Research Hotspots Mining and Trend Analysis of Blockchain

Ji Fang1, Wenyan Tu1, Fangyu Jiao, and Yutong Ci1*

1 Library, Huazhong University of Science and Technology, Wuhan, Hubei, 430074, China

*Corresponding author’s e-mail: cba@hust.edu.cn

Abstract. In this paper, the bibliography included in Web of Science is taken as the data sample. This paper analyzes articles, reviews and proceedings papers published from 2013 to 2020 in the field of blockchain to reveal the major countries and institutions. We use co-word method to identify research hotspots in the field of blockchain. The results show that the research on blockchain is an emerging research area. China and The United States are always in the leading position. Beijing University of Posts & Telecommunications publish most papers in the field of blockchain. Smart contract, Internet of Thing, bitcoin, security and privacy are the hottest research topics in the field of blockchain.

1. Introduction
Blockchain is a new application model of computer technology such as distributed data storage, peer-to-peer transmission and encryption algorithm. Blockchain originated from bitcoin. The concept of blockchain was proposed by Satoshi Nakamoto in 2008. How to use blockchain technology to promote the development of human society has become the common focus of the industry and academia. The research on blockchain covers computer science, engineering, social science, economics, business, chemistry, clinical medicine and agricultural science. It may be widely used in variety of fields in the future. Makhdoom[1], Salah[2]and Fernandez[3] gave a detailed overview of blockchain technology about Internet of Things, AI and smart factories. Most of the studies are qualitative. Therefore, we adopt bibliometrics methods to analyze academic papers, in order to understand the latest research progress and development tendency in the field of blockchain.

2. Data sources and methods
2.1. Data sources
Data in this paper is collected from SCIE, SSCI, A&HCI, CPCI-S and CPCI-SSH in Web of Science core collection database. The search strategy is TS=(“blockchain” or “block chain” or “bitcoin”). The time span is 1996-2020. The literature type is limited to article, review and proceedings paper. Finally, 8020 related papers are retrieved. Download bibliography information (including title, author, institution, journal name, publication year, keywords, etc.) of the 8020 papers from Web of Science. We retrieved the data on March 1, 2021.

2.2. Analysis methods and tools
In this paper, the bibliometrics method is used to analyze the papers in the field of blockchain. Social network analysis is used to reveal the global trend, major countries and institutions in this field. Social network analysis can visually present the cooperative relationship between countries, and describe the
characteristics of the cooperative network between countries in the field of blockchain, through constructing the national cooperative network. Co-word analysis is used to identify research hotspots in this field. Co-word analysis can help draw the knowledge map of the blockchain and analyze the research hotspots in this field. In this paper, we use Vosviewer to draw the knowledge map of the blockchain.

3. Empirical analysis

3.1 Posting trend

Figure 1 shows the time distribution of papers in the field of blockchain. From 2013 to 2020, 8020 papers were published. The number of papers shows an increasing trend. The first paper was published in 2013 which shows that blockchain emerged late. From 2013 to 2016, the number of papers grew slowly. After 2017, the growth rate became faster and the research heat began to increase. The number of papers published in 2017 was five times as many as 2016. This reflects the fact that blockchain technology is an emerging field.

![Figure 1. Chronological distribution of papers in the field of blockchain.](image)

3.2 Locations analysis

Table 1 shows the major countries or regions which published most papers in the field of blockchain. China published 2003 papers which rank first around the world. The United States published 1535 papers which has the highest citations and higher influence. The top ten countries or regions in the blockchain field publish international cooperation papers accounting for more than 30%, which indicates that the international cooperation in this field is extensive. Other top 10 countries or regions are: United Kingdom, India, South Korea, Australia, Germany, Canada, Italy and France.

| Country/Region     | Paper Number | Citations | International Cooperation Ratio |
|--------------------|--------------|-----------|---------------------------------|
| China              | 2003         | 14283     | 41.99%                          |
| USA                | 1535         | 16176     | 48.27%                          |
| United Kingdom     | 613          | 4864      | 64.27%                          |
| India              | 532          | 2511      | 40.23%                          |
| South Korea        | 435          | 2932      | 29.43%                          |
| Australia          | 427          | 3834      | 67.45%                          |
| Germany            | 368          | 2743      | 38.04%                          |
| Canada             | 363          | 2644      | 63.64%                          |
| Italy              | 323          | 2154      | 43.34%                          |
| France             | 233          | 1631      | 58.80%                          |
If n papers belong to both country A and country B, a link with weight n is established between node A and B in the national cooperation network. Set every two years as a time window. We construct four national cooperative networks of 2013-2014, 2015-2016, 2017-2018 and 2019-2020. In order to further explain the evolution of national cooperation in the field of blockchain, we calculate the indicators of national cooperation network in Table 2. The result shows that the number of nodes and edges in the national cooperative network presents an increasing trend, while the network density generally presents a decreasing trend. The number of edges in the national cooperative network in t3 time window increased to 237, which indicated new partnerships came into being and knowledge fusion within the field was improved.

Table 2. The indicators of national cooperation network.

| Time Window | Node | Edge | Density |
|-------------|------|------|---------|
| t1          | 5    | 1    | 0.1     |
| t2          | 21   | 22   | 0.1048  |
| t3          | 79   | 237  | 0.0769  |
| t4          | 91   | 273  | 0.0667  |

3.3. Institutions analysis
Table 3 shows the major institutions that published most papers in the field of blockchain. Beijing University of Posts & Telecommunications published 145 papers which has the highest citations influence around the world. The second is University of Electronic Science & Technology of China that published 91 papers. Most of the top 10 institutions are academic institutions. Besides, IBM is the only company that ranks top 10.

Table 3. The top ten institutions in the volume of papers.

| Institutions | Paper Number | Citations |
|--------------|--------------|-----------|
| Beijing University of Posts & Telecommunications | 145 | 1437 |
| University of Electronic Science & Technology of China | 91 | 1023 |
| Xidian University | 79 | 552 |
| Nanyang Technological University | 77 | 919 |
| University of New South Wales Sydney | 70 | 1095 |
| IBM | 69 | 822 |
| Tsinghua University | 67 | 525 |
| Shanghai Jiao Tong University | 67 | 541 |
| Hong Kong Polytechnic University | 66 | 703 |
| King Saud University | 65 | 590 |

3.4. Highly cited papers analysis
Table 4 shows the top five highly cited papers in the field of blockchain. The most cited paper is "Blockchains and Smart Contracts for the Internet of Things", published in IEEE Access in 2016. This paper is cited 1020 times. This paper discusses how the combination of smart contracts with the Internet of Things can facilitate the sharing of services\textsuperscript{[4]}. The second most cited paper is "Decentralizing Privacy: Using Blockchain to Protect Personal Data", published in 2015 IEEE security and privacy workshops in 2015, with a total of 607 citations. This paper describes a more secure and distributed personal data management system that allows users own and control their data\textsuperscript{[5]}. The third most cited paper is "An Overview of Blockchain Technology: Architecture, Consensus, and Future Trends" published in 2017, with a total of 510 citations. This paper gives a detailed overview of blockchain and introduces smart contracts briefly\textsuperscript{[6]}.

"Industry 4.0: state of the art and future trends" was published in International Journal of Production Research in 2018, with a total of 487 citations. This paper gives a detailed overview of blockchain architecture, typical algorithms, technical...
challenges, recent developments and future trends.\cite{7} “Hawk: The Blockchain Model of Cryptography and Privacy-Preserving Smart Contracts” was published in 2016 IEEE symposium on security and privacy in 2016, with a total of 471 citations. This paper implements a compiler model named "HAWK", which can protect privacy based on smart contracts.\cite{8}

| Article Title | Year | Source | Citations |
|---------------|------|--------|-----------|
| Blockchains and Smart Contracts for the Internet of Things | 2016 | IEEE Access | 1020 |
| Decentralizing Privacy: Using Blockchain to Protect Personal Data | 2015 | 2015 IEEE Security and Privacy Workshops | 607 |
| An Overview of Blockchain Technology: Architecture, Consensus, and Future Trends | 2017 | 2017 IEEE 6th International Congress on Big Data | 510 |
| Industry 4.0: state of the art and future trends | 2018 | International Journal of Production Research | 487 |
| Hawk: The Blockchain Model of Cryptography and Privacy-Preserving Smart Contracts | 2016 | 2016 IEEE Symposium on Security and Privacy | 471 |

3.5. Research hotspots analysis

Keywords represent the main content of the article. High-frequency keywords are the words that appear frequently in a certain research field. To some extent, they reflect research hotspots of the field. The top 20 high-frequency keywords in Table 5 appear more than 120 times. Among them, blockchain, smart contract, Internet of Thing, security, bitcoin, privacy, ethereum, cryptocurrency, cloud computing and trust are in the top 10 of all keywords.

| Keywords | Frequency | Keywords | Frequency |
|-----------|-----------|-----------|-----------|
| blockchain | 4983 | architecture | 203 |
| smart contract | 1055 | authentication | 194 |
| internet of thing | 810 | supply chain | 187 |
| security | 748 | framework | 186 |
| bitcoin | 594 | access control | 177 |
| privacy | 488 | distributed ledger | 167 |
| ethereum | 400 | big data | 163 |
| cryptocurrency | 326 | consensus | 162 |
| cloud computing | 288 | distributed ledger technology | 133 |
| trust | 224 | edge computing | 127 |

However, the research hotspots in the field cannot be revealed only through high-frequency keywords, the co-word analysis method should be used to further reveal the relationships among high-frequency keywords. The co-word analysis method uses the co-occurrence frequency of keyword pairs to reflect the correlation strength between keywords, and displays the research structure of the field in the form of visualization. In this paper, Vosviewer is used to draw the knowledge map of co-occurrence network of keywords in the field of blockchain, as shown in Figure 2. From the point of node centrality, cryptography, technology, management, privacy, blockchain, security, cryptocurrency and trust have highest network centrality. These words are in the central position of the network and associated with several other nodes. In addition, cloud computing, Internet of Thing, bitcoin, access control, smart contract, ethereum and big data also have high centrality and also play an important role in the network.

Combined with high-frequency keywords and keywords with high centrality, it can be seen that the research hotspots in the field of blockchain focus on smart contract, Internet of Things, bitcoin, security and privacy issues. Smart contracts are scripts on top of blockchain technology. Smart contracts represent a form of automation by reducing or completely replacing mediation layers. As a
result, blockchain smart contract systems reduce transaction and execution costs as well as processing time. Griggs created a system where the sensors communicated with a smart device that called smart contracts and wrote records of all events on the blockchain[9]. Internet of Things (IoT) is the "Internet of everything connected". It is an extension of the internet and an extended network, which combines all kinds of information sensing devices to form a huge network. Internet of Things realizes the interconnection of people, machines and things at any time and any place. Bitcoin brought blockchain technology to the public. In the future, blockchain will greatly promote the application of digital currency and other fields. The application and development of blockchain need secure system security. Various countries and authorities focus on the research of the security of blockchain. Privacy issues in the field of blockchain have been raised. While blockchain technology can improve efficiency and reduce costs, blockchain technology also faces serious privacy leakage problems, which has received extensive attention from researchers.

Figure 2. Co-occurrence network of keywords.

4. Conclusion
In this paper, bibliometrics methods and co-word analysis are used to analyze the development tendency of blockchain. We analyze the publishing trends, major countries or institutions and research hotspots based on papers in the field of blockchain. The analysis draws the following conclusions: the number of papers published in the field of blockchain is increasing obviously. China and the United States are the leading producers in this field. The research hotspots focus on smart contract, Internet of Things, bitcoin, security and privacy issues.

At the same time, this paper also has some shortcomings. Data source is relatively single, for example, we only selected the data from Web of Science. In the following research, we will expand the data source and use more diversified research methods, so as to provide some reference for scholars to
research blockchain technology.

References
[1] Makhdoom, I., Abolhasan, M., Abbas, H., et al. (2019) Blockchain's adoption in IoT: The challenges, and a way forward. Journal of Network and Computer Applications, 125:251-279.
[2] Salah, K., Rehman, M., Nizamuddin, N., et al. (2018) Blockchain for AI: Review and Open Research Challenges. IEEE Access, 7:10127-10149.
[3] Fernandez-Carames, T., Fraga-Lamas, P. (2019) A Review on the Application of Blockchain for the Next Generation of Cybersecure Industry 4.0 Smart Factories. IEEE Access, 1-1.
[4] Christidis, K., Devetsikiotis, M. (2016) Blockchains and Smart Contracts for the Internet of Things. IEEE Access, 4:2292-2303.
[5] Zyskind, G., Zekrifa, D., Alex, P., et al. (2015) Decentralizing Privacy: Using Blockchain to Protect Personal Data. In: IEEE Security & Privacy Workshops. San Jose. pp. 180-184.
[6] Zheng, Z.B., Xie, S.A., Dai, H.N., et al. (2017) An Overview of Blockchain Technology: Architecture, Consensus, and Future Trends. In: 6th IEEE International Congress on Big Data. Honolulu. pp. 557-564.
[7] Xu, L., Xu, E., Li, L., et al. (2018) Industry 4.0: state of the art and future trends. International Journal of Production Research, 56: 2941-2962.
[8] Kosba, A., Miller, A., Shi, E., et al. (2016) Hawk: The Blockchain Model of Cryptography and Privacy-Preserving Smart Contracts. In: 2016 IEEE Symposium on Security and Privacy (SP). San Jose. pp. 839-858.
[9] Griggs, K., Olya, O., Kohlios, C., et al. (2018) Healthcare Blockchain System Using Smart Contracts for Secure Automated Remote Patient Monitoring. Journal of Medical Systems, 42:130.