The Development Trend of Artificial Intelligence in Cyberspace Security: A Brief Survey

ZiyuanLi*, MingXian1, Jian Liu1 and HuimeiWang1

1College of Electronic Science and Technology, National University of Defense Technology, Changsha, Hunan,410073, China

*Corresponding author’s e-mail:303272245@qq.com

Abstract. In recent years, artificial intelligence and network security have attracted much attention. After decades of practice, artificial intelligence and relevant theoretical methods have provided new ideas for the traditional field of cyberspace security. Knowledge mapping is a tool to show the relationship between the development process and structure of science and technology. At present, there is no knowledge mapping that can describe the development trend of artificial intelligence in the field of cyberspace security. So we used Bibliometric, Citespace and VOSviewer to analyze more than 1,000 quality articles downloaded from Web of Science (WoS), include literature quantity analysis, co-citation analysis, keyword co-occurrence analysis and Research hotspot analysis.

1. Introduction

With the growing network traffic, artificial intelligence plays a more important role in civil network and military development[1], and the monitoring scheme and attack weapon in cyberspace security are more effective in the development of artificial intelligence and related methods.

In the future, the extension of artificial intelligence technology in the cyberspace security is also likely to lead to a significant change in the national strategy, so it is important to study the development trend of artificial intelligence in the field of cyberspace security.

As a branch of information science, Bibliometry, using the database, knowledge mapping and other technologies, summarizes the development trends of some areas through literature analysis.[2],[3], [4] are examples of investigation and induction of a area this way.

This paper is derived from the Web of Science (WOS), which is used to ensure the authority and reliability of the literatures, only retrieving the article literatures of sci-e in the retrieval settings, and eliminating the conference articles. The time range is set to 1999-2018, for setting the relevant topic index strategy, and finally retrieving 1,043 documents as the source of data to analysis.

2. Research tools and data sources

With the accelerated development of information society and the promotion of network application, scholars have realized the visualization of subject knowledge by drawing the scientific knowledge mapping of literature data information, so as to conduct research on the growth of research theory, the transformation of research paradigm, the evolution of subject field and the identification of subject structure.

In our research, we analyze these literatures by Bibliometrictic, Citespace and VOSviewer. This method is used to obtain the main research methods and research hotspot in this field[5].
Web of Science is an online subscription-based scientific citation indexing service originally produced by the Institute for Scientific Information, later maintained by Clarivate Analytics[6], that provides a comprehensive citation search. It gives access to multiple databases that reference cross-disciplinary research, which allows for in-depth exploration of specialized sub-fields within an academic or scientific discipline[7].

3. Analysis and results
The data of our research were downloaded from The Web of Science Core Collection. We used “artificial intelligence” and “cyberspace security” as the retrieval term to conduct subject retrieval in The Web of Science Core Collection. The literature type was limited to article and review. The Years is set for 1999-2018. Then we got 1,043 articles. Then, we analyze the article from the following four aspects:

3.1 Quantitative analysis
The change of the number of papers can reflect the development trend of a field[8]. Through bibliometric analysis, we make a total number of documents in recent years, and we find that the total volume of the literature is exponentially higher. The detail of distribution of papers is shown in figure 1.

![Figure 1. The amount of literature in each year](image)

Through bibliometric analysis, we found that before 2015, the main country of literatures contribution in this field was the US. After 2015, the contribution rate of Chinese literatures exceeded that of the US. By 2018, the contribution rate of Chinese literatures had reached the highest 42.8%.

In the development process of this field, many technologies are optimization schemes proposed for a specific problem and in-depth work carried out after inheriting the experience of predecessors. There are also numerous research citations and collaborations between countries. After analyzing the literature, we drew the following article citation diagram in figure 2.
Among them, the US has the highest citation rate, and most of the literatures have been cited many times by other countries, which can reflect that the research points of the literatures produced by the US are representative to some extent, so they are widely and deeply studied in the research field.

### 3.2 Co-citation analysis

Co-citation analysis can reflect the research emphasis of researchers in a certain direction in this field. The more times one paper is cited, the higher its academic value will be. Through the literature clustering function in citespace, we adjusted the clustering function and pruning operation to ensure the maximum simplicity of the obtained results[9]. The result of cluster analysis is shown in figure 3.

![Figure 3. The results of co-cited cluster analysis](image)

After simplified screening of co-cited analysis results, we can find that the literatures appearing in the center of clustering results are all expanded around "model", "recognition", "system", etc.

Similarly, we can find the most cited literatures, which can reflect the basic theory of the current research in this field. The highly cited literatures in this field are listed as follows:
Table 1. The most cited literatures

| Literature name                                                      | Published time | Cited the number | The paper source       |
|---------------------------------------------------------------------|----------------|------------------|------------------------|
| Deep learning                                                       | 2015           | 40               | NATURE                 |
| LIBSVM: A library for support vector machines                       | 2011           | 36               | ACM T INTEL SYST TEC   |
| ImageNet Classification with Deep Convolutional Neural Networks     | 2012           | 37               | ADV NEURAL INFORM PR   |
| Deep Residual Learning for Image Recognition                        | 2016           | 20               | PROC CVPR IEEE NEURAL NETWORKS |
| Deep learning in neural networks - Department of Economics          | 2015           | 20               | IEEE NEURAL NETWORKS   |
| The use of computational intelligence in intrusion detection systems: A review | 2010           | 18               | APPL SOFT COMPUT       |
| A survey of data mining and machine learning methods for cyber security intrusion detection | 2016           | 17               | IEEE COMMUN SURV TUT   |

The most frequently cited articles in the above table are all about the implementation methods of artificial intelligence, such as machine learning, support vector machine and other modified machine algorithms, which indicates that most of implementation approaches needed in this field are inseparable from these basic artificial intelligence algorithms. At the same time, there are also representative literatures that use artificial intelligence to solve practical problems in the field of network security and related literatures that can use artificial intelligence to analyze.

3.3 Keywords co-occurrence analysis

Hot keywords that appear repeatedly in the literatures can also reflect popular topics in the field in some way. VOSviewer software was used to conduct keyword co-occurrence analysis on the literature, and the keyword selection node was set as 4, which could reduce the emergence of vocabulary with low importance, speed up the literature analysis, and extract more representative hot keywords[10]. After removing duplicate references from CiteSpace and importing them into the VOSviewer in the specified format, keywords are generated for clustering, and density graphs of their occurrence frequency are generated, as shown in figure 4.

According to the annual distribution of keywords, it is possible to study which keyword information is most favored by researchers in each time period. Through citesepace’s time zone clustering function, it is possible to visualize the change in the number of citations of certain keywords, as shown in figure 5.

From the above analysis, we can draw the following three preliminary conclusions:

- The most widely used keyword is machine learning. As a major algorithm, machine learning is applied to various fields related to cyberspace security.
In the application of artificial intelligence, there are a wide range of keywords including prediction, classification, system, etc. These keywords are practical problems that can be solved by artificial intelligence, but the practical problems can be solved by artificial intelligence are more extensive, which indicates that in the security of cyberspace, the related technologies of artificial intelligence still have more room for development.

In the field of cyberspace security, there are a wide range of keywords: intrusion detection, anomaly detection, network security, malware detection, privacy protection, etc., indicating that in the current field of cyberspace security, the use of artificial intelligence technology is still limited[11].

In general, there are still few existing artificial intelligence technologies that can be applied in the field of cyberspace security. There are still many difficulties in the field of traditional cyberspace security that have not yet applied artificial intelligence technologies, and they are also constantly advancing each other in the verification and implementation[12].
3.4 Research hotspot analysis

Citespace exports the data report after clustering the co-cited documents. When using the explosion point discovery function, the cluster frequency index can be obtained[13]. This indicator reflects the frequency of occurrence of a hot spot, that is, the heat of research hotspots. The result is shown in Table 2.

| Hotspot                     | Frequency indicator | Hotspot                      | Frequency indicator |
|-----------------------------|---------------------|------------------------------|---------------------|
| industrial mobile-iot network | 0.4                 | intelligent wireless network | 0.15                |
| malware threat              | 0.4                 | cloud computing              | 0.15                |
| pluvial flood               | 0.16                | fast detection method        | 0.15                |
| defensive technique        | 0.15                | zero-day malware detection   | 0.09                |
| security threat             | 0.15                | iot security technique       | 0.05                |
| comprehensive survey       | 0.15                | cyber-physical attack        | 0.05                |

Table 2 show that the current research hotspots mainly include the Internet of Things network, malware detection, cloud computing security, etc. With the development of the times and technologies, artificial intelligence has had a huge impact on cyberspace security. It not only provides algorithms such as machine learning to solve traditional problems, but also redefines the major proposition of what cyberspace is, expanding the domain area for implementation.

4. Conclusions

In our research, we downloaded a total of 1,043 articles from The Web of Science Core Collection. Then we used Citespace and VOSviewer conducted paper distribution analysis, co-citation analysis, keyword analysis and hotspot analysis. We can preliminarily draw the following conclusions:

- According to the distribution and growth of literatures, the number of literatures in recent years presents a "blowout" growth. According to the development situation, it can be predicted that the artificial intelligence technology will continue to develop in the field of cyberspace security.

- From the results of the literatures co-citation, in the field of cyberspace security, simple artificial intelligence algorithms are used to replace traditional algorithms, and data mining is used to analyze, such as network traffic data.

- From the co-occurrence results of keywords, The machine learning method still dominates, and the new algorithm can bring new solutions to the problem of cyberspace security. The literatures on cloud computing security, Internet of Things security, and big data security has begun to appear and is increasing. It can be seen that research in this field will be dominated by these points in the future, focusing on security research on new intelligent applications.

In general, the application of artificial intelligence technology to the field of cyberspace security is still in its infancy. There are different solutions to specific security issues, and there are still many opportunities and challenges.

For future work, we're going to do some more in-depth and practical research, such as do some research about Cloud computing security and IoT security.

References

[1] Xinhuanet. (2010) Artificial intelligence and big country strategy. http://www.xinhuanet.com/globe/2017-03/29/c_136168263.
[2] Guo Wangshu, WangHuimei, XianMing, Ye Tian. (2019) Research on "Cyberspace Security Testing and Evaluation" Technology Development Trend. In: CISCE.2019. Haikou, China. pp. 363-367.

[3] He Jiawei, Wang Huimei, Xian Ming, Liu Jian. (2019) The Application of Big Data in Cyberspace: A Survey. In: CISCE.2019. Haikou, China. pp. 570-574.

[4] Tan Yejin, Liu Jian, Wang Huimei, Xian Ming. (2019) The Development Trend Analysis of 5G Network. In: CISCE.2019. Haikou, China. pp. 382-385.

[5] Lv Wenjing, Xu Li, Liu Jin, Chen Jin. (2018) Ten Years Review of Artificial Intelligence Research in China——Based on Bibliometrics and Knowledge Mapping Analysis from 2008 to 2017. J. technical economy, 37(10):73-78+116.

[6] Analytics, Clarivate. (2017) Acquisition of the Thomson Reuters Intellectual Property and Science Business by Onex and Baring Asia Completed. www.prnewswire.com.

[7] Drake, Miriam A. (2004) Encyclopedia of Library and Information Science. New York, N.Y.

[8] Shengmin Guo, Li Wang, Yujie Xie, Xi Luo, Shaojun Zhang, Linbo Xiong, Haibo Ai, Zhihao Yuan, Jianxiong Wang. (2019) Bibliometric and Visualized Analysis of Stem Cells Therapy for Spinal Cord Injury Based on Web of Science and CiteSpace in the Last 20 Years. J. World Neurosurgery.

[9] Zhang Xingzhi, Yang Chao. (2019) Bibliometrics and Knowledge Mapping Based on CNKI and CiteSpace. J. Contemporary Library, 2019(09): 12-17+64.

[10] Song Xiufang, Chi Peijuan. (2016) A Comparative Study of Vosviewer and Citespace Applications. J. Information Science, 2016, 34(07): 108-112+146.

[11] Zhang Yang. (2018) The impact of artificial intelligence on cyberspace security. J. China Information Security, 2018 (05): 65-66.

[12] Ma Yizhen. (2019) Application of Artificial Intelligence in Network Security Defense System. J. Computer and Network, 2019, 45(15): 48-49.

[13] Zeng Wei, Li Jianqiu. (2019) The Hot Spot and Frontier Analysis of Domestic TRIZ Research Based on CiteSpace. J. Science and Technology Management Research, 2019, 39(18): 260-265.