南 京 大 学

期中课程论文

课程名：数据挖掘

授课教师：石进

学院：信息管理学院
成员：高　伦 171860544
成员：朱思成 171860559
时间：2020 年 4 月

目录

区块链研究热点及其变迁趋势.....2
——基于 WOS2016-2020 年的文献分析.....2
1 引言.....2
2 数据来源与研究方法.....3
2.1 数据来源.....3
2.2 研究工具.....3
3 研究过程与结果分析.....5
3.1 年度发文量.....5
3.2 学科分布.....6
3.3 核心引文.....7
3.4 研究热点.....8
3.4.1 主题词共现图谱聚类.....8
3.4.2 作者共被引.....9
3.4.3 研究热点总结.....9
3.5 研究热点变迁趋势.....9
3.5.1 时间线图谱.....9
3.5.2 突现词.....12
4 结论.....12
参考文献.....13
附录.....13

区块链研究热点及其变迁趋势

——基于 WOS2016-2020 年的文献分析

高 伦 171860544，朱思成 171860559
南京大学信息管理学院

摘要：本研究通过 CiteSpace、HistCite、文献计量在线分析平台三种可视化软件对 2016-2020 年间 WOS 核心数据库收录的 2294 篇区块链领域研究文章从摘要、关键词、作者、引用关系等角度绘制知识图谱并进行数据挖掘，分析发现区块链研究的研究热点、变迁趋势等。

关键词：区块链；研究热点；变迁趋势；WOS；CiteSpace

1 引言

区块链是分布式数据存储、点对点传输、共识机制、加密算法等计算机技术的新型应用模式。区块链是一种防篡改、共享的数字化账本，用于记录公有或私有对等网络中的交易，并由集体维护交易数据的分布式数据库。

2008 年，中本聪第一次提出了区块链的概念，在随后的几年中，区块链成为了电子货币比特币的核心组成部分。近年来，世界对比特币的态度起起落落，但作为比特币底层技术之一的区块链技术日益受到重视。区块链在金融、保险、物联网、公共服务、数字版权等诸多领域都得到了广泛的应用。2019 年 10 月 24 日，在中央政治局第十八次集体学习时，习近平总书记强调，“把区块链作为核心技术自主创新的重要突破口”“加快推动区块链技术和产业创新发展”。

在此背景下， 本文基于 Web of Science 数据库，对 2016-2020 年间国际区块链研究领域相关文献进行计量及对比分析，以期更好地把握该领域的研究热点、变迁趋势及规律。

2 数据来源与研究方法

2.1 数据来源

从开题目的出发，为了获取区块链领域的最新研究进展，本研究将文献检索范围限定在 2016-2020 年间的 Web of Science 核心数据库。检索规则为：TS=("blockchain" OR "block chain" OR "block-chain") AND PY=(2016-2020)。文献类型仅选取研究类文献 Article，最终得到文献 2294 篇。

获取方法：Github 上的 wos_crawler 工具是一个 wed of science 核心集合爬虫（链接：https://github.com/tomleung1996/wos_crawler）。该工具支持爬取任何合法高级检索式的检索结果，并将结果保存为 Plain text 格式，最后将爬取结果解析导入数据库，包含文献的基本信息（标题、摘要、关键词、被引量等）以及作者机构，分类，参考文献等信息。笔者通过使用此工具，获取了 5538 篇文献，再通过 citespace 的文献数据处理功能，分类出文献类型为 Article 的 2294 篇文献。

2.2 研究工具

2.2.1 CiteSpace

CiteSpace 是一种多元、分时、动态的指数图谱绘制工具，因其可视化的突出特点，受到了国内学者们的高度关注并得到了广泛应用，截止 2020 年 4 月，仅 CNKI 就有近 3800 篇文献应用了 CiteSpace 计量工具。

CiteSpace 的分析原理主要分以下几类。

其一，共被引分析（Co-citation analysis）。论文引用其他论文可以看作知识重组产生新知识的过程；而论文被引用可以看作这个过程的延续。学科领域就是在这样的引用和被引用中不断产生新的知识和突破，渐渐就会形成引文网络。1970 年，由美国情报学家 Small 和苏联情报学家 Marshakova 几乎同时提出了文献共被引的概念，它是指两篇及以上的文献同时被其他文献所引用，则这两篇文献构成共被引关系；同时引用这两篇文献的文献篇数被用来衡量共被引关系的强度，强度越大，表示这两篇文献的相似程度越大，关系越密切。实现共被引分析需要先从文献信息中归纳得到引证矩阵，在此基础上通过矩阵原理生成共被引矩阵，再通过图论的原理将共被引矩阵进行网络化、可视化。

其二，共词分析（Co-word analysis）。专业的学者，可以很好地提炼出最能代表其作品的关键词来反映其作品内容，同时，学者标注其作品关键词时也会受到领域内其他重要成果的影响而学习并使用相似的关键词来标注。词频是指某词语在文献中出现的次数，针对词频的文献分析可以通过从文献中提取词频较高的词语并结合关键词来研究领域热点和发展动向。共词分析是在词频分析的基础上进行更高层次的分析，其基本原理是对一组词两两统计其在同一组文献中出现的次数，以此衡量其关联程度。其基本步骤与共被引分析类似，先从文献中得到关键词信息，形成矩阵，在此矩阵基础上通过矩阵数学原理生成共词矩阵，最后

可视化，形成共词网络。关联程度高的多个词语的结点，便会形成更大的聚类点，便于研究人员对领域热点以及趋势把握分析。

其三，聚类分析。俗语说“物以类聚，人以群分”，自古以来，人类就知道依靠经验和知识来进行分类，聚类分析依靠更科学的数学算法将分类的过程定量化。聚类分析又称群分析，是以相似性为基础的分析方法。许多聚类算法都依靠欧几里得或者曼哈顿距离来度量点与点之间的相似性。在一个聚类中的点之间比不在同一聚类中的点之间具有更高的相似性。CiteSpace 的聚类算法的依据是谱聚类算法，这是一种基于图论的算法，会将样本看作结点，样本的相似度看作带权的边，如此一来，聚类的问题就会演变为图的分割问题，其对类似共引网络这种关注链接关系甚于结点本身的聚类要求有着天然的优势，这是因为一般的 K-Means 等聚类算法都是建立在凸球星样本空间上的，一旦样本控件非凸，算法结果就会只关注部分而失去对全局的把控。图论的方法很好地解决了这一问题，使得谱聚类能够识别任意形状的样本空间并收敛于全局最优解。

其四，高频与高中介中心性。高频，代表文献拥有较高的被引量，说明其是某个甚至多个领域的重要知识基础。中介中心性，指的是一个结点担任其他两个结点之间最短路的桥梁的次数。拥有高中介中心性的文献代表其与多篇文献形成共被引关系，其与多篇文献均有关联，是领域内的联系枢纽，也是领域内的重要文献。若一篇文献同时具有高频性和高中介中心性，则可以代表领域内该时期的研究热点。

本文基于 CiteSpace 知识图谱，对最终检索得到 2294 篇文献进行文献计量分析，包括共被引分析，共词分析等，并绘制知识图谱，分析区块链领域的研究热点及趋势，以了解近五年国际区块链技术研究的最新进展，掌握其发展的方向和趋势演变特征。

2.2.2 HistCite

HistCite 是一款功能强大的引文分析工具，可以快速绘制出某个研究领域发展的脉络，快速锁定领域的重要文献。

LCS（Local Citation Score）是指某一文献被本地数据集中的文献所引用的次数。导入 HistCite 的文献都与设置的检索词有关，因此可以认为这些文献都属于本文的研究方向，若一篇文献被本地数据集中的文献引用次数较高，则可以说明它在这个领域内的重要程度较高；若文献的 LCS 值在本地数据集中数一数二，则意味着此文献很可能是该领域的开山之作，因为之后其他文献都与其有关。

LCR（Local Cited References）是指本地被引参考文献数，也即某一篇文献的参考文献在本地数据集中的数量。LCR 值的大小可以作为衡量一篇文献是否为关注该领域的重要文献，因为一篇文献引用本地数据集中的文献数量越多就越能说明其十分关注本领域的研究动向，从该文献很可能找到领域发展的新方向。

在 Histcite 中导入区块链领域的本地数据集，得到文献的 LCS 值可以作为衡量文献在领域内的重要程度甚至开创程度；得到文献的 LCR 值可以作为发现文献在领域内发展的新方向的依据。本问通过该工具对 2294 篇文献进行引文网络分析，绘制出文献之间的引用关系网络，计算具有较高被引量的文献的 LCS 和 LCR，寻找区块链领域近五年重要程度较高的文献和发展方向更新颖的文献，并据此分析区块链领域的研究热点和变迁趋势。

2.2.3 文献计量在线分析平台

文献计量在线分析平台（https://bibliometric.com/），是一个集文献总量，合作关系，影响分析，关键词，引用关系等功能为一体的在线数据分析平台。该平台获得过 2013 年中国科学院国家科学图书馆“科研教育开放信息创新应用大赛”三等奖，其分析结果直观，图形美观大方。

该平台通过 web 端服务，以图形化的方式，协助研究人员对引文数据进行文献计量分析。简单，易上手，是其对比其他文献计量分析工具最大的特点，用户只需上传指定格式的数据集即可一键获得分析结果，从大体上把握领域的研究动向，避免了参数调节等复杂工序。

本文通过导入平台指定的 utf-8 格式的数据，分析了多个国家近五年在区块链领域发文的数量情况、国家间的合作关系、关键词的占比情况以及数量变化趋势，以求分析不同国家地域在区块链领域研究的热点问题和发展趋势。

3 研究过程与结果分析

3.1 年度发文量

论文的发表数量是衡量某领域学术发展水平和科研成果的重要指标。图 1 是 2016-2020 年各年份文献发表数量，从图中可以看出：在 2016 年，研究者对于区块链的研究较少，全年仅有 37 篇。而从 2017 年开始，区块链技术在实际应用领域取得了广泛的成功，文献发表数呈现迅猛增长的态势，在 2019 年达到一个极大值 1227 篇。鉴于本文撰写时间（2020 年 4 月），2020 年份发文量仅有 400 余篇。

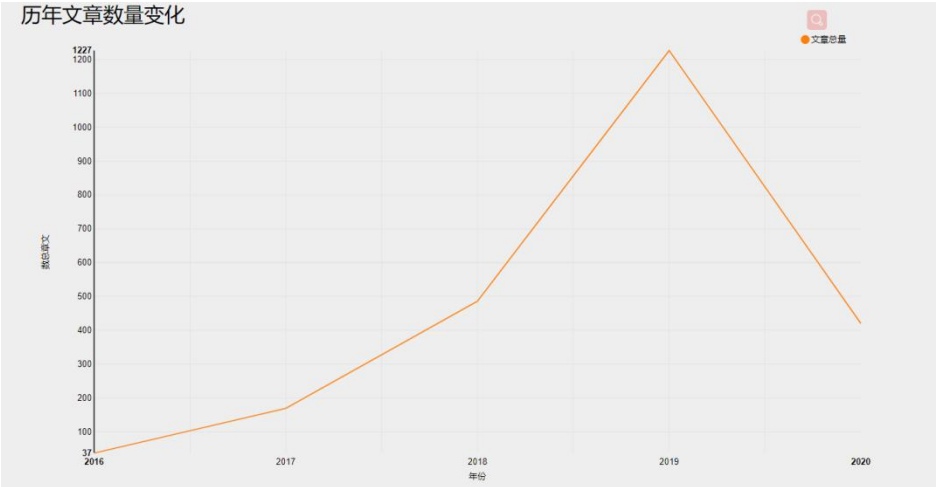


图 1 区块链研究各年份发文量

对区块链文献的发文国家进行分析与探讨，不仅有利于理清各国的研究现状，也有有利于进一步研究区块链的适用国情和研究范式。

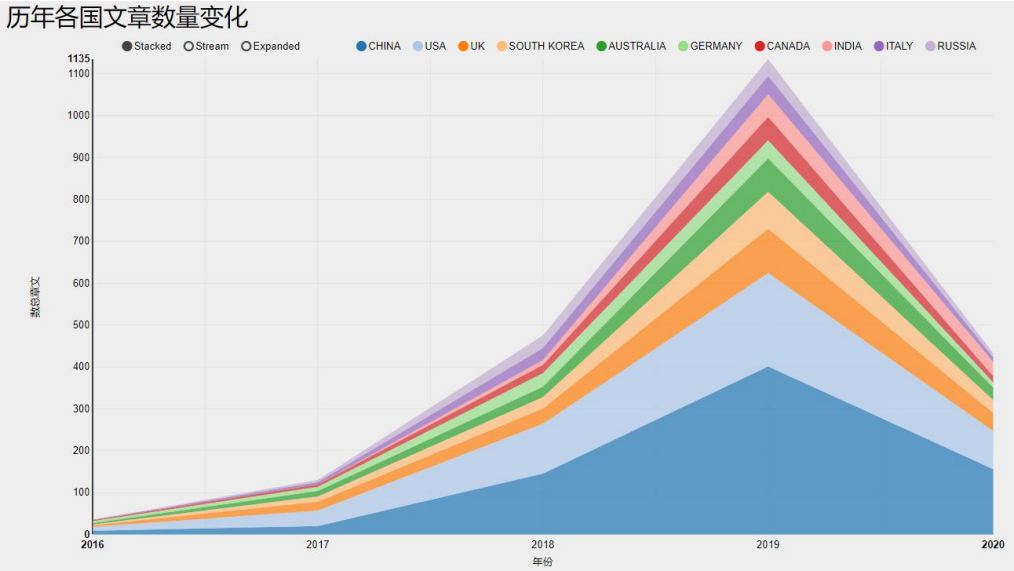


图 2 核心国家及其发文量

从图 2 可以看出，中国和美国的发文量远远领先于英国、韩国、澳大利亚等国家，这在很大程度上与中美两国相对完善的科学研究体系和金融业发展情况息息相关。从时间分布来看，中国对区块链研究的进展最为迅速，在 2017 年-2020 年间中国的发文量大量增长，赶超美国成为全球第一，其他国家的发文量也都有不同程度的增长。

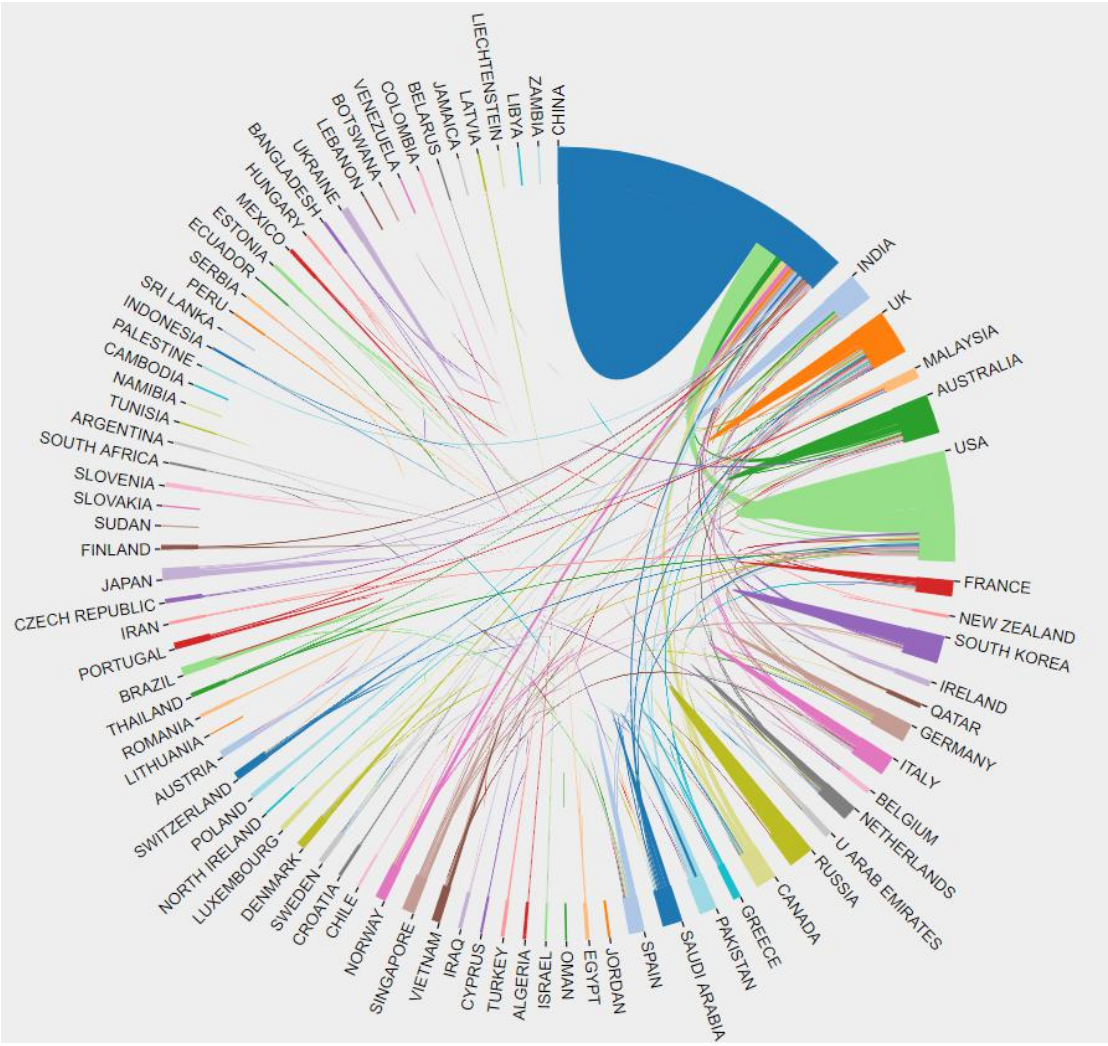


图 3 国家间发文合作关系

从图 3 来看，区块链研究已经得到了众多国家的广泛关注，由于综合国力、人才培养、政策支持等差异，不同国家的区块链研究实力存在差距，但在中国、美国、英国、澳大利亚等国家牵头的合作研究环境下，越来越多的国家开始积极投入到区块链的研究当中。

3.2 学科分布

学术期刊既是学术成果传播的重要纽带和载体，也是开展学术研究的基础。从期刊所在的学科分布即可看出文献所在的学科分布，以更好地把握区块链研究现状。在 CiteSpace 中绘制双图叠加（dual-map overlay）图谱，如图 4 所示。

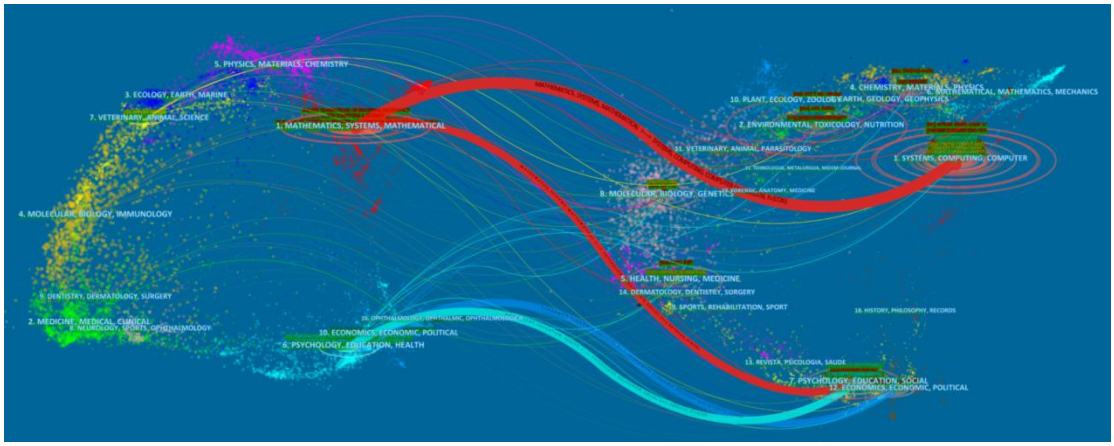


图 4 区块链研究学科分布

在这幅图中， 左侧是施引文献所在的期刊分布，代表了区块链所属的主要学科（如左侧的 Mathematics，Systems，mathematical；psychology，enucation，health；）；右侧是对应被引文献所在的期刊分布，代表了区块链领域主要引用了哪些学科（如右侧的 Systems,computing,computer；psychology，education，social；economics，economic，political）。前者可以看做是区块链的领域应用，包括数学、物理、化学、地理学、生物学、医药、教育、经济等，应用研究十分广泛；后者可以看做是区块链的研究基础，核心为系统、计算、环境、社会研究等。

3.3 核心引文

引文分析可以帮助研究者寻找某个研究领域的发展脉络，快速锁定某个研究方向的重要文献和核心作者，还可以找到某些具有开创性成果的无指定关键词的论文。使用 HistCite 对文献的本地被引、全部被引、本地引用、全部引用进行整理，按照本例被引倒序排列得到 TOP10 的核心引文，如图 5 所示。

http://127.0.0.1:1925/ - HistCite - download_01.txt - Internet Explorer

File Analyses View Tools Help HistCite™

Untitled Collection
List of All Records

Grand Totals: LCS 15768, GCS 24288, CR 168192
Collection span: - 2020

Records: 5512, Authors: 12275, Journals: 2082, Cited References: 104238, Words: 5651
Yearly output | Document Type | Language | Institution | Institution with Subdivision | Country

|< << < > >> >|

#	Date / Author / Journal	LCS	GCS	LCR	CR
1	197 Christidis K, Devetsikiotis M **Blockchains and Smart Contracts for the Internet of Things** IEEE ACCESS. 2016; 4: 2292-2303	536	612	0	53
2	167 Kosba A, Miller A, Shi E, Wen Z, Papamanthou C **Hawk: The Blockchain Model of Cryptography and Privacy-Preserving Smart Contracts** 2016 IEEE SYMPOSIUM ON SECURITY AND PRIVACY (SP). 2016; : 839-858	233	271	0	53
3	438 Zheng ZB, Xie SA, Dai HN, Chen XP, Wang HM **An Overview of Blockchain Technology: Architecture, Consensus, and Future Trends** 2017 IEEE 6TH INTERNATIONAL CONGRESS ON BIG DATA (BIGDATA CONGRESS 2017). 2017; : 557-564	227	242	4	51
4	199 Tschorsch F, Scheuermann B **Bitcoin and Beyond: A Technical Survey on Decentralized Digital Currencies** IEEE COMMUNICATIONS SURVEYS AND TUTORIALS. 2016; 18 (3): 2084-2123	215	254	0	242
5	213 Azaria A, Ekblaw A, Vieira T, Lippman A **MedRec: Using Blockchain for Medical Data Access and Permission Management** PROCEEDINGS 2016 2ND INTERNATIONAL CONFERENCE ON OPEN AND BIG DATA - OBD 2016. 2016; : 25-30	194	208	0	11
6	142 Eyal I, Gencer AE, Sirer EG, van Renesse R **Bitcoin-NG: A Scalable Blockchain Protocol** 13TH USENIX SYMPOSIUM ON NETWORKED SYSTEMS DESIGN AND IMPLEMENTATION (NSDI '16). 2016; : 45-59	152	156	0	57
7	265 Underwood S **Blockchain Beyond Bitcoin** COMMUNICATIONS OF THE ACM. 2016 NOV; 59 (11): 15-17	149	168	0	1
8	2541 Aitzhan NZ, Svetinovic D **Security and Privacy in Decentralized Energy Trading Through Multi-Signatures, Blockchain and Anonymous Messaging Streams** IEEE TRANSACTIONS ON DEPENDABLE AND SECURE COMPUTING. 2018 SEP-OCT; 15 (5): 840-852	148	169	0	23
9	706 Lansiti M, Lakhani KR **THE TRUTH ABOUT BLOCKCHAIN** HARVARD BUSINESS REVIEW. 2017 JAN-FEB; 95 (1): 119-127	144	165	0	0
10	182 Luu L, Chu DH, Olickel H, Saxena P, Hobor A **Making Smart Contracts Smarter** CCS'16: PROCEEDINGS OF THE 2016 ACM SIGSAC CONFERENCE ON COMPUTER AND COMMUNICATIONS SECURITY. 2016; : 254-269	142	164	0	28

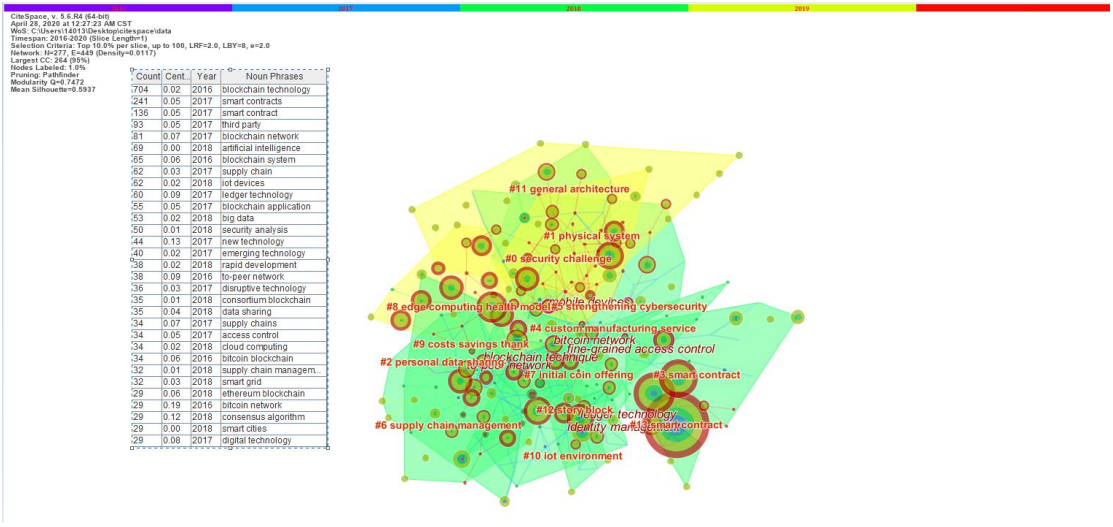
图 5 本地被引频次排序

研读图 5 中出现的 10 篇文献发现，其中 4 篇文献对区块链的基础核心技术做出了改进和展望，3 篇文献是研究基于区块链的智能合约技术，2 篇文献是讨论区块链的安全性和隐私性，还有 1 篇文献是较早地应用了区块链技术到具体领域中。其中智能合约技术的有关文献具有最高的被引频次，说明智能合约不仅是区块链的核心基础应用方案，还一直得到广泛的研究和应用；而其余 7 篇文献其中 6 篇涉及基础技术研究和基本安全问题，还有一篇是最早的优质区块链学术应用，由于其研究开展和文献发表时间较早，取得了更高的地位优势，是后来的研究避免不了的引文，被引频次也较高。

3.4 研究热点

3.4.1 主题词共现图谱聚类

关键词是一篇论文的核心概括，对论文关键词进行分析可对文章主题窥探一二。而在以往的研究中，研究者们也往往是直接对论文给出的关键词进行主题分析。这种方法的确能比较精确地概括文章的主题，但不能很好地发现文章中使用的分析工具。出于这种考虑，笔者团队将标题、摘要、关键词综合考虑，作为主题词。一篇论文中的主题词一定存在某种关联，这种关联可以用共现的频次来表示。一般认为，词汇在同一片文献中出现的次数越多，则这两个主题的关系越紧密。共词分析法就是利用文献集中词汇或名词短语共同出现的情况，来确定该文献集所代表学科中各主题之间的关系。统计一组文献的主题词两两之间在同一篇文献出现的频率，便形成一个由这些词对关联所组成的共词网络，进行简单聚类后如图 6 所示。



Count	Cent.	Year	Noun Phrases
704	0.02	2016	blockchain technology
241	0.05	2017	smart contracts
136	0.05	2017	smart contract
93	0.05	2017	third party
81	0.07	2017	blockchain network
69	0.00	2018	artificial intelligence
65	0.06	2016	blockchain system
62	0.03	2017	supply chain
62	0.02	2018	iot devices
60	0.09	2017	ledger technology
55	0.05	2017	blockchain application
53	0.02	2018	big data
50	0.01	2018	security analysis
44	0.13	2017	new technology
40	0.02	2017	emerging technology
38	0.02	2018	rapid development
38	0.09	2016	to-peer network
36	0.03	2017	disruptive technology
35	0.01	2018	consortium blockchain
35	0.04	2018	data sharing
34	0.07	2017	supply chains
34	0.05	2017	access control
34	0.02	2018	cloud computing
34	0.06	2016	bitcoin blockchain
32	0.01	2018	supply chain managem...
32	0.03	2018	smart grid
29	0.06	2018	ethereum blockchain
29	0.19	2016	bitcoin network
29	0.12	2018	consensus algorithm
29	0.00	2018	smart cities
29	0.08	2017	digital technology

图 6 主题词共词网络

其中节点数就是图中的主题词个数，边数就是主题词之间的连线数。只要主题词在同一片文献中出现过，两者之间就会有一条连线。图中圆圈大小代表的是关键词频次，频次越大，圆圈越大。线条代表主题词之间的联系，线条颜色与图中上方年份对应，用于标志每一年有哪些主要主题词。

尽管从视觉上看，各聚类分隔不够清楚，但网络的模块度达 0.7472，表明这一共词聚类可以比较清楚地界定出区块链的各个子领域。聚类效度评价的另一个指标——平均轮廓值也达到了 0.5937，在网络中有部分小聚类存在的情况下，这个分值已经比较令人满意了。从图中可以看到，智能合约（smart contract）与定制制造服务（custom manufacturing service）被聚为一类，分布式账本（ledger technology）、身份管理（identity management）和物联网（iot environment）被聚为一类，总体架构（general architecture）、物理系统（physical system）和安全挑战（security challenge）被聚为一类，总体而言大型节点的聚类结果符合区块链技术常识。

3.4.2 作者共被引

作者拱北因分析方法的基本假设是：当两个作者的文献同时被第三个作者的文献引用，则称这两个作者存在共引关系；如果这两位作者共被引频次越高，则说明他们的学术关系越密切，“距离”越近。选取各年份被引频次排名前 50 的作者（称为核心作者）对应的导入 CiteSpace 绘制作者共被引网络图谱，进行简单聚类后得到与图 5 性质相似的图 7。其中网络模块度达到 0.8106，平均轮廓值尽管只有 0.3972，但图中小聚类还比较多，这个分值还可以接受，聚类效果良好。

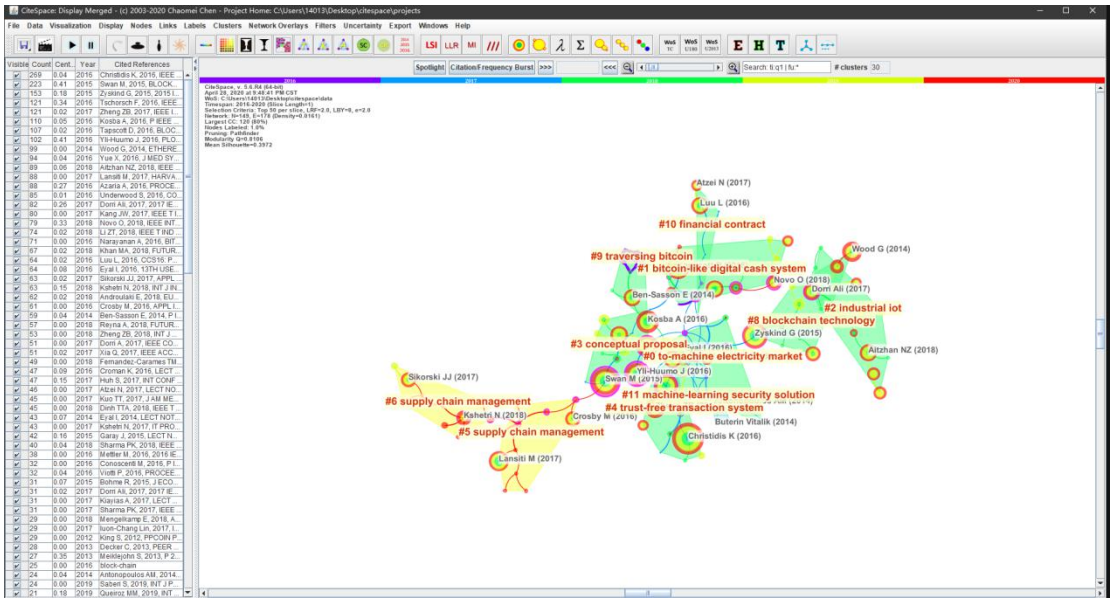


图 7 作者共被引聚类

从图 7 中可以看出，Lansiti M、Kshetri N 和 Sikorski JJ 等作者共同建立了供应链管理（supply chain management）的学术研究子群，Swan M 则是机器学习（machine-learning）相关领域的核心作者，Atzei N 和 Luu L 等人则专攻金融合约（financial contract）方向。从各个领域聚类情况来看，大多都有独当一面的引领着，作者间形成了稳定的合作关系，形成了多个以某一个或某几个作者为核心的研究团队。

3.4.3 研究热点总结

研究热点是指在某一特定时间段内，被许多研究者所一致关注并研究探讨的话题，从知识理论的角度看，中心性和高频词的关键词可以作为代表。中心性作为衡量节点权力的大小，反映了该点在网络中的重要性。关键词的共现频次越高，点中心性越高，说明节点在该领域越重要。

根据前两节的聚类成果和研究结论，区块链领域的前 5 研究热点主题为智能合约（smart contract）、比特币（bitcoin）、数字加密货币（cryptocurrency）、隐私/私有（privacy）、物联网（iot）。具体来讲，区块链的技术核心和理论研究方法已经日趋成熟，目前做的相关工作主要是将区块链技术应用到具体领域中，进行创新式的商业发现，特别是智能合约和数字货币技术。从整体上看，区块链创新研究目前还处于知识体系整合的变革阶段，创新点主要在应用方面。

3.5 研究热点变迁趋势

3.5.1 时间线图谱

为了更加清晰地了解每个时间段的研究热点，看出区块链技术从 2016 年到 2020 年的发展脉络及研究热点的变化，探寻区块链技术的研究前沿，从主题词（沿用 3.3.1 中设定）和核心作者两个角度生成时间线图谱。

3.5.1.1 主题词

图 8 和图 9 用两种图谱显示模式模式从主题词的角度反映了区块链技术从 2016 年到 2020 年的研究热点变化态势及研究热点间的联系。图中圆圈代表某篇文献，圆环的大小代表了相应文献的被引频次，连线代表引用关系，颜色代表文献的发表年份，文献排序从上到下为从新到旧排列。

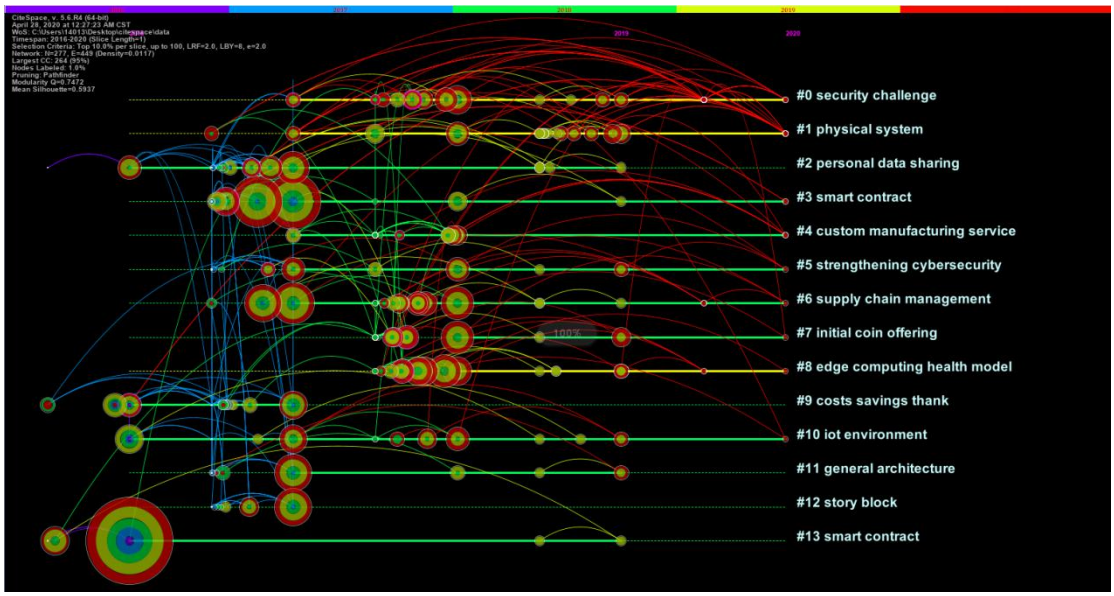


图 8 区块链研究主题词时间线图谱-1

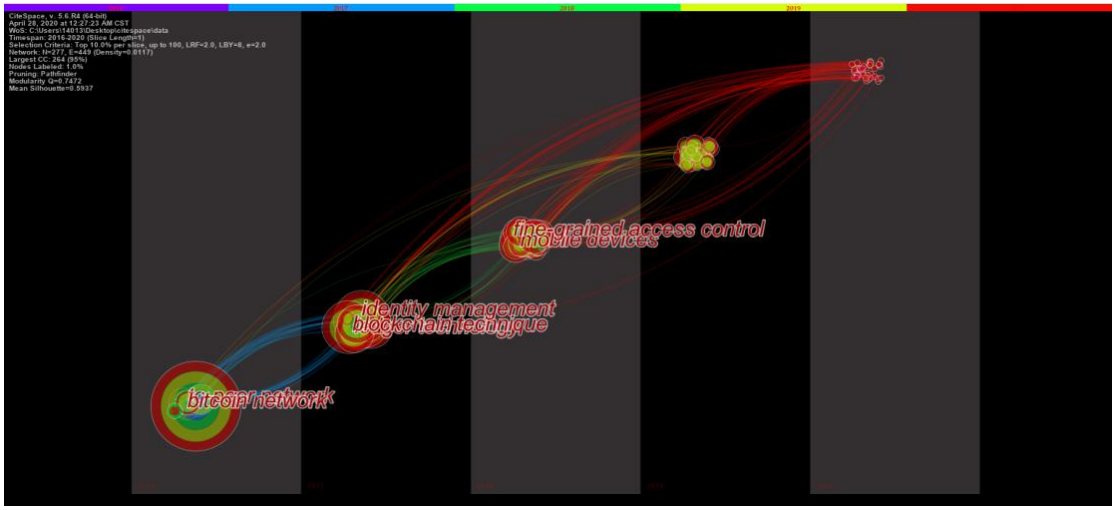


图 9 区块链研究主题词时间线图谱-2

结合图 8 和图 9 可以看出，2016 年到 2018 年的研究热点依次是比特币网络（bitcoin network）、身份管理（identity management）、移动设备（mobile devices）；区块链技术（block chain technique）一直是各年研究的重点。随着研究方向的变化和社会发展的进步，以及各国的政策支持变化情况，很多研究热点随之改变，最近两年细粒度访问控制（fine-grained access control）、边缘计算健康模型（edge computing health model）、物理系

统（physical system）和安全挑战（security challenge）等成为了区块链领域的研究前沿。

3.5.1.2 核心作者

图 10 和图 11 用两种图谱显示模式模式从主题词的角度反映了区块链技术从 2016 年到 2020 年的研究热点变化态势及研究热点间的联系。

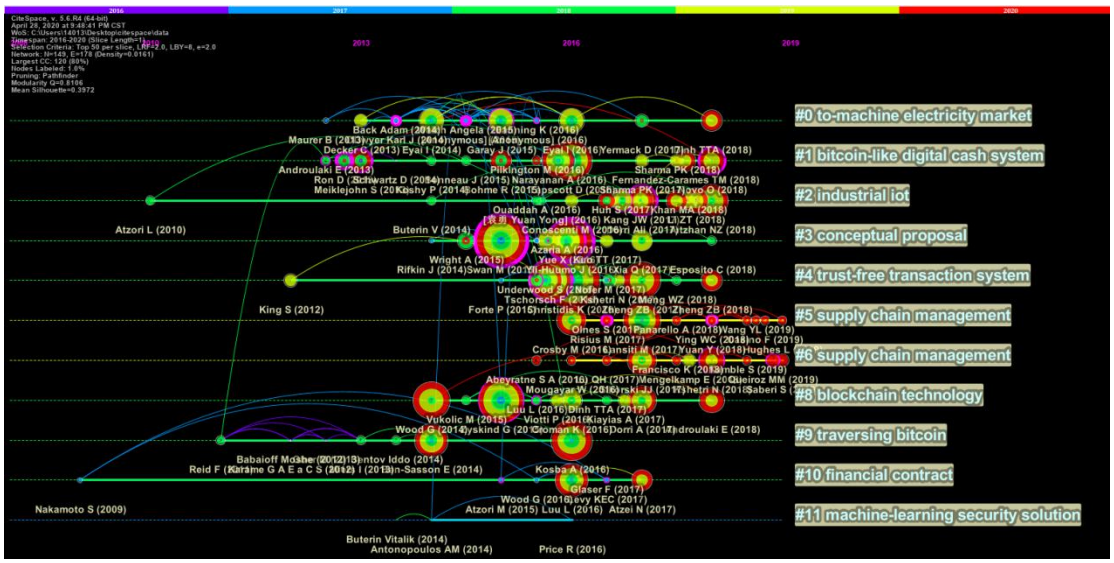


图 10 区块链研究核心作者时间线图谱-1

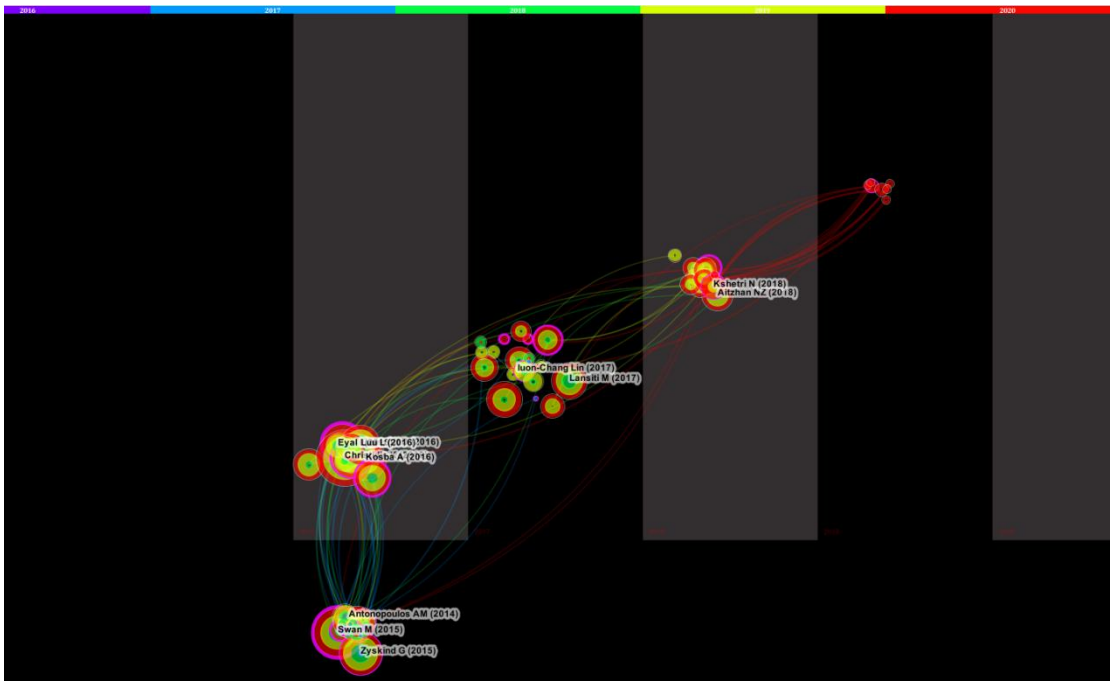


图 11 区块链研究核心作者时间线图谱-2

结合图 9 和图 10 可以看出，2016 年的核心作者主要有 Luu L、Kosba A、Eyal I 等，他们的主要研究方向为比特币（bitcoin）和类比特币数字货币系统（bitcoin-liKe digital cash system）；2017 年的核心作者为 Iuon-Chang Lin 和 Lansiti M，他们的主要研究方向机器学习安全解决方案（machine-learning security solution）；2018 年的核心作者为 Aitzhan NZ 和 Kshetri N，他们的主要研究方向为概念性提案（conceptual proposal）；物联网（iot）和机电市场（machine electricity market）相互关联，在 2018 年出现后知道 2020 年依旧火热；区块链技

术（blockchain technology）贯穿整条时间线；特别地，供应链管理（supply chain management）从 2017 年提出以来，直到 2020 年还是研究热点之一，也是 2020 年最重要的研究热点。

3.5.2 突现词

突现词指在较短时间内出现较多或使用频率较高的词，根据突现词的次品变化可以判断研究领域的前沿与趋势。根据 CiteSpace 的相关功能，得到区块链领域突现主题词及对应的突显率和被引热度曲线，如图 12 所示。

突现率前三的主题词和突现时间分别为数字货币（digital currency,2017）、公共账本（public ledger,2016）和金融业（financial sector,2017）；突现率共同排行 4-7 的主题词分别为比特币系统（bitcoin system）、公有链（public blockchain）、区块链协议（blockchain protocol）和数据资产（digital assets），突现时间均为 2017 年。这在一定程度上说明，目前区块链研究前沿主要集中在这些领域。

Top 15 Terms with the Strongest Citation Bursts

Terms	Year	Strength	Begin	End	2016 - 2020
digital currency	2016	5.4197	2017	2018	
public ledger	2016	4.8599	2016	2018	
financial sector	2016	4.165	2017	2018	
bitcoin system	2016	3.7473	2017	2018	
public blockchain	2016	3.7473	2017	2018	
blockchain protocol	2016	3.7473	2017	2018	
digital assets	2016	3.7473	2017	2018	
financial system	2016	3.3298	2017	2018	
block chain	2016	3.2649	2016	2017	
blockchain transaction	2016	2.9126	2017	2018	
block chain technology	2016	2.9126	2017	2018	
smart devices	2016	2.9126	2017	2018	
blockchain model	2016	2.9126	2017	2018	
financial markets	2016	2.4957	2017	2018	
key challenge	2016	2.4957	2017	2018	

图 12 区块链关键词凸显率

4 结论

本文利用可视化软件 Cite Space、HistCite 和文献计量在线分析平台，以 Web of Science 核心合集中 2016-2020 年 2294 篇关于区块链研究文献为数据池，对区块链研究文献分别进行年度发文量分析、学科分布分析、核心引文分析、主题词和作者共现聚类分析、时间线和突现词分析等，在客观数据基础上进行讨论，梳理出区块链领域的热点研究问题和变迁趋势。

研究结论表明：

（1）区块链技术研究文献从 2017 年后出现快速发展的趋势。区块链技术发展时间不长，各国支持力度打，学术领域还有广阔的研究空间。

（2）中美两国在区块链技术领域处于世界领先地位。中国在该领域的发文量位居世界第一，美国的发文量位居世界第二。

（3）区块链技术研究领域的最近五年的热点文献有 Blockchains and Smart Contracts for the Internet of Things;Hawk: The Blockchain Model of Cryptography and Privacy-Preserving Smart Contracts;An Overview of Blockchain Technology: Architecture, Consensus, and Future Trends;Bitcoin and Byond: A Technical Survey on Decentralized Digital Currencies;MedRec: Using Blockchain for Medical Dat Access and Permission Management;Bitcoin-NG: A Scalable Blockchain Protocal;Blockchain Beyond Bitcoin;Security and Privacy in Decentralized Energy Trading Through Multi-Signatures, Blockchain and Anonymous Messaging Streams; THE TRUTH ABOUT BLOCKCHAIN;Making Smart Contracts Smarter，这些文献对于区块链技术领域的研究具有重要的参考意义。

（4）区块链技术研究领域最近五年的高影响力人物有 Luu L、Kosba A、Eyal I、luon-Chang Lin、Lansiti M、Aitzhan NZ、Kshetri N 等人。

（5）区块链技术研究领域最近五年的研究热点包括比特币（bitcoin）、密码协议（cryptography）、物联网（internet of thing）、智能合约（smart contract）、安全（security）、隐私（privacy）、供应链管理（supply chain management）等。

（6）区块链技术研究领域最近五年逐渐地持续地从核心技术的开发和改进，转移到应用技术层面，最重要的是更加广泛地应用于金融业和信息产业。研究前沿主要是智能合约（smart contract）、密码协议（cryptography）和安全（security）等。

参考文献

[1] 许振宇,吴金萍,霍玉蓉.区块链国内外研究热点及趋势分析[J].图书馆,2019(04):92-99.

[2] 程豪,张峥.基于 CiteSpace 分析的区块链技术可视化研究[J].物流科技,2019,42(02):7-11+22.

[3] Chen C. CiteSpace Ⅱ: Detecting and visualizing emerging trends and transient patterns in scientific literature[J].Journal of the American Society for Information Sci-ence and Technology,2006,57(3):359-377.

[4] 王娟,陈世超,王林丽,杨现民.基于 CiteSpace 的教育大数据研究热点与趋势分析[J].现代教育技术,2016,26(02):5-13.

[5] 陈超美（2016）CiteSpace 的分析原理. 科学知识图谱：前沿与实践. 北京，高等教育出版社.

[6] 邱均平,马瑞敏,李晔君.关于共被引分析方法的再认识和再思考[J].情报学报,2008,27(1):69-74. DOI:10.3969/j.issn.1000-0135.2008.01.011.

附录

完整项目文件链接：https://github.com/ZSC2017IM/datamining_20200430