

大数据环境下基于分工协作的情报系统构建方法研究

张明宝 秦 琪

(南京航空航天大学经济与管理学院信息管理与电子商务研究所 南京 211106)

摘要: [研究目的] 将大数据技术引入构建服务于中小规模组织的情报系统时面临技术难度大、资源获取困难以及用户使用门槛高等困难。因此,提出一种大数据环境下基于分工协作的情报系统构建方法,该方法可以有效解决中小规模组织构建情报系统时所面临的这些问题。[研究方法] 通过对基于第三方的大数据环境下情报系统构建的三层模型和基于分工协作构建大数据情报系统的典型场景的阐述,提供了一种初步的可行技术解决方案,最后对基于分工协作构建大数据环境下的情报系统所面临的管理困难进行了探讨。[研究结论] 基于分工协作构建情报系统是大数据环境下情报学研究的一个重要方向,通过建立技术参考模型和典型应用场景可以为相关技术标准和实现方法的进一步确立提供参考。

关键词: 大数据; 社会化协作; 情报服务; 情报系统; 行业知识库

中图分类号: TP391

文献标识码: A

文章编号: 1002-1965(2022)02-0029-06

引用格式: 张明宝, 秦 琪. 大数据环境下基于分工协作的情报系统构建方法研究[J]. 情报杂志, 2022, 41(2): 29-34.

Research of Collaboration-based Information System Construction Methodology in a Big Data Environment

Zhang Mingbao Qin Qi

(Institute of Information Management and Electron Business, College of Economy and Management,
Nanjing University of Aeronautics and Astronautics, Nanjing 211106)

Abstract: [Research purpose] The introduction of big data technology into the construction of intelligence system serving small and medium-sized organizations is faced with great technical difficulties, resource acquisition difficulties and high user threshold difficulties. In this paper, we propose an intelligence system construction method based on division of labor and cooperation in the environment of big data. This method can effectively solve these problems faced by small and medium-sized organizations when they construct intelligence systems. [Research method] It provides a preliminary feasible technical solution by elaborating the three-tier model of intelligence system construction under big data environment and typical scenarios of building big data intelligence system based on division of labor and cooperation. Finally, it discusses the management difficulties faced by intelligence system construction under big data environment based on division of labor and cooperation. [Research conclusion] Building information system based on division of labor and cooperation is an important direction of information science research under the environment of big data. The establishment of technical reference model and typical application scenarios can provide reference for the further establishment of relevant technical standards and implementation methods.

Key words: big data; social collaboration; intelligence services; intelligence system; industry knowledge base

0 引言

大数据作为一种崭新的信息技术生态环境已经渗

透进各行各业的应用之中。大数据技术与情报学研究之间的关系很早便被关注。文献[1]和文献[2]分别从不同视角对大数据对企业竞争情报工作的影响进行

了分析,指出情报研究要与大数据结合;文献[3]则有针对性的探讨了大数据时代开展情报工作的策略和方法;文献[4]则明确提出在大数据时代发展情报系统需要变革和创新,并分析了可能的路径。文献[5]对大数据环境中的情报学发展方向进行了讨论。目前在情报学领域对大数据的研究主要还是局限于讨论大数据对改造传统情报工作模式的影响,探讨大数据作为未来情报学发展新方向的可行路径选择。具体的将大数据技术体系应用于构建新型情报系统的理论探讨还不充分。既有的基于大数据的情报系统研究工作主要围绕具体应用场景中的技术分析层面的问题展开,对新的基于大数据的情报服务应用模式的研究还很少见报道。

大数据应用以海量数据源和功能强大的大数据技术工具体系为基础,但是由于大数据的4V特征和大数据技术平台的高复杂性特征使得中小规模组织(企业)在导入大数据建设应用系统时面临巨大挑战^[6]。这使得在大数据环境中基于对新的服务与被服务之间关系的构造来实现易于建设和使用的大数据应用系统变得尤为关键。基于大数据构造情报服务系统同样面临技术和资源获取困难以及用户使用门槛高等困难,所以大数据时代的情报系统建构研究同样需要从用户需求 and 系统工作模式出发,通过寻求新的情报服务实现方式来重构情报系统建设的理念和方法。本文从这一思路出发提出在大数据环境中通过引入分工与协作的思想来构建情报服务系统,以实现既发挥大数据的技术优势又降低大数据情报系统构建难度的目标。论文对在大数据环境中基于分工协作构建情报系统的思想根源、技术参考模型和典型应用场景进行了系统阐述,并且对该方法体系的应用所可能面对的困难进行了分析。

1 大数据环境中基于分工协作构建情报系统的原因

在大数据环境中构建情报系统属于大数据技术应用的范畴,大数据技术应用模式的发展可以借鉴个人计算机、互联网和万维网等IT技术领域的成功经验。在这些技术领域,早期用户基于这些技术构建应用系统时同样普遍面临资源缺乏和进入门槛高的困难。为了解决这些问题,在IT技术领域逐步形成以标准为基础,通过专业化的社会分工和基于分工的有效协作来实现各种满足用户需求的应用系统的开发模式^[7]。个人计算机的成功在于引入了总线结构^[8]。基于总线标准,各种零部件供应商开发制造具有标准接口的零部件,用户根据需要灵活构建自己所需要的计算机系统,这才促成了个人计算机的广泛流行。互联网的成功应归功于ISO OSI标准和TCP/IP协议的制定^[9]。ISO

OSI标准对开放环境中的网络系统互联所需解决的问题进行了归纳和定义,有利于理论研究分工协作;TCP/IP协议则实现了各种网络设备的简单、低成本的互联互通。万维网的快速普及源自于其开放的技术体系和技术标准^[10]。它使得浏览器提供商、web服务器提供商、网页开发者、网站运营商和用户之间实现了有效的分工协作,共同实现了万维网中的各种类型的应用模式。

社会分工的不断深入是社会进步的必然结果,通过社会分工既可以有效使用资源,又可以极大提高产出效率。每一次新的社会分工都是在巨大社会需求的拉动和强有力的技术支撑推动之下产生的。大数据的发展正好处于这样一种大背景下,所以在大数据环境下构建情报系统仍然可以采用这一模式。将专业化分工与社会化协作的思想引入基于大数据的情报系统构建存在如下三个方面的现实考量。

首先,大数据的数据来源广泛,并且不再局限于某一公司或组织内部的信息系统,这就需要大数据情报系统与数据源提供者进行分工合作。随着社会性网络、移动互联网、物联网等的快速发展,这些领域的服务运营商累积的原始业务数据会成为大数据的主要数据来源,那么与这些运营商共享数据就成为大数据采集的一个无法回避的问题。另一方面,对于绝大多数需要基于大数据构建情报系统的组织来说,依靠其自身集聚大数据分析所需的巨量数据资源也是不现实的。

其次,大数据的成功关键在于数据分析,而大数据分析的复杂性对分析者的专业性提出了很高的要求。当将多样化的情报服务需求所催生的大数据分析任务完全交由某一组织来完成时,这往往会成为不可能完成的任务。大数据与传统的数据挖掘和智能分析最主要的区别在于:由于大数据的数据量巨大、数据类型多样、数据分析的实时性要求高等特征使得现有技术系统根本无法承载。技术难度的增加客观上要求各种不同类型的技术系统公司只有通力协作才可能实现既定的任务目标。今天,数据分析领域的技术外包已经非常普遍,大数据时代这样的外包与协作只会在更高的层面上更加频繁地出现。

第三,大数据分析的动态性和实时性,再加上用户需求的不确定又使得基于单一组织构建大数据情报系统困难重重。发端于上世纪90年代的供应链管理、客户关系管理、第三方物流等管理思想以及盛行于本世纪初的面向服务的架构SOA、网格、云计算等IT技术理念等无一不是借助于分工协作理念在实践中获得了成功。中小规模组织受制于自身的技术能力,在构建大数据情报系统时面对用户的不确定性需求自然也只

能遵循这一理念,通过社会分工以及基于分工的有效协作来快速聚集相关技术资源形成个性化的大数据服务能力。

2 基于分工协作的大数据情报系统的三层结构模型

图1展示了一种基于分工协作的大数据情报系统的三层结构模型,它可以为大数据环境中的情报系统构建提供思想指导。最底层为支撑层,主要包括网络基础设施建设、行业技术标准建立、相关法律法规设立以及人力资源建立和储备这几个方面的工作。这一层主要为大数据情报系统构建提供资源和基础设施支撑。中间层为业务层,主要描述基于分工协作构建大数据情报服务所涉及到的业务实体及其之间关系。该层最重要的实体是大数据情报服务提供者,它的职责在于面向用户个性化的情报服务需求,通过自身强大的综合情报服务能力整合各类情报服务资源快速形成所需的大数据情报服务。大数据情报服务提供者作为综合情报服务提供商只作自己最擅长的情报服务工作,其所欠缺的服务能力通过外包给其他服务提供商或者与其他服务提供商合作来实现。该层的数据资源提供者、计算资源提供者和数据分析技术提供者都是拥有某一特殊资源或能力的实体。它们通过封装其资源或能力对外提供服务,以此谋求自身的发展。最上层是资源层,这一层主要由各种大数据情报服务中介组成。随着分工协作思想的盛行,会出现数量庞大的专业服务提供商,如何帮助用户或者供需双方找到合适的合作伙伴就成为一个大问题。借助于这些中介可以有效整合大数据情报服务领域的各类资源,更好地为用户服务。

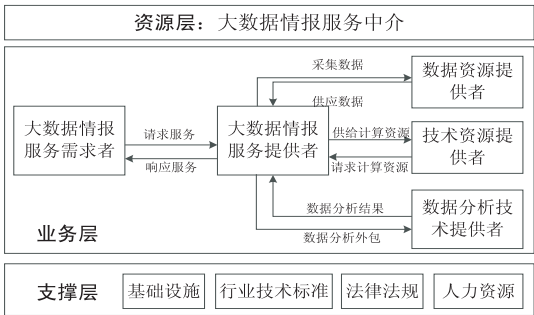


图1 基于第三方情报服务的大数据情报服务系统分层模型

按照图1所示的分层模型构建大数据情报系统的优势在于:a.通过分工协作,使得大数据情报服务工作分解落实到各类专业服务提供者,既使原本复杂的服务系统构建任务难度降低,使其实现成为可能,又可实现“让最合适的人做最适合他的事”这一目标,提高效率;b.通过社会分工,有利于大数据情报系统基础设施的构建,有利于整个社会大数据情报资源的发掘利用;c.通过分工协作,有利于大数据情报服务者能够快速

提供满足用户需求的个性化的情报服务,通过定制来解决大数据应用所面临的需求不确定的难题。

3 基于分工协作构建大数据情报系统的典型场景

大数据环境下通过分工协作构建情报服务系统需要首先产生一批专业化的情报服务实体(企业或者专门机构)。由于大数据处理和分析的复杂性,不同的情报服务实体将专注于不同的情报服务业务领域,譬如,有的情报服务实体专长在于发掘数据源并汇聚大数据,有的情报服务实体则专注于大数据的分析运算服务,有的情报服务实体则擅长于各类挖掘分析算法的设计等等。由于不同的情报服务实体在不同的业务层面提供服务,这就会产生不同的情报服务场景。设计并分析这些应用场景有利于挖掘出针对大数据情报服务系统构建的应用需求。图2至图6显示了五种基于分工协作的大数据情报系统服务的典型应用场景,每一图中的数字表明服务者与被动服务者之间采取交互活动的顺序编号。

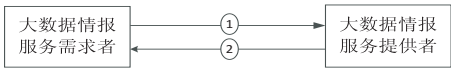


图2 确定环境下的简单情报服务应用场景

图2所示为确定环境下的简单情报服务场景。大数据服务提供者专门针对某一领域提供综合的情报挖掘和分析服务,大数据情报服务需求者对自身情报服务的需求和大数据情报服务提供者的能力都有明确的理解。服务过程如下:

第1步,大数据情报服务需求者向大数据情报服务提供者提出服务请求;

第2步,大数据情报服务提供者根据服务需求设计分析算法,进行数据处理,最后向大数据情报服务需求者传送分析结果。

这一场景产生于大数据情报服务需求者自身数据挖掘能力不足或者数据资源获取受限并且存在有针对性的大数据情报服务提供者的情况之下,这是一种最简单的情报服务方式。

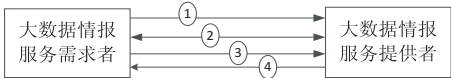


图3 确定环境下基于紧密协作的情报服务应用场景

图3所示为确定环境下基于紧密协作的情报服务场景。该场景不同于图2之处在于大数据情报服务提供者自身不具备数据分析所需的数据源,数据源需要大数据情报服务需求者提供。服务过程如下:

第1步,大数据情报服务需求者首先向大数据情报服务提供者提出服务请求;

第2步,大数据情报服务提供者响应服务请求并且通过二者之间的交互签订了服务契约;

第3步,大数据情报服务需求者向大数据情报服务提供者递送待分析的数据集合(一般是大数据情报服务需求者的业务相关数据);

第4步,大数据情报服务提供者根据服务需求和数据集合特征设计分析算法,进行数据处理,最后向大数据情报服务需求者传送分析结果。

这一场景的特点在于:大数据情报服务提供者需要大数据情报服务需求者提供数据源,由于二者之间的紧密合作关系可以实现深度的数据挖掘服务。这一场景实现的难点在于基于信任的协作关系的建立,因为大数据情报服务需求者向大数据情报服务提供者公开自身的数据资源会产生信息安全问题,如何既实现数据共享又不致信息泄露是一个关键问题。

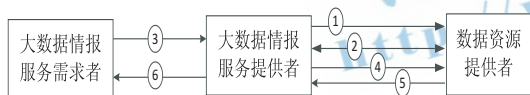


图4 确定环境下复杂情报服务应用场景

图4所示为确定性环境下的复杂情报服务应用场景。该场景不同于图3之处在于大数据情报服务提供者不使用大数据情报服务需求者的数据,而是使用数据资源提供者提供的各类海量数据资源。该场景的服务过程如下:

第1步,大数据情报服务提供者为了形成自己的情报服务能力主动寻找数据源;

第2步,大数据情报服务提供者通过和数据资源提供者之间的交互签订数据供给契约,形成完整的服务能力;

第3步,大数据情报服务需求者向大数据情报服务提供者提出服务请求;

第4步,大数据情报服务提供者根据服务需求者的服务需求向数据资源提供者提出数据需求;

第5步,数据资源提供者向大数据情报服务提供者提供待分析的数据集合;

第6步,大数据情报服务提供者进行数据处理,最后向大数据情报服务需求者传送分析结果。

这一场景的特点在于:大数据情报服务提供者和数据资源提供者通过合作形成完整的情报服务能力,大数据情报服务提供者不再需要大数据情报服务需求者提供数据源,这样大数据情报服务需求者和大数据情报服务提供者之间不需要进行紧密的协作。大数据情报服务提供者和数据资源提供者之间则可以通过契约将协作关系固定下来。

图5所示为不确定环境下的简单情报服务应用场景。该场景不同于前述场景之处在于大数据情报服务需求者和大数据情报服务提供者彼此相互不了解,需要借助大数据情报服务中介来相互发现。该场景的完

整服务过程如下:

第1步,大数据情报服务提供者描述自己的情报服务能力,并且按照大数据情报服务中介的要求向其提交自己的服务能力描述,由大数据情报服务中介公开发布;

第2步,大数据情报服务需求者在大数据情报服务中介上寻找满足自身服务需求的大数据情报服务提供者;

第3步,大数据情报服务中介提供检索服务,并且返回给大数据情报服务需求者有关的大数据情报服务提供者的服务能力信息;

第4步,大数据情报服务需求者和大数据情报服务提供者直接进行交互,并且签订服务契约;

第5步,大数据情报服务需求者向大数据情报服务提供者提出服务请求;

第6步,大数据情报服务提供者根据服务请求开展数据分析工作并返回情报服务结果。

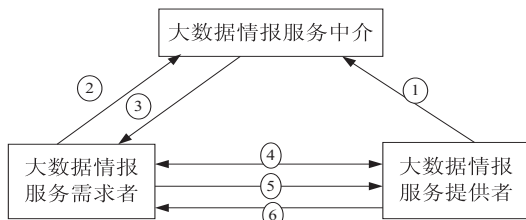


图5 不确定环境下的简单情报服务应用场景

这一场景的特点在于:大数据情报服务需求者和大数据情报服务提供者之间的服务与被服务的关系是借助于大数据情报服务中介动态建立的,这种关系是由服务需求驱动的,当请求者的服务得到满足以后,二者可以解除这种服务与被服务的关系。双方都可以继续借助于第三方的大数据情报服务中介去寻找新的合作伙伴,组建新的服务与被服务的关系。这一场景的优势在于情报服务的构建具有了敏捷性并且可以在更为广阔的空间中集聚情报服务资源,但是这一场景的实现高度依赖一个能够被各方所共信的第三方的大数据情报服务中介,并且它要求所有的参与者的情报系统可以直接对接交互。标准和规范的建立在此场景中具有关键性的作用。

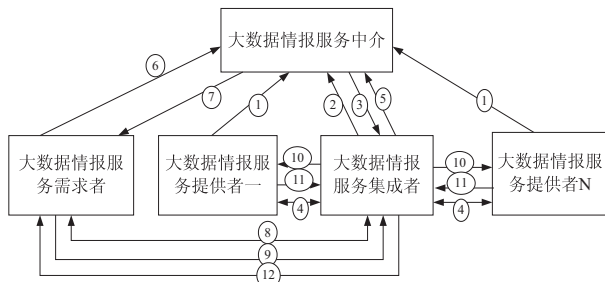


图6 不确定环境下基于服务复合的复杂情报服务应用场景

图6所示为不确定环境下基于服务复合的复杂情

报服务应用场景。该场景不同于图5之处在于大数据情报服务集成者作为一个完整情报服务能力提供者本身也需要通过大数据情报服务中介寻找自身所缺乏的服务能力,通过将上游服务提供商的服务能力集成形成自己的完整情报服务能力。其服务过程如下:

第1步,各个服务提供者(如图中的服务提供者一和服务提供者N等)描述自己的服务能力并将其服务能力描述递交大数据情报服务中介,通过该平台发布;

第2步,大数据情报服务集成者通过大数据情报服务中介寻找自身所需要的情报服务能力;

第3步,大数据情报服务中介通过服务能力匹配,将满足服务集成者需要的大数据情报服务提供者相关信息返回给大数据情报服务集成者;

第4步,大数据情报服务集成者直接和这些服务提供者交互,签订服务服务契约;

第5步,大数据情报服务集成者对自己通过聚合所形成的情报服务能力进行描述并且提交给大数据情报服务中介由其公开发布;

第6步,大数据情报服务需求者在大数据情报服务中介上寻找满足自身服务需求的大数据情报服务提供者;

第7步,大数据情报服务中介提供检索服务,并且返回给大数据情报服务需求者有关的大数据情报服务提供者(此处为大数据情报服务集成者)的服务能力信息;

第8步,大数据情报服务需求者和大数据情报服务集成者直接进行交互,并且签订服务契约;

第9步,大数据情报服务需求者向大数据情报服务集成者提出服务请求;

第10步,大数据情报服务集成者根据该服务请求向大数据情报服务提供者一和大数据情报服务提供者N提出服务请求;

第11步,大数据情报服务提供者一和大数据情报服务提供者N对大数据情报服务集成者的服务请求进行响应,提供大数据情报服务集成者所需要的服务能力;

第12步,大数据情报服务集成者响应大数据情报服务需求者,返回其所需要的情报服务内容。

这一场景的特点在于:大数据情报服务集成者可以借助于第三方的大数据服务中介发现自己所需要的服务能力,然后聚合这些服务商,通过构建情报服务供应链形成自己的综合情报服务能力。大数据情报服务提供商之间的聚合策略可以是“强强结合”,也可以是“优势互补”。它最能体现大数据环境下情报服务专业化和社会化的特征。这一场景实现的难度是最大的,它要求所有的大数据情报服务提供者在大数据情

报服务集成者的调控下实现同步的协作。

4 基于分工协作构建情报系统所面临的问题及其思考

推广应用基于协作分工构建情报系统的理念还面临诸多挑战,这些挑战的本质在于如何快速扩大服务者和用户的数量规模,所以相关的管理问题便成为进一步研究的关键。

4.1 服务提供者亟需构建专业化的情报服务能力

尽管理论分析的结果表明专业化的情报服务提供者的出现是大数据时代的必然趋势,但是,提供什么样的专业化情报服务能力确是每一个情报服务提供者的艰难的战略抉择。目前还很缺乏对这一类问题的深入探讨。传统的情报系统包括数据采集、传输、组织、存储、应用和再生等等工作内容,可以按照这些不同的工作内容进行分工协作。大数据情报系统也包含这些工作内容,但是大数据情报系统数据量大、数据类型多样、来源广泛、分析复杂、实时性要求高这些特点使得简单照搬传统情报系统的分工模型是行不通的,需要结合大数据自身特点、大数据技术条件和应用环境来研究大数据情报系统的分工业务模型,确定分工的标准和原则。譬如,需要研究确定具体场景中大数据情报系统需要哪些具体的参与实体?每一个实体的业务职能范围是什么?这些实体的对外服务能力是否是大数据时代所必需的?这些专业化服务的市场需求足够大吗?等等。

4.2 服务需求者寻求外部情报服务的观念转变

传统的自给自足的情报服务理念在大数据时代要彻底抛弃。大数据意味着更加专业化的社会分工,而这样的社会分工要能够存在并且发展壮大需要传统情报服务者改变观念,变自我服务为依托外部专业化服务。情报服务需求者要在观念上认知“利用外部资源”的价值,然后要通过自觉的自我改造来实现与外部资源的对接。情报服务需求者要学习如何区分可自我满足的情报需求和不可自我满足的情报需求;对于不可自我满足的情报需求要学会分析是否需要外部服务;对于需要外部服务的情报需求,要学会描述自己的需求,并且利用外部资源去有效的发现可用的外部资源,并且做好与外部情报服务对接的准备。从自给自足到依赖外部服务不是简单的服务能力的外化,而是基于对自我核心能力清楚认知基础上的对外协作。这种协作应该是组织情报服务能力达到一定水平时的自然选择。简言之,在社会化的第三方情报服务发展的同时,传统情报服务需求者也需要不断自我发展和修炼,以形成有效使用外部情报服务的能力。

4.3 情报服务关系的构建 当组织的情报需求、获

取和消费从组织内部扩展至组织外部之后,情报服务过程中所涉及到的各方之间的关系构建就变得尤为重要。情报服务者与被服务者之间的关系可以从管理方法、业务流程和系统接口三个层面来描述。管理方法涉及情报服务契约、服务过程跟踪与控制、服务计费策略、服务质量评价和服务策略等等内容;业务流程主要是指情报服务信息系统与企业情报服务信息系统之间的协作过程。

4.4 建立激励机制推动分工协作 基于分工协作的情报系统要有生命力,必须强调大众参与,强调发挥群体智慧进行相互协作。情报服务能力的构建和情报需求实现的外包都会改变用户应用情报系统的模式,这种改变必然会带来阻力,所以建立相应的激励机制鼓励参与就非常重要。

5 结 语

基于分工协作构建情报系统是大数据时代建设自动化和智能化的情报系统的一种必然趋势。基于分工协作构建情报系统既需要研究技术支撑环境又需要探讨应用模式和管理方法。本文建立的大数据环境中构建情报系统的分层模型和基于分工协作构建情报系统的典型场景为继续研究这些问题提供了初步的探索。基于分工协作的思想构建情报系统更为重要的是要在情报服务的全生命周期过程中转变所有参与者的传统

思维观念并且将面向服务的需求驱动作为系统设计的出发点,这无疑是最困难的。后继我们将通过原型系统的构建和使用来探讨解决这些问题的可能途径。

参 考 文 献

- [1] 吴金红,张 飞,鞠秀芳. 大数据:企业竞争情报的机遇、挑战及对策研究[J]. 情报杂志,2013,23(1):5-9.
- [2] 董 克,邱均平. 论大数据环境对情报学发展的影响[J]. 情报学报,2017,36(9):886-893.
- [3] 李广建,杨 林. 大数据视角下的情报研究与情报研究技术[J]. 图书与情报,2012(6):1-8.
- [4] 曾建勋,魏 来. 大数据时代的情报学变革[J]. 情报学报,2015,34(1):37-44.
- [5] 苏新宇. 大数据时代情报学与情报工作的回归[J]. 情报学报,2017,36(4):331-337.
- [6] 孟小峰,慈 祥. 大数据管理:概念、技术与挑战[J]. 计算机研究与发展,2013,50(1):146-169.
- [7] 张明宝,刘义俊,李 雨. 互联网环境下的非结构化数字信息组织与共享方法研究[J]. 情报学报,2012,31(10):1034-1043.
- [8] 姚莉娟,高 优. 浅述计算机总线技术的发展[J]. 数字通信世界,2018,4(5):64.
- [9] 周 倜. 计算机网络标准综述[J]. 航天标准化,2010,4(3):31-35.
- [10] 丁 玲. Web 标准与 Html5 重点技术探讨[J]. 科技与创新,2019,4(24):135-136.

(责编/校对:王平军)