

5G 雾计算环境下图书馆“个性环绕贴身式”服务技术方案*

张 凯 赵国甫 陈 沅

摘 要 随着 5G、雾计算和推荐系统技术的成熟,图书馆服务有望在数字图书馆的基础上再次升级。本文基于 5G 雾计算环境提出了一种“个性环绕贴身式”图书馆服务模式。首先,利用软件定义网络(SDN)技术对 5G 图书馆网络进行垂直结构和水平业务切片设计;其次,利用边缘计算(MEC)技术和雾模型进行图书馆网络边缘计算和雾计算设计;接着利用网络功能虚拟化(NFV)技术进行图书馆网络读者边缘设备的“雾化”设计;然后分别使用文献目录细分和人工智能技术对文献进行“颗粒化”设计;最后,利用个性推荐和数据库技术进行图书馆读者个性信息推送软件设计。文章探讨的是一种新的图书馆服务模式,可以为读者提供超高速、贴身环绕、个性化、精准的信息推送服务。图 10。表 2。参考文献 29。

关键词 图书馆服务 5G 雾计算 个性推荐

分类号 G251

A Technical Scheme of “Personal Surrounding and Close Fitting” for Library Service under 5G and Fog Computing

ZHANG Kai,ZHAO Guofu & CHEN Yuan

ABSTRACT

In the explosive growth of information today, how to use the more advanced information technology, quickly and accurately find valuable information for readers, improve the efficiency of reading literature, become urgent problems to be solved in the library. 5G has become a hot new technology and is expected to upgrade the library service. This research has certain theoretical significance and practical reference value. At present, there are three problems to be solved in library service: The first is speed. Nowadays, e-library services, especially multimedia services, are seriously stuck, and their network speed has not kept up with the needs of readers. Secondly, the popularity of mobile phones makes readers look forward higher to personal services. The third is personalized services. At present, almost all libraries are standardized “one to many” services, and their one-to-one personalized services are not strong. For this reason, the authors put forward a library service mode of “personalized around and close to the body” under 5G and fog

* 本文系国家自然科学基金项目“第四范式下数据新闻业务流程集成框架重构研究”(编号:15BXW012)的研究成果之一。(This article is an outcome of the project “Research on the Reconstruction of Data News Business Process Integration Framework under the Fourth Paradigm”(No. 15BXW012) supported by National Social Science Foundation of China.)

通信作者:张凯,Email:zhangkai@zuel.edu.cn,ORCID:0000-0001-7417-9929(Correspondence should be addressed to ZHANG Kai,Email:zhangkai@zuel.edu.cn,ORCID:0000-0001-7417-9929)

computing, which requires using 5G network, fog computing and personalized recommendation technology to provide readers with super high speed and “around and close to the body” personalized information push service, so that readers can enjoy the service, literature delivered to readers’ mobile phones, according to their personal preferences at any time and any place, even in high-speed mobile. This mode is actually a combination of “5G network + fog computing + personalized recommendation” and library services. 5G has the characteristics of ultra-high speed, large connection, ultra-high reliability, ultra-low time delay, etc. The combination of 5G network edge computing and fog computing technology will make the cloud computing service (“Literature Cloud”) of the electronic library become a fog computing environment (“Literature Fog”) closer to the reader end. The combination of the granulation method of library documents and the personalized recommendation technology is expected to make the service of the library become a “personalized” service. From the perspective of existing technology and application conditions, the service mode of “personalized surrounding and close fitting” library under 5G and fog computing is just a concept and assumption. It is a forward-looking and academic exploration of the library service mode, and what the author puts forward is just a technical framework. If this concept is implemented in every library, it needs to be designed in detail according to the specific situation. 10 figs. 2 tabs. 29 refs.

KEY WORDS

Library service. 5G. Fog computing. Personal recommendation.

0 引言

网络的发展日新月异,正在改变人们的工作、生活和学习。截至2020年6月,我国手机网民9.32亿,其中使用手机上网比例达99.2%^[1]。用户月均使用移动流量达7.2GB,为全球平均水平的1.2倍;网络视频用户7.73亿,占网民整体的85.6%^[2]。随着网速越来越快,手机在多媒体方面的应用将成为一种趋势。在此背景下,图书馆应加强移动服务工作,这样不仅可以吸引更多手机读者,也可促使其服务质量有较大的提升。

在2000版ISO 9000系列中,顾客满意度成为产品服务质量的度量标准^[3]。产品服务不再是批量化的公众服务,它由面向所有人群或针对某一个人群的服务转向只针对每一个人或某一个人的某种需求的个性化服务。另外,“懒人经济”正推动O2O(Online To Offline,在线离线/线上到线下)消费升级^[4],快递小哥、同城配送等个性服务已将送货上门落实到每一位客户^[5]。如果将这种思想理念用于图书馆服务,

即个性化的“送货上门”服务,那么图书馆的服务质量将会升级。图书馆的个性化服务即针对读者的一对一服务,图书馆“送货上门”服务可以理解为根据读者搜索历史挖掘个性化阅读需求,并将文献资料甚至更准确的知识推送送到读者手机上。

在信息爆炸性增长的今天,文献数量剧增,使读者可选资料范围扩大,但同时也使读者查阅知识的难度增加。如何利用信息技术快速而又准确地为读者找到有价值的信息,提高读者查阅文献的效率,成为图书馆迫切需要解决的问题。大数据、人工智能和5G已成为当今炙手可热的新技术,这些新技术使图书馆服务再次面临前所未有的挑战,同时也是其难得一遇的发展机会。5G的出现使边缘计算技术异军突起,它使计算机相关设备贴身读者服务成为可能;各种小型智能计算设备的普及,比如手机、交换机、路由器等,使环绕读者服务的能力提高;人工智能技术的快速发展使读者准确获取文献中所需知识点更容易、更便捷,而且还能个性化推送到读者手机上。5G雾计算环境下的图

图书馆服务研究具有较强实践应用价值和理论意义。从实践层面来看,对服务技术的探索很迫切,也很重要,将对图书馆技术服务能力的提升和完善起到积极的推动作用;从理论层面来看,服务方式的改变将推动图书馆服务理念的转变和服务价值的提升。

1 研究现状及存在的问题

1.1 相关技术及应用的研究现状

5G 及其应用研究早已开始并日益成为焦点。5G 网络速率将达 10Gbps,是 4G 的 25 倍,5G 网络切片技术已有相关研究介绍^[6]。网络功能虚拟化 (Network Functions Virtualization, NFV) 与软件定义网络 (Software Defined Network, SDN) 是实现 5G 网络切片的主要技术支撑^[7]。目前很多国家已经开始 5G 商用^[8],各领域的应用也在逐渐展开。

雾计算方面的应用已有学者开始研究,比如物联网与云雾计算融合网络框架^[9],云计算与边缘计算^[10],雾计算与虚拟化在线学习系统^[11]。“5G+云雾边缘计算”模式的应用也有学者讨论,比如边缘计算网络 and 平台应用场景^[12],5G 智能物联边缘计算^[13],5G 网络的增强现实和 VR 直播等场景应用^[14]。

关于推荐系统和信息个性推送研究,近年来有不少学者涉及,比如基于微信公众平台的图书馆个性阅读推送系统^[15],基于用户兴趣感知挖掘的个性化服务推送^[16],基于读者个性特征的信息推送服务^[17]。

5G 在图书馆应用方面的研究主要包括 5G 时代的图书馆变革与发展策略^[18],5G 时代智慧图书馆如何发展^[19],5G 环境下图书馆馆员能力提升探析^[20],5G 与智慧图书馆建设^[21]。另外,基于用户需求的图书馆智慧推荐系统研究^[22]也有学者涉足。

学界对 5G、雾计算、个性推荐技术及其在图书馆方面的应用进行了诸多探讨。已有研究成果从概念角度分析了 5G、雾计算的未来发展趋

势,但 5G 技术在图书馆领域的具体应用还尚未发现;个性推荐技术在图书馆服务方面已有涉及,但 5G、雾计算和个性推荐技术结合的应用还未出现,也没有 5G、雾计算和个性推荐技术相结合应用到图书馆中的研究。本文试图探究利用 5G、雾计算和个性推荐技术来改进图书馆的服务模式。

1.2 图书馆服务目前存在的问题

第一,速度问题。目前的数字图书馆服务“卡顿”现象严重,其网络速度已经跟不上读者的需求。读者在阅读电子文献时,特别是使用图书馆多媒体服务时受网速影响较大,用户较多时,这种现象更加明显。如果读者在超高速运动中(比如高铁上),图书馆服务“卡顿”现象更严重。在非 5G 网络下,这个问题很难得到彻底解决。

第二,手机贴身服务问题。图书馆“贴身性”服务有待提升。即使在电子服务模式(云计算)下,图书馆“远离”读者的现象依旧突出。这是因为目前图书馆服务多采用云计算和局域网技术,虽然支持手机图书馆,但与雾计算和 5G 边缘计算相比,其对读者的“贴身性”服务明显不足。如果是纸质文献为主的图书馆服务模式,其与读者的“距离”会更远。

第三,个性化服务问题。目前,几乎所有的图书馆都是“千面一律”标准化的“一对多”服务模式。尽管图书馆已经电子化,但其服务方式依然没有发生实质性变化,仍是“一个”图书馆对“多个”用户的服务模式,而不是一对一的个性化服务,针对性不强。

针对上述问题,本文基于 5G 网络、边缘计算、切片技术、云计算、雾计算等技术和个性推荐系统的属性特征,提出“个性环绕贴身式”图书馆服务模式,并依次探讨设计图书馆 5G 网络切片、图书馆网络边缘计算和雾计算、图书馆网络读者边缘设备的“雾化”、图书馆文献的“颗粒化”和图书馆读者个性信息推送软件等,构建为读者提供超高速、贴身环绕、个性化、精准的信息推送服务的技术方案。

2 研究基础

2.1 5G 网络、边缘计算和切片技术

2012 年我国启动 5G 移动通信研究。5G 通信具有广域覆盖、高容量、低功耗、大连接、超高速等特点。5G 网络切片将物理网络分割成多个虚拟网络,每个虚拟网络间逻辑独立、相互隔离,任何一个虚拟网络出现故障都不会对其他虚拟网络造成影响。边缘计算是 5G 网络的核心技术之一。在传输网络架构中,5G 计算功能由靠近中心的云计算向靠近接入侧(用户端)的边缘服务器转移,这样使用户边缘计算的能力增强,也使图书馆服务更靠近用户。5G 网络、边缘计算和切片技术有望为图书馆服务升级提供强有力的技术支撑。

2.2 云计算和雾计算技术

云计算具有强大的计算能力和存储能力,可为图书馆信息化建设提供强有力的技术支撑^[23]。但云服务器距离读者“较远”,“长距离”信息传输要占用大量网络带宽且会增加传输时延,降低读者工作或学习效率。

雾计算概念由思科公司于 2011 年提出^[24],引起了业内学者的广泛关注。雾计算命名来源于“雾是更接近地面的云”这一概念,它是一种比云计算更贴近用户、半虚拟化的服务计算框架模型。雾计算与云计算相比有以下优点^[25]:较低时延,定位本地网络边缘,分布式计算模式,服务器节点数量多,支持移动性。不像云计算那么“遥远”,雾计算的数据存储及处理更依赖本地设备。这些优点都注定了雾计算环境下的图书馆服务模式会优于云计算下的图书馆服务模式。

2.3 个性化推荐系统

进入网络传播时代,个人信息需求一直被服务业务置于首位。满足个人信息需求的方式有两种:一是通过客户端软件到服务器查询数据,称为“拉技术”;二是在手机与服务器之间建立一个 TCP 长

连接,设置一定的推送规则,服务器实时地将符合用户需求的数据推送到客户端,称为“推技术”。

推送系统始于 20 世纪 90 年代,由美国 Pointcast 公司提出,可根据客户信息进行挖掘,以发现客户的个性需求,实现针对用户的信息服务个性化推送。推送技术使用户“获取信息的费力程度降到最低”^[26],让用户能在最短的时间、用最少的费用和精力获取自己想要的信息。个性化推荐系统是基于海量数据挖掘基础上的一种智能平台,可帮助读者过滤不需要的信息,从海量数据中挖掘出有用的信息,并将其推荐给读者^[27]。个性化推送系统基于发现算法挖掘读者感兴趣的内容,将其与待预测信息进行匹配,通过对读者历史数据的学习模拟,并根据读者的偏好变化,智能地修正将要推送的信息。

3 图书馆“个性环绕贴身式”服务模式

3.1 图书馆服务模式的演化

站在图书馆为读者服务的视角看,从纸本图书馆到数字图书馆,解决了读者从“很远”的地方跑到图书馆查阅书刊资料,转而利用计算机到数字图书馆“近距离”查阅电子资料的问题。读者不需要“长距离”奔波去图书馆,其体力和时间被节省,查阅资料的速度也加快。从数字图书馆服务到 5G 雾计算环境下的图书馆服务,将要解决读者从数字图书馆查阅电子资料,转而由图书馆向读者实时个性推送所需信息的问题。信息与读者之间更加“亲近”,读者获取所需信息的速度进一步加快,时间被进一步节省。最重要的是,由图书馆“拉”信息(将电子资料下载到读者的计算机上)服务模式演变为图书馆“推”信息(将电子资料上载到读者的 5G 手机上)服务模式。图书馆服务模式出现颠覆性变化,将给读者带来全新的服务体验。

3.2 新服务模式的定义

5G 雾计算环境下“个性环绕贴身式”图书馆服务模式,将利用 5G 网络、雾计算和个性推荐技

术,为读者提供超高速、“环绕贴身”的个性信息推送服务,使读者在任何时间任何地点,甚至在高速移动中,均能享受图书馆根据读者个性偏好推送到手机上的信息服务。该服务模式实际上是“5G 网络+雾计算+个性推荐”三种技术与图书馆服务的结合。5G 具有超高速、大连接、超高可靠性、超低时延等特点,5G 网络边缘计算与雾计算技术的结合使数字图书馆的云计算服务(“图书云”)变成更靠近读者的雾计算环境(“图书雾”),图书馆文献颗粒化方法与个性推荐技术的结合使图书馆的服务成为“个性化”服务。

3.3 新服务模式的功能

(1)图书馆可以提供超高速文献查阅服务。在 5G 网络环境下,读者在高速运动中(即使在高铁上)都可以非常流畅地查阅图书馆的文献资料(包括超高清视频,甚至视频直播或多媒体交互服务)。

(2)图书馆可以提供“环绕贴身式”服务。5G 边缘计算和雾计算将“远端”云计算下移到离读者很近的雾计算中,此时,5G 边缘计算、雾服务器和雾设备群就可以向读者提供“贴身”和“环绕”的服务。

(3)图书馆可以提供一种个性化知识推送服务。图书馆的各种资料被“颗粒化”后,存储在核心网的“知识云”中。推荐系统会根据读者个人偏好,将知识云中的知识颗粒下载到雾计算环境(“知识雾”)中,然后有针对性地向读者

推送其所需的信息。

4 图书馆网络切片设计

针对 5G 雾计算环境下“个性环绕贴身式”图书馆服务功能需求,本节将进行新服务模式下图书馆网络总体结构的设计。垂直网络切片和水平网络切片设计是首要工作。垂直网络切片设计主要是针对 5G 网络与现有图书馆网络系统重组后,对整个图书馆网络系统结构进行设计;水平网络切片设计是对图书馆业务进行水平的 5G 网络设计,包括图书馆行政管理、基础工作和读者服务三个子切片。

4.1 垂直网络切片结构设计

新模式下图书馆网络总体结构设计需要软件定义网络(Software Defined Network,SDN)技术的支撑。SDN 是一种新的网络虚拟化架构,其核心技术 OpenFlow 将网络设备的控制功能与数据功能分离,网络控制功能作为应用程序独立运行在集中式逻辑控制器中,其目标是通过软件 API 使控制部分可编程,实现网络流量的灵活控制,使网络管理和操作智能化,为新模式下图书馆核心网络构建和应用创造良好环境。

在以太网和“知识云”环境下,5G 图书馆垂直网络切片架构如图 1 所示。它主要包括以下功能。①图书馆网络管理。它是 5G 网络切片的入口,可将图书馆各种业务需求转化为端到端

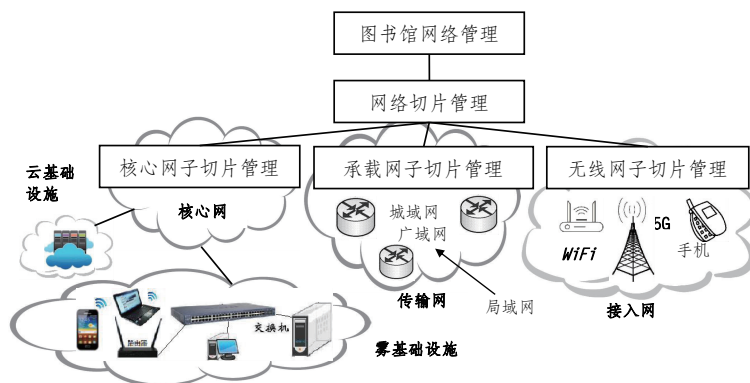


图 1 垂直网络切片架构

的网络切片需求,并下发到网络切片管理层。
②网络切片管理。当网络切片管理层收到图书馆网络管理层对网络切片的需求后,会产生一个切片实例,并将它转化成对网络子切片的需求,然后下发到网络子切片管理层。
③网络子切片管理。当网络子切片管理层接收到网络切片管理需求后,将其转换为网络子切片管理需求。网络子切片管理功能包括核心网子切片管理、承载网子切片管理和无线网子切片管理,这三个子功能将共同服务于图书馆各种业务需求。

4.2 水平网络业务切片设计

图书馆业务工作主要包括行政管理工作、

基础工作和读者服务工作三个方面,因此 5G 云计算环境下“个性环绕贴身式”图书馆服务模式的网络方案是一个三层 5G 切片架构。如图 2 所示,第一层为图书馆行政管理工作 5G 网络切片,第二层为图书馆基础工作 5G 网络切片,第三层为图书馆读者服务工作 5G 网络切片。数据中心(Data Center, DC)是 5G 网络切片的核心。这种分层切片只是本文的大致建议,每个图书馆可以根据自己的需要调整其层级和网络结构。5G 网络切片为图书馆网络提供了一种“按需分配”(即按图书馆业务和部门划分或重组需求)应用模式。

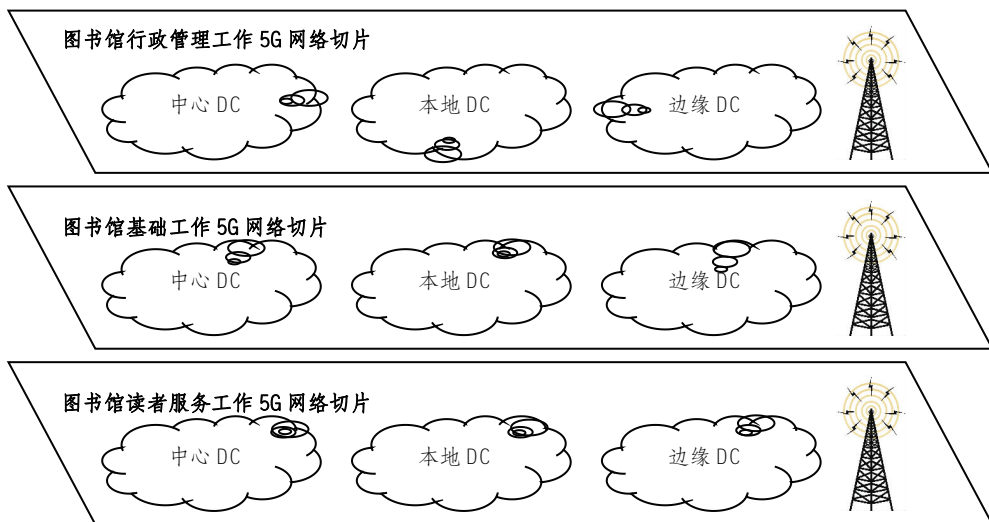


图 2 图书馆业务 5G 网络切片

将图书馆三方面业务工作具体化,行政管理工作包括人事管理、财务管理、设备管理、环境建设、规章制度五个方面;基础工作包括书刊采购、验收登记、分类、编目、贴标上架、总括登录等;读者服务工作包括外借、阅览、修补图书、新书宣传、书评、统计分析、读者咨询、编撰二次文献和读书活动组织等。具体的 5G 水平网络业务切片细化方案如图 3 所示。

除了图书馆的业务工作外,5G 还将为图书馆开创三大应用场景:第一是增强移动宽带

(Enhanced Mobile Broadband, eMBB),它是在现有移动宽带业务(比如 2G、3G 或 4G 等)场景基础上,提供一种大流量移动宽带业务,使读者使用超高清视频直播的体验有较大提升;第二是超可靠超低时延通信(ultra-Reliable and Low Latency Communications, uRLLC),它可以保证读者在通信过程中信息传输的超可靠和超低时延,手机读者无论打开多大的视频或文件都不可能出现“卡顿”现象;第三是海量机器类通信(Massive Machine Type of Communication, mMTC),它

主要针对大规模移动海量接入业务需求,可以 超大型图书馆的数字服务不会出现瞬间信息过
为瞬间接入的百万级读者提供服务,这将保证 载现象。

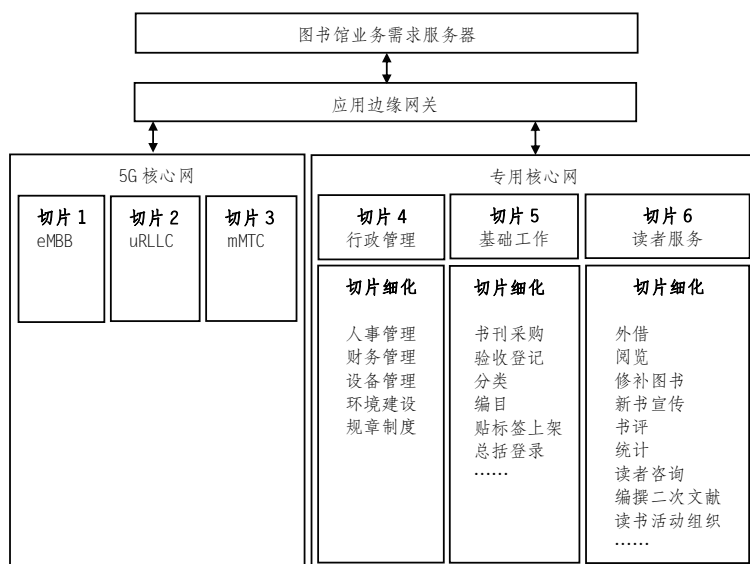


图3 5G水平网络业务切片细化方案

5 图书馆网络边缘计算和雾计算设计

为实现5G雾计算环境下图书馆“个性环绕贴身式”服务,需要进行5G网络边缘计算和雾计算设计。5G网络边缘计算技术的应用将使图书馆的服务更“贴近”读者,服务速度更快;雾计算技术的应用将使图书馆的服务可以“环绕”读者。

5.1 网络边缘计算设计

边缘计算是5G网络核心技术之一。5G移动边缘计算(Mobile Edge Computing, MEC)可以利用无线接入网络就近为读者提供信息服务和云计算服务,同时提高图书馆各项服务和应用的下载速度。MEC可以将远端“图书云”的计算功能下移到近端读者边缘,这时持有移动设备的读者可以请求接入网内虚拟资源,使源地址和目的地址之间的通信链路变短,减少网络信息传播量,降低端到端的时延,使通信效率提高。边缘计算是通信技术和计算技术的融合,

打破了传统计算和通信技术间的阻隔。在传输网络架构中,5G边缘计算将网络核心能力向下迁移到近端读者边缘,以满足读者超高带宽和超低时延需求,这使得网络计算如云如雾,包围读者,无处不在。它的计算功能由靠近中心的云计算向靠近读者的边缘服务器转移,使读者边缘的阅读需求计算能力增强(见图4)。

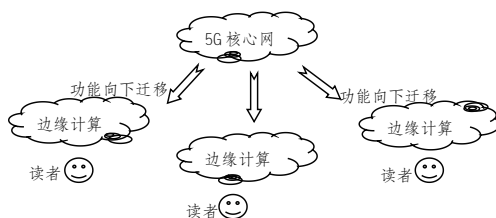


图4 5G网络边缘计算示意

5.2 网络雾计算设计

雾计算由性能较弱且布局分散的各类计算设备组成,是一种去中心化、分布式数据处理方式。它不需要将数据传输到云端,而是将数据计算、分析和处理功能在读者近端完成,降低数

据通过云计算处理的响应延迟,并降低读者终端设备能耗,从而延长终端设备的待机时间,甚至可降低网络传输数据的安全风险,很好地改善读者的体验(见图5)。

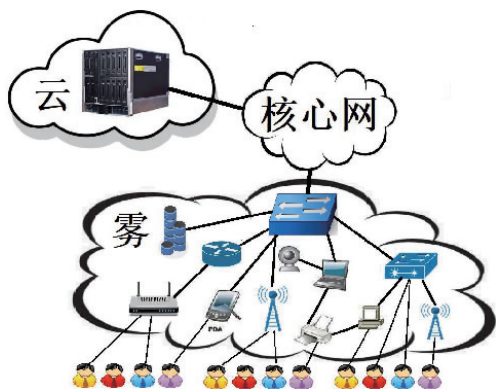


图5 雾计算服务器群

5G网络的分布式功能使个性智能推荐功能增强,同时雾计算所具备的泛在计算技术可对读者端庞大数据量进行雾计算并就近处理,使大量终端设备实现高效协同工作。读者在查阅文献时,希望尽快且准确找到自己想要的信息,这就需要图书馆设备有较高的运行效率。雾计算具有较强的分布式存储能力和计算能力,一方面,能够在终端设备附近进行本地雾计算,及时查询到读者需要的信息,主动营造“知识环绕”的场景氛围和效果;另一方面,也能将“图书云”端信息

通过核心网提前传输到雾计算服务环境中。

6 图书馆网络读者边缘设备的“雾化”设计

图书馆服务的“雾化”作为一种理念和模式,要想落实到具体应用中并最终实现“个性环绕贴身式”服务,必须首先对服务器进行“雾化”设计,然后在雾服务器中响应读者需求,最终实现任务分解与结果推送功能。

6.1 边缘设备“雾化”整体设计

网络功能虚拟化(Network Functions Virtualization, NFV)作为一种网络架构,是5G基础设施的关键技术,它利用虚拟化技术将网络节点切割成几个功能区块,功能区块分别通过软件方式实现,不受网络硬件架构限制。NFV将图书馆“个性环绕贴身式”服务功能从专用硬件平台剥离到通用服务器上运行,或在服务器虚拟机上运行,把应用程序、业务流程与基础设施用软件手段整合起来,使之成为灵活可编程的网络,由此实现网络功能虚拟化。

图书馆边缘设备的“雾化”,需要在读者和云计算之间增加一个“雾化”层,利用离读者最近的“雾化”层设备提供计算资源(见图6)。读者端智能路由器、交换机等都可以成为雾设备,甚至读者的手机、iPad、计算机等也可以作为雾设备,其部分功能可被雾服务器调用。

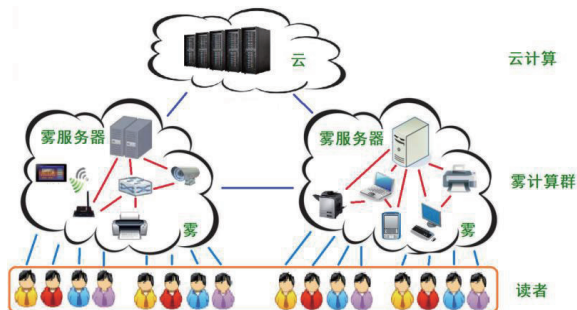


图6 读者边缘设备的“雾化”

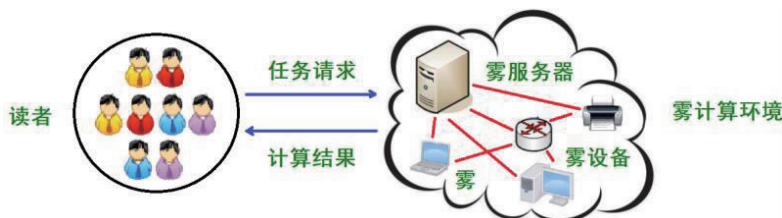
6.2 读者需求与雾服务器响应设计

雾计算环境包括雾服务器和被雾服务器管

理的雾设备(见图7)。雾服务器作为雾计算环境的本地协调器,接受雾计算任务请求,协调分

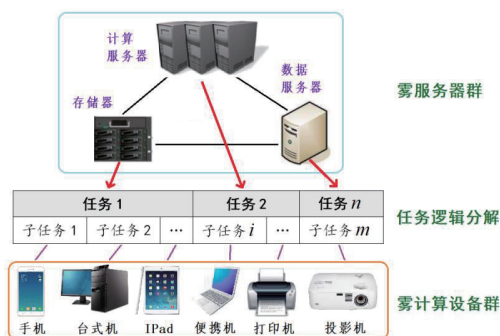
配计算资源,但其不承担具体的计算任务。在雾服务器的统一管理下,读者向雾服务器发送

阅读请求,雾服务器为阅读请求分配计算资源,并协同完成阅读请求的计算。



6.3 任务分解与结果推送设计

5G 网络边缘的雾设备负责为读者端提供计算服务。因为单个雾设备的计算能力有限,读者的阅读请求计算任务会被划分为多个子任务,然后被提交给雾服务器^[28]。雾服务器根据子任务的性质从本地服务设备列表中选取合适的雾计算设备予以响应(见图8)。雾计算设备完成计算任务后,向读者推送计算结果。



根据读者需求,有针对性地提供“最小化”的信息服务,以缓解读者的阅读疲劳。

为了区分馆藏文献“知识单元”的规模 and 大小,现给出知识单元的粒度定义,即知识块、知识点、知识颗粒、知识子粒、知识微粒等。知识块最大,其他知识单元的规模依次减小。

图书馆资源包括图书、期刊、报纸、统计数据、专利、标准、多媒体资源、百科/参考工具等及其对应的数字产品^[29]。需要“颗粒化”的文献通常是规模、体量、内容相对较大或知识单元很难定位的文献,比如书籍、研究报告、影像视频资料以及体量较大的论文(期刊论文、学位论文、会议论文)等。读者在查阅这类文献资料时,大多数以全书、整个报告、整篇论文、整个视频影像资料的形式提供。比较而言,知网中的期刊论文、硕博学位论文“小块化”“颗粒化”程度相对较好,其可以进行三级目录查询,甚至部分文献可以做到知识元检索。但目前大多数图书馆存储的电子书籍,只提供整本电子书的文献传递或查阅,并最多提供到三级目录导航。显然,这种文献服务太“粗糙”,难以满足读者所需内容的“精准”定位。而知识单元很难定位的文献则需要用智能手段解决。

一般来说,读者在查阅文献时可能只需要一部分章节,一小段内容,甚至是一句话,因此,希望图书馆能提供“精准”内容,即要查阅的内容在文献的哪一段落并高亮显示出来,或只精准呈现读者需要的内容,而不是把全文交给读者让其自己去查找。如果图书馆只是全文呈

7 图书馆文献的“颗粒化”设计

为了精准地向读者推送信息,除了网络设备的“雾化”,图书馆文献也要实现“颗粒化”设计。

7.1 图书馆文献“颗粒化”方法

图书馆文献“颗粒化”是指馆藏文献的“模块化”“小块化”“单元化”等工作。信息爆炸容易使读者阅读疲劳,图书馆文献“颗粒化”可以

现,读者就只能自己在整个文献中慢慢查找所需内容,这会使读者的查阅负担加重,白白浪费掉大量宝贵时间。解决这个问题比较好的方法是将整本电子文献通过文本化的方式划分成一个个知识单元,不能文本化的图片、视频、音频也要以知识单元为单位进行“颗粒化”。具体方法有两个:一是增加电子文献目录的层级,例如,书籍的目录至少要到三级、四级、五级,最好细化到最末级;二是在云雾服务器上部署“知识元检索”服务,按读者检索请求在符合条件的所有文献中进行知识元检索,并将得到的知识推送到读者端,这样读者就能非常快地找到所需内容。

我们首推在云雾服务器上部署“知识元检索”服务,这样可以一劳永逸地解决知识“颗粒化”问题。但是,就目前技术看,由于“知识元检索”算法较复杂,成本也比较高,完全做到“知识元检索”还有一个漫长的过程。因此,图书馆文献目录细化工作是当前比较可行的文献“颗粒化”方式。另外,人工智能颗粒化文献也是图书馆文献服务值得尝试的一种方法。

7.2 图书馆文献目录细化方法

规模体量较大的书籍、研究报告、音视频资料等形式的文献,因为知识点较多,其“颗粒化”的工作就是章节目录细化到最末级。以《法学概论》一书为例,目录共5级,内容如下:

第1章 法学基础理论

1.1 法律概述

1.1.1 法、法律的概念

1.1.2 法律的基本特征

1.1.3 法律的作用

1.1.4 法律的分类

(一)成文法和不成文法

(二)实体法和程序法

(三)根本法和普通法

(四)一般法和特别法

(五)国内法和国际法

(六)公法与私法:①公法;②私法。

目前图书馆的电子文献有两类:一是文本格式的电子文献,这类文献的编目细化很简单;二是图像格式的电子文献,它们是纸质文献电子扫描的结果。如果图像格式的电子文献未按章节目录细分到最末级,其章节细化编目的工作量将会非常大。

只有面对“颗粒化”的文献,读者在使用计算机检索、查阅所需信息时才更有针对性,也更容易准确地将读者所需信息“精准”推送到读者终端设备上。如果某个图书馆所有图像格式的电子文献都被“颗粒化”,即目录细化到最末级,读者查阅文献才会非常方便。

对图像格式电子文献的“颗粒化”工作是一项非常有意义但耗资很大的工程,显然这项工作不能用纯人工方式去完成,如果采用计算机图像汉字识别方法,加上人工智能的手段,将是一个值得研究的方向。

现假设图像格式电子文献已经完成三级、四级甚至五级目录细化工作,其编目如表1所示。

表1 目录级别表

一级目录	二级目录	三级目录	四级目录	五级目录
第1章				
	1.1			
		1.1.1		
			(一)	
				①

论文形式的文献,如果其规模和体量较大,包含的知识点或知识颗粒较多,也可以进行类似的细化工作;音视频资料的目录细化工作与此类似。

7.3 图书馆文献智能颗粒化设计

图书馆电子文献的形式多种多样,比如文本、图像、音频、视频、多媒体等。人工智能颗粒化文献指采用人工智能的方法,通过寻找、识

别、定位、提取、摘要等手段对文献进行颗粒化处理,以便得到读者所要查找的内容。目前,这类技术手段比较多,现将其分为两类讨论:一类针对文本形式的文献资料,另一类针对图像、音频、视频或多媒体形式的文献资料。

对于文本形式文献资料的知识颗粒化,可采用以下三种技术。①自然语言处理技术。即利用人工智能中的自然语言理解方法来判断文献中的哪一段最有可能是读者需要的知识颗粒。具体方法是利用读者输入的关键词或机器提取的读者偏好,通过文本语义分析获取文献段落的主题观点,进而判断该段落是否为读者寻找的知识颗粒。当判断文献的某一段落最有可能是读者需要的知识颗粒时,计算机程序将其标识在文献中的位置(起止地址),并将其存入雾端数据库,以便为读者推荐做准备。比如,百度智能云自然语言处理NLP开源软件工具包就具有情感倾向分析、对话情绪识别、文章分类、观点抽取、文章标签、新闻摘要、地址识别等方面的功能。②信息提取技术。有些文献可能包含读者需要的知识颗粒,而有些软件具有提取知识颗粒的能力,比如“应用信息提取V1.0.1”软件可以提取图片、版本号、版本名、信息摘要等;“窗口信息提取器V1.0”可以提取网页窗口中的文本信息、句柄、类名等。当知识颗粒被提取后可以存入读者雾端数据库,供读者阅读时取用。③文献摘要技术。目前有些软件(比如“超级自动摘要v1.3”软件、Auto-Ab免费自动摘要工具)已经具备了为文献自动撰写摘要的功能,这类软件可以根据读者需要及文献的信息内容自动为读者撰写所需的知识颗粒,然后将生成的知识颗粒存入读者雾端数据库。

对于图像、音频、视频和多媒体四种形式文献资料的知识颗粒化,可以采用以下四种方法。①图像形式文献资料的颗粒化分为两类。一是含有文字的图像,只需利用图像文字识别软件将其转化为文字并判断其是否为读者需要的知识颗粒,这类技术已经成熟(比如知网图像汉字识别软件);二是图像的识别和匹配,其方法是

扫描图像,如果图像与读者期望一致,即匹配成功,说明其为读者需要的图像类型的知识颗粒,这类技术也很成熟(比如Google的图像识别技术)。这两种情况都很容易为读者标识知识颗粒在文献中的位置,为其知识颗粒的雾端数据库存储做准备。②音频形式文献资料的颗粒化分为两类。一是含有文字信息的语音,只需利用语言识别软件将其转化为文本文字,并扫描判断哪一段文字是读者所需的知识颗粒,这类技术已成熟(比如腾讯云智能语音文字识别软件);二是音频的识别和匹配,其方法是扫描音频信号,如果某一段音频与读者期望一致,即匹配成功,说明这一段音频是读者所需的音频类型的知识颗粒,这类技术也很成熟(比如Dual-Eyes音频同步软件可自动同步匹配视频的录音并剪辑)。这两种情况都很容易为读者标识信息的位置,为其知识颗粒的雾端数据库存储做准备。③视频形式文献资料的颗粒化。利用人工智能和模式识别原理的算法,通过智能分析模块对视频画面进行扫描和识别,然后切割扫描识别的段落(包括起止地址)并将其作为读者需要查找的知识颗粒。由于视频画面的复杂性,为读者精准识别和定位视频信息的技术有待进一步完善。目前视频人脸检测、跟踪和识别的技术相对成熟,比如KnowU视频人脸识别系统,可以为读者截取某人行踪的视频段。另外,视频中的语音提取和识别的技术也相对成熟,比如网易云视频内容识别软件。④多媒体形式文献资料的颗粒化。多媒体是多种媒体信息的混合,其信息的识别和定位相对难度更大,甚至无法满足读者的需求。

8 图书馆读者个性信息推送软件设计

本节首先设计“图书雾”端数据库,然后根据读者偏好进行雾端数据库和图书馆个性推荐系统及软件的设计。

8.1 “图书雾”端数据库设计

在计算机中,文献一旦有了类似于表1的目

录级别表,就很容易通过计算机编程的方式实现自动检索,定位到相应章、节、目、条等,甚至实现“知识块”“知识点”或“知识颗粒”的调用和显示。

文献“颗粒化”完成后,文献及其目录表均可以放置在核心网云端数据库中(见图9)。在雾计算环境中,“图书雾”是硬件和软件两个部分的集合。其中,硬件物理结构设计在前文已经完成,而软件部分则包括雾端数据库和个性推荐软件。雾端数据库结构主要是字段、字段类型和字段长度,字段包括主题名、作者、出版机构、日期、摘要、知识单元、数据库记录地址等。

以前文《法学概论》一书为例,雾端数据库字段的“知识单元”在进行文献粒度分级时规定,第五级目录中的公法或私法,可称为“知识颗粒”;第四级目录下“(一)成文法和不成文法”“(二)实体法和程序法”“(三)根本法和普通法”等内容,可称为“知识点”,知识颗粒的目

录级别比知识点的目录级别低一级;“1.1.4 法律的分类”属于第三级目录,可称其为知识块,它比知识点的目录级别高一级;文献知识中更大的粒度是书籍的章(第一级目录)和节(第二级目录)。由于书籍的章节包含的内容太多,它们不属于本文粒度细化的讨论范围,在此从略。如果存在第六级目录的内容,可称为“知识子粒”,如果存在第七级目录的内容,则可称为“知识微粒”,依次类推。意思完整的句子可以作为一个知识单元,比如某个概念的科学定义可以作为一个知识单元。

图书馆的文献类型很多,比如图书、期刊、报纸、统计数据、专利、标准、多媒体资源、百科/参考工具等。当目录细分方法不适用时,可以考虑使用知识颗粒智能识别和定位方法。使用该方法时,注意及时获取主题名、作者、出版机构、日期、摘要、知识单元(知识颗粒)、记录地址等信息,这将为雾端服务器存储做好准备工作。

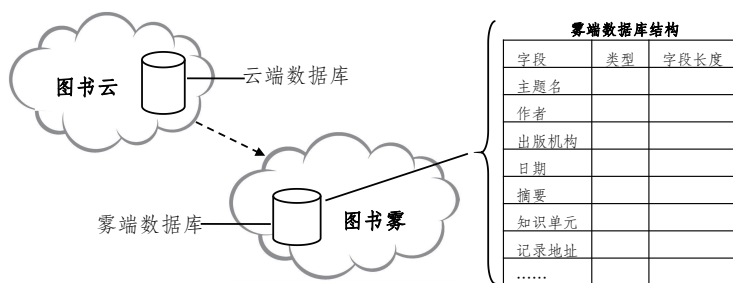


图9 “图书雾”及数据库

8.2 雾端数据库与读者偏好表

在读者的APP应用端,系统会自动识别读者的个性偏好。例如计算机一旦判断该读者想了解“公法”“知识单元”的含义,就会立刻将

“公法”“知识单元”的超级链接(核心网云端数据库中的记录地址)传输到雾计算环境设备的数据库中,并映射到读者偏好表中。读者偏好表结构如表2所示。

表2 读者偏好表结构

序号	读者IP地址	主题名	作者	出版机构	日期	摘要	知识单元	数据库记录地址	
1	xxx	xxx	xxx						
2	xxx	xxx	xxx						
3	xxx	xxx	xxx						

由于读者偏好列表的数据量较小,一旦读者与图书馆建立了联系(比如安装了图书馆软件),雾计算即可将“读者偏好表”的结构空表提前下载到读者的终端设备上(如手机、台式机或其他便携设备)。当数字图书馆有新的文献增加到核心网“图书云”中时,计算机系统马上根据读者偏好向其终端设备的“读者偏好列表”传输该文献的“知识单元”数据。这些数据就是图书馆个性推荐系统软件所要推送的素材。

8.3 图书馆个性推荐系统及其软件设计

图书馆个性推荐系统及其软件包括读者端和图书馆端,读者端为个性推荐 APP 软件,图书馆端为推荐系统(见图 10)。推荐系统包括硬件

和软件,硬件放置于雾计算环境的某个设备中,软件则包括推荐模块、读者信息收集模块和读者描述(画像)模块。读者信息收集模块负责收集读者的请求信息、点击行为信息、查阅行为信息,并将其存于访问日志,输入读者描述(画像)模块。然后,推荐模块根据读者描述(画像)模块汇聚的内容,提取读者兴趣特征需求,并与雾计算环境中“知识块”“知识点”或“知识颗粒”进行匹配,利用推荐算法(比如内容推荐算法、协同过滤推荐算法、关联规则推荐算法、知识推荐算法和组合推荐算法等)进行计算筛选,找到读者可能感兴趣的推荐对象(“知识单元”),然后将其推荐给读者。个性推荐 APP 软件的需要调用雾端数据库的数据。

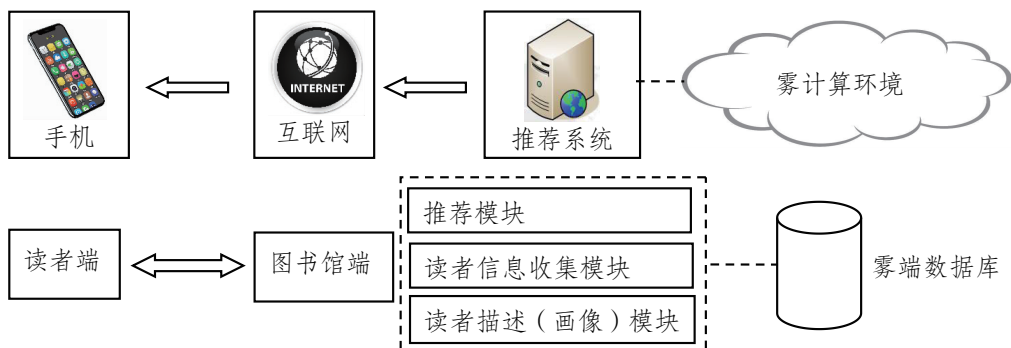


图 10 图书馆个性推荐系统

9 结语

综上所述,5G 及相关新技术的发展使图书馆服务面临前所未有的挑战和难得一遇的发展机会。5G 雾计算环境下“个性环绕贴身式”图书馆服务模式是“5G 网络+雾计算+个性推荐”三种技术与图书馆服务的结合,是图书馆服务模式的一种创新,它较当前图书馆服务在查阅速度、环绕贴身和个性化服务方面有比较大的优势。

2019 年 6 月我国政府批准 5G 商务应用,目前四大电信运营商正在紧锣密鼓地进行 5G 基站部署,各大企事业单位(包括图书馆)也在积

极规划 5G 应用。从目前的技术发展和应用条件看,本文提出的 5G 雾计算环境下“个性环绕贴身式”图书馆服务模式只是一种服务模式的前瞻构想和学术探索,或是一种技术框架构想。每个图书馆的具体落实尚需根据具体情况进行有针对性的详细设计。比如在 5G 网络部署时,如何与现有的局域网进行接口,现有的局域网是否需要升级,雾计算逻辑服务器如何划分和配置,边缘设备如何虚拟化,图书馆服务的个性推荐软件如何开发(联合开发,独立开发,还是外包开发),软件部署在哪个服务器上,等等,所有这些问题均需进一步在应用层面进行学术讨论,这将是我們下一步的研究工作。

参考文献

- [1] 中国互联网络信息中心. 第46次《中国互联网络发展状况统计报告》[R/OL]. [2020-10-01]. <https://www.cnnic.net.cn/hlwfyj/hlwzbg/hlwtjbg/202009/P020200929546215182514.pdf>. (China Internet Network Information Center(CNNIC). The 46th China statistical report on Internet development[R/OL]. [2020-10-01]. <https://www.cnnic.net.cn/hlwfyj/hlwzbg/hlwtjbg/202009/P020200929546215182514.pdf>.)
- [2] 中国互联网络信息中心. 第44次《中国互联网络发展状况统计报告》[R/OL]. [2020-10-01]. <https://www.cnnic.net.cn/hlwfyj/hlwzbg/hlwtjbg/201908/P020190830356787490958.pdf>. (China Internet Network Information Center(CNNIC). The 44th China statistical report on Internet development[R/OL]. [2020-10-01]. <https://www.cnnic.net.cn/hlwfyj/hlwzbg/hlwtjbg/201908/P020190830356787490958.pdf>.)
- [3] 钟庆龙. 六西格玛与ISO 9000质量管理体系的比较分析[D]. 广州:暨南大学,2003. (Zhong Qinglong. Comparative analysis of Six Sigma and ISO 9000 quality management systems [D]. Guangzhou:Jinan University,2003.)
- [4] 周琳,倪元锦,龚雯. 掘金“懒人经济” O2O 引领消费升级[J]. 中国中小企业,2017(11):45-47. (Zhou Lin,Ni Yuanjin,Gong Wen. Gold digging “lazy economy” O2O leads consumption upgrading [J]. China Small and Medium Enterprises,2017(11):45-47.)
- [5] 王晓平. “懒人经济”下同城配送问题研究[J]. 山西经济管理干部学院学报,2015,23(3):56-59. (Wang Xiaoping. Research on “lazy economy” city distribution questions [J]. Journal of Shanxi Institute of Economic Management,2015,23(3):56-59.)
- [6] 吴泉. 5G 通信技术应用场景及关键技术分析[J]. 网络安全技术与应用,2019(5):61-62. (Wu Quan. 5G communication technology application scenario and key technology analysis [J]. Network Security Technology and Application,2019(5):61-62.)
- [7] 安琪,刘艳萍,孙茜,等. 基于SDN与NFV的网络切片架构[J]. 电信科学,2016,32(11):119-126. (An Qi,Liu Yanping,Sun Xi,et al. Network slicing architecture based on SDN and NFV [J]. Telecommunications Science,2016,32(11):119-126.)
- [8] 徐湘禹,崔颖强. 解读5G网络技术研究现状和发展趋势[J]. 信息技术与信息化,2019(3):91-92,95. (Xu Xiangyu,Cui Yingqiang. Interpretation of 5G network technology research status and development trend [J]. Information Technology & Informatization,2019(3):91-92,95.)
- [9] 李树磊,樊毅,杜剑波. 基于混合云计算的物联网架构[J]. 邮电设计技术,2018(2):1-4. (Li Shulei,Fan Yi,Du Jianbo. The architecture of IOT based on mixed cloud/flog computing [J]. Designing Techniques of Posts and Telecommunications,2018(2):1-4.)
- [10] 崔佳. 云计算与边缘计算[J]. 电子技术与软件工程,2020(10):159-160. (Cui Jia. Cloud computing and edge computing [J]. Electronic Technology and Software Engineering,2020(10):159-160.)
- [11] 魏斌,曾青松. 基于雾计算与虚拟化技术的在线学习系统的设计与实现[J]. 福建电脑,2018,34(11):32-35. (Wei bin,Zeng Qingsong. Design and implementation of online learning system based on fog computing and virtualization technology [J]. Fujian Computer,2018,34(11):32-35.)
- [12] 耿鹏飞. 金山云曹亚孟:5G时代云计算与边缘计算将迎来新机遇[J]. 通信世界,2019(19):33. (Geng Pengfei. Jinshanyun Cao Yameng:cloud computing and edge computing will usher in new opportunities in 5G era [J]. Communication World,2019(19):33.)
- [13] 吴友文,林溪声,易景明. 5G边缘计算驱动智能物联新时代[J]. 新经济导刊,2019(2):22-27. (Wu Youwen,Lin Xisheng,Yi Jingming. 5G edge computing drives intelligent IOT new era [J]. New Economy Weekly,2019(2):22-27.)
- [14] 刘亚利. 5G网络中移动边缘计算的应用展望[J]. 中国新通信,2018,20(23):160-161. (Liu Yali. Application prospect of mobile edge computing in 5G network [J]. China New Communications,2018,20(23):160-161.)
- [15] 刘垵. 基于微信公众平台的图书馆个性阅读推送系统的设计与实现[D]. 武汉:湖北工业大学,2017. (Liu Yi. Design and implementation of library personalized reading push system based on WeChat public platform [D]. Wuhan:Hubei University of Technology,2017.)

- [16] 王福. 基于用户兴趣感知挖掘的个性化服务推送研究[J]. 图书馆理论与实践,2013(4):14-17. (Wang Fu. Research on personalized service push based on user interest awareness mining [J]. Library Theory and Practice,2013(4):14-17.)
- [17] 覃运芬,成云启,董坤. 基于读者个性特征的信息推送服务[J]. 内蒙古科技与经济,2010(15):158-159. (Qin Yunfen,Cheng Yunqi,Dong Kun. Information push service based on readers' personality [J]. Inner Mongolia Science Technology and Economy,2010(15):158-159.)
- [18] 李歌维. 5G 时代的图书馆变革与发展策略[J]. 图书与情报,2018(5):94-97. (Li Gewei. The reform and development strategy of library in the 5G era [J]. Library and Information,2018(5):94-97.)
- [19] 任广阔. 5G 时代下智慧图书馆的发展[J]. 信息通信,2019(6):98-100. (Ren Guangkuo. The development of intelligent library in the 5G era [J]. Information & Communication,2019(6):98-100.)
- [20] 刘小芳. 5G 环境下图书馆馆员能力提升探析[J]. 创新创业理论与实践,2019,2(12):175-176. (Liu Xiaofang. Analysis on the ability improvement of librarians in 5G environment [J]. The Theory and Practice of Innovation and Entrepreneurship,2019,2(12):175-176.)
- [21] 刘炜,陈晨,张磊. 5G 与智慧图书馆建设[J]. 中国图书馆学报,2019,45(5):42-50. (Liu Wei,Chen Chen,Zhang Lei. 5G and smart library construction [J]. Journal of Library Science in China,2019,45(5):42-50.)
- [22] 赵岩. 基于用户画像的数字图书馆智慧阅读推荐系统研究[J]. 图书馆学刊,2018(7):121-124. (Zhao Yan. Research on smart reading recommendation system of digital library based on user portrait [J]. Journal of Library Science,2018(7):121-124.)
- [23] 张红艳. 云计算平台下基于信息共享空间模式的图书馆管理技术研究[J]. 现代电子技术,2019(17):180-182,186. (Zhang Hongyan. Library management technology based on information sharing spatial pattern in cloud computing environment [J]. Modern Electronics Technique,2019(17):180-182,186.)
- [24] 吴磊. 思科:“雾计算”的 buzz 与干货[J]. IT 经理世界,2014(11):68-69. (Wu Lei. Cisco: buzz and dry goods of fog computing [J]. CEOCIO China,2014(11):68-69.)
- [25] 叶卓映,丰海. 云计算技术与雾计算技术的比较研究[J]. 通信与信息技术,2019(3):47-48. (Ye Zhuoying,Feng Hai. Comparative study on cloud computing and fog computing [J]. Communication and Information Technology,2019(3):47-48.)
- [26] 宋建武. 智能推送为何易陷入“内容下降的螺旋”——智能推送技术的认识误区[J]. 人民论坛,2018(17):116-119. (Song Jianwu. Why is intelligent push easy to fall into the “spiral of content decline”: the misunderstanding of intelligent push technology [J]. People's Forum,2018(17):116-119.)
- [27] 杨云,李文如. 基于 Mahout 的图书推荐系统研究[J]. 重庆科技学院学报(自然科学版),2019(4):98-102. (Yang Yun,Li Wenru. Book recommendation system based on Mahout [J]. Journal of Chongqing University of Science and Technology(Natural Science Edition),2019(4):98-102.)
- [28] 汤琳煜,蒋加伏,谷科. 基于雾计算的计算资源分配方案[J]. 计算机工程与应用,2019,55(19):96-104. (Tang Linyu,Jiang Jiafu,Gu Ke. Computing resource allocation scheme based on fog computing [J]. Computer Engineering and Applications,2019,55(19):96-104.)
- [29] 许莉. 基于特色专题的公共图书馆多媒体资源建设与服务策略研究[J]. 图书馆理论与实践,2019(10):15-19,49. (Xu Li. Research on the construction and service strategy of multimedia resources of public library based on special topics [J]. Library Theory and Practice,2019(10):15-19,49.)

张 凯 中南财经政法大学信息与安全工程学院教授,博士生导师。湖北 武汉 430073。

赵国甫 中南财经政法大学信息与安全工程学院博士研究生。湖北 武汉 430073。

陈 沅 中南财经政法大学信息与安全工程学院硕士研究生。湖北 武汉 430073。

(收稿日期:2019-09-23;修回日期:2020-10-23)