在中心化架构中运行。智能合约是运行于区块链之上的图灵完备编程语言,需要对传统推荐算法进行智能合约化改造,才能使其适用于区块链运行环境,并使推荐逻辑和数据使用过程透明化,以满足对数据使用过程跟踪溯源的前提。

第五,推荐算法与边缘计算技术的结合。推荐系统属于计算密集型应用,推荐模型的训练和推荐结果的生成需要基于用户提供的反馈数据进行大量的运算,与区块链融合之后会造成计算和存储成本上升,增加平台负担。边缘计算是近年来兴起的一种计算模型,利用具有计算资源的终端设备在本地处理海量数据和大规模运算,旨在解决集中式的云计算实时性不高、网络带宽不足、计算能耗大和数据安全保护不足等问题^[51]。因此,应基于区块链去中心化的结构特点,探索将推荐算法与边缘计算技术结合,使计算负荷被合理分配给区块链中的节点,提高基于区块链的推荐系统效率。

区块链的兴起为推荐系统带来了根本性的变革。从推荐系统和区块链技术实践的发展来看,区块链与推荐系统的结合必将成为重要的发展趋势。因此,通过对区块链推荐系统领域进行系统的文献回顾,有助于区块链推荐系统研究的发展,并更好地指导和服务实践。

参考文献:

- [1] SCHAFFER J B, KONSTAN J A, RIEDL J. E-commerce recommendation applications[J]. *Data Mining and Knowledge Discovery*, 2001, 5(1-2): 115-153.
- [2] JANNACH D, ZANKER M, FELFERNIG A, et al. *Recommender systems: an introduction*[M]. Cambridge: Cambridge University Press, 2010.
- [3] SMUTKUPT P, KRAIRIT D, ESICHAIKUL V. Mobile marketing: implications for marketing strategies[J]. *International Journal of Mobile Marketing*, 2010, 5(2):126-140.
- [4] CHEN P T, HSIEH H P. Personalized mobile advertising: its key attributes,

- trends, and social impact[J]. *Technological Forecasting and Social Change*, 2012, 79(3): 543–557.
- [5] NAJAFABADI M K, MOHAMED A, ONN C W. An impact of time and item influencer in collaborative filtering recommendations using graph-based model[J]. *Information Processing & Management*, 2019, 56(3): 526–540.
- [6] ADOMAVICIUS G, TUZHILIN A. Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2005, 17(6): 734–749.
- [7] FAN W Q, MA Y, LI Q, et al. Graph neural networks for social recommendation[C]// *Proceedings of the 19th World Wide Web Conference*. New York: ACM Press, 2019: 417–426.
- [8] MAN T, SHEN H W, JIN X L, et al. Cross-domain recommendation: an embedding and mapping approach[C]// *Proceedings of the 26th International Joint Conference on Artificial Intelligence*. California: International Joint Conferences on Artificial Intelligence Organization, 2017: 2464–2470.
- [9] VAN DYKE T P, MIDHA V, NEMATI H. The effect of consumer privacy empowerment on trust and privacy concerns in e-commerce[J]. *Electronic Markets*, 2007, 17(1): 68–81.
- [10] FREY R, WÖRNER D, ILIC A. Collaborative filtering on the blockchain: a secure recommender system for e-commerce[C]// *Proceedings of the 22nd Americas Conference on Information Systems*. Atlanta: AIS Electronic Library, 2016: 1–5.
- [11] GENTRY C. A fully homomorphic encryption scheme[D]. Palo Alto: Stanford University, 2009.
- [12] LINDELL Y, PINKAS B. Secure multiparty computation for privacy-preserving data mining[J]. *Journal of Privacy and Confidentiality*, 2009, 1(1): 59–98.
- [13] WANG X H, YANG X X, GUO L, et al. Exploiting social review-enhanced convolutional matrix factorization for social recommendation[J]. *IEEE Access*, 2019, 7: 82826–82837.
- [14] RESHMA R, AMBIKESH G, THILAGAM P S. Alleviating data sparsity and cold start in recommender systems using social behaviour[C]// *Proceedings of 2016 International Conference on Recent Trends in Information Technology*. Piscataway: IEEE Press, 2016: 1–8.
- [15] KHERN-AM-NUAI W, KANNAN K, GHASEMKHANI H. Extrinsic versus intrinsic rewards for contributing reviews in an online platform[J]. *Information Systems Research*, 2018, 29(4): 871–892.
- [16] 刘彦松, 夏琦, 李柱, 等. 基于区块链的链上数据安全共享体系研究[J]. *大数据*, 2020, 6(5): 92–105.
- LIU Y S, XIA Q, LI Z, et al. Research on secure data sharing system based on blockchain[J]. *Big Data Research*, 2020, 6(5): 92–105.
- [17] MOBASHER B, BURKE R, BHAUMIK R, et al. Toward trustworthy recommender systems[J]. *ACM Transactions on Internet Technology*, 2007, 7(4): 23.
- [18] WILLIAMS C A, MOBASHER B, BURKE R. Defending recommender systems: detection of profile injection attacks[J]. *Service Oriented Computing and Applications*, 2007, 1(3): 157–170.
- [19] TONG C, YIN X, LI J, et al. A shilling attack detector based on convolutional neural network for collaborative recommender system in social aware network[J]. *The Computer Journal*, 2018, 61(7): 949–958.
- [20] MEHTA B, HOFMANN T. A survey of attack-resistant collaborative filtering algorithms[J]. *IEEE Data Engineering Bulletin*, 2008, 31(2): 14–22.
- [21] YUAN F, YAO L N, BENATALLAH B.

- Adversarial collaborative neural network for robust recommendation[C]// Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: ACM Press, 2019: 1065–1068.
- [22] SINHA R, SWEARINGEN K. The role of transparency in recommender systems[C]// Proceedings of 2002 CHI Conference Extended Abstracts on Human Factors in Computing Systems. New York: ACM Press, 2002: 830–831.
- [23] ZYSKIND G, NATHAN O, PENTLAND A S. Decentralizing privacy: using blockchain to protect personal data[C]// Proceedings of 2015 IEEE Security and Privacy Workshops. Piscataway: IEEE Press, 2015: 180–184.
- [24] DENNIS R, OWENSON G. Rep on the roll: a peer to peer reputation system based on a rolling blockchain[J]. International Journal for Digital Society, 2016, 7(1): 1123–1134.
- [25] CARBONI D. Feedback based reputation on top of the bitcoin blockchain[J]. arXiv preprint, 2015, arXiv:1502.01504.
- [26] SCHAUB A, BAZIN R, HASAN O, et al. A trustless privacy-preserving reputation system[C]// Proceedings of IFIP International Conference on ICT Systems Security and Privacy Protection. Cham: Springer International Publishing, 2016: 398–411.
- [27] 杜兰, 朱叶, 田越, 等. 基于Fabric的图书推荐系统的设计与实现[J]. 新型工业化, 2020, 10(9): 50–54.
DU L, ZHU Y, TIAN Y, et al. Design and implementation of book recommendation system based on Fabric[J]. The Journal of New Industrialization, 2020, 10(9): 50–54.
- [28] 赵子军, 应作斌, 杨钊, 等. 结合区块链和车辆社交网络的车队成员推荐[J]. 西安电子科技大学学报, 2020, 47(5): 122–129.
ZHAO Z J, YING Z B, YANG Z, et al. Recommendation of platoon members by combining the blockchain and vehicular social network[J]. Journal of Xidian University, 2020, 47(5): 122–129.
- [29] 杨立, 左春, 梁赓. 基于区块链的保险产品推荐模型[J]. 计算机系统应用, 2019, 28(1): 61–68.
YANG L, ZUO C, LIANG G. Insurance product recommendation model based on blockchain[J]. Computer Systems & Applications, 2019, 28(1): 61–68.
- [30] LI X L, DU E X, CHEN C, et al. Blockchain-based credible and privacy-preserving QoS-aware web service recommendation[C]// Proceedings of 2019 International Conference on Blockchain and Trustworthy Systems. Singapore: Springer Singapore, 2019: 621–635.
- [31] CASINO F, PATSAKIS C. An efficient blockchain-based privacy-preserving collaborative filtering architecture[J]. IEEE Transactions on Engineering Management, 2020, 67(4): 1501–1513.
- [32] BOSRI R, RAHMAN M S, BHUIYAN M Z A, et al. Integrating blockchain with artificial intelligence for privacy-preserving recommender systems[J]. IEEE Transactions on Network Science and Engineering, 2021, 8(2): 1009–1018.
- [33] LIN L J, TIAN Y C, LIU Y. A blockchain-based privacy-preserving recommendation mechanism[C]// Proceedings of 2021 IEEE 5th International Conference on Cryptography, Security and Privacy. Piscataway: IEEE Press, 2021: 74–78.
- [34] UMEKWUDO J O, SHIM J. Blockchain technology for mobile applications recommendation systems[J]. The Journal of Society for e-Business Studies, 2019, 24(3): 129–142.
- [35] ARIF Y, NOPEMBER I T S, NURHAYATI H, et al. Blockchain-based data sharing for decentralized tourism destinations recommendation system[J]. International Journal of Intelligent Engineering and Systems, 2020, 13(6): 472–486.
- [36] 陈亚辉, 吴基成, 孙澜. 金融用户精准可信推荐研究[J]. 广西质量监督导报, 2019(11):

- 195–196.
- CHEN Y H, WU J C, SUN L. Research on accurate and credible recommendation for financial users[J]. Guangxi Quality Supervision Guide Periodical, 2019(11): 195–196.
- [37] YAN B W, YU J G, WANG Y, et al. Blockchain-based service recommendation supporting data sharing[C]// Proceedings of 2020 International Conference on Wireless Algorithms, Systems, and Applications. Cham: Springer International Publishing, 2020: 580–589.
- [38] LISI A, DE SALVE A, MORI P, et al. Rewarding reviews with tokens: an Ethereum-based approach[J]. Future Generation Computer Systems, 2021, 120: 36–54.
- [39] YAN B W, DONG A M, CHAI B B, et al. Blockchain-assisted collaborative service recommendation scheme with data sharing[J]. IEEE Access, 2021, 9: 40871–40883.
- [40] 董学文, 刘昊哲, 乔慧, 等. 支持冷启动用户推荐的区块链服务发布方案[J]. 通信学报, 2021, 42(1): 57–66.
- DONG X W, LIU H Z, QIAO H, et al. Blockchain-based service publishing scheme with recommendation for cold start users[J]. Journal on Communications, 2021, 42(1): 57–66.
- [41] 黎孟雄, 李杨. 基于区块链的教育资源智能分发平台研究[J]. 长沙大学学报, 2020, 34(2): 49–54.
- LI M X, LI Y. Research on intelligent distribution platform of education resource based on blockchain[J]. Journal of Changsha University, 2020, 34(2): 49–54.
- [42] ARORA M, CHOPRA A B, DIXIT V S. An approach to secure collaborative recommender system using artificial intelligence, deep learning, and blockchain[C]// Proceedings of 2018 Intelligent Communication, Control and Devices. Singapore: Springer Singapore, 2020: 483–495.
- [43] ABBAS K, AFAQ M, AHMED KHAN T, et al. A blockchain and machine learning-based drug supply chain management and recommendation system for smart pharmaceutical industry[J]. Electronics, 2020, 9(5): 852.
- [44] WANG S, HUANG C C, LI J J, et al. Decentralized construction of knowledge graphs for deep recommender systems based on blockchain-powered smart contracts[J]. IEEE Access, 2019, 7: 136951–136961.
- [45] CAI W H, DU X, XU J L. A personalized QoS prediction method for web services via blockchain-based matrix factorization[J]. Sensors, 2019, 19(12): 2749.
- [46] SRIDEVI S, KARPAGAM G R, KUMAR B V. Incorporating blockchain for semantic web service selection (SWSS) method[J]. Sādhanā, 2021, 46(2): 1–14.
- [47] PORKODI S, KESAVARAJA D. A trust-based recommender system built on IoT blockchain network with cognitive framework[M]// Recommender system with machine learning and artificial intelligence: practical tools and applications in medical, agricultural and other industries. Hoboken: John Wiley & Sons, Inc., 2020: 293–311.
- [48] LISI A, DE SALVE A, MORI P, et al. A smart contract based recommender system[C]// Proceedings of 2019 International Conference on the Economics of Grids, Clouds, Systems, and Services. Cham: Springer International Publishing, 2019: 29–42.
- [49] YEH T Y, KASHEF R. Trust-based collaborative filtering recommendation systems on the blockchain[J]. Advances in Internet of Things, 2020, 10(4): 37–56.
- [50] LISI A, DE SALVE A, MORI P, et al. Practical application and evaluation of atomic swaps for blockchain-based recommender systems[C]// Proceedings

of the 3rd International Conference on Blockchain Technology and Applications. New York: ACM Press, 2020: 67-74.

- [51] 施巍松, 张星洲, 王一帆, 等. 边缘计算: 现状与展望[J]. 计算机研究与发展, 2019, 56(1): 69-89.

SHI W S, ZHANG X Z, WANG Y F, et al. Edge computing: state-of-the-art and future directions[J]. Journal of Computer Research and Development, 2019, 56(1): 69-89.

作者简介



许小颖(1987-), 男, 博士, 华南理工大学工商管理学院副教授, 主要研究方向为推荐系统、机器学习、大数据分析 and 区块链应用。



陈熙(1996-), 女, 华南理工大学工商管理学院硕士生, 主要研究方向为推荐系统、区块链应用。



陈源(1998-), 男, 华南理工大学工商管理学院硕士生, 主要研究方向为推荐系统、区块链应用。



谢永靖(1993-), 男, 华南理工大学工商管理学院硕士生, 主要研究方向为管理信息系统、区块链应用。

收稿日期: 2021-06-01

通信作者: 谢永靖, xieyongjing2021@163.com

基金项目: 国家自然科学基金资助项目(No.72071083); 广东省自然科学基金资助项目(No.2021A1515012003); 广州市哲学社科规划2021年度课题(No.2021GZQN09)

Foundation Items: The National Natural Science Foundation of China (No.72071083), The Guangdong Natural Science Foundation (No.2021A1515012003), Guangzhou Philosophy and Social Science Planning Project 2021 (No.2021GZQN09)