Research on the Application of Artificial Intelligence in Cyber Range

Jiujiang Han\textsuperscript{1\!*}, Ming Xian\textsuperscript{1}, Jian Liu\textsuperscript{1} and Huimei Wang\textsuperscript{1}

\textsuperscript{1}College of Electronic Science and Technology, National University of Defense Technology, Changsha, Hunan, 410073, China

\*Corresponding author’s e-mail: hanjiujiang20@nudt.edu.cn

Abstract. In recent years, innovative research and application promotion of new theories of cyberspace security represented by cyber range have gradually emerged around the world. With the in-depth development of artificial intelligence technology, cyber range has begun a new wave of intelligent applications. Currently, there are relatively few literatures on the development trend of artificial intelligence in the field of cyber range. Therefore, we use the knowledge map to visually analyze the 993 related articles downloaded from WOS, including article quantitative analysis, co-citation analysis, and keyword co-occurrence analysis. Finally, we obtained the development prospects, important research achievements and research hotspots in this field.

1. Introduction

With the rapid development of technologies and applications such as the Internet and artificial intelligence (AI), cyberspace security has become more and more important in our lives\cite{1}. The degree of cyber threats is becoming more and more widespread, bringing huge security risks to personal life, enterprise production and even national key infrastructure.

In order to gain cyberspace security advantages, countries all over the world have launched cyberspace security strategies and initiated national cyber range construction plans. The cyber range is a network information security scientific research test platform or information system aimed at the field of information security offense and defense, covering typical application scenarios in the field.

From the beginning of the "Cyber Manhattan" project in 2008 to the present, it has been 13 years for the cyber range from concept to implementation. Starting in 2020, the cyber range has entered the intelligent stage. Intelligent development is inseparable from artificial intelligence technology. The landing and promotion of artificial intelligence technology in the security field has become a new wave. In the United States, where the development of cyber range is relatively advanced, some manufacturers have begun to research and promote the intelligent application of cyber range. Therefore, it is of great significance to study the development trend of artificial intelligence in the field of cyber range.

As a branch of information science, bibliometrics uses technologies such as databases and knowledge graphs to summarize development trends in some fields through literature analysis. There are many examples of using this method to study and summarize a field\cite{2-4}. This article finally retrieved 993 relevant documents from the Web Of Science (WOS) as a data source for analysis, reviewed and summarized the research results and development trends of artificial intelligence in cyber range, and provided references for the future theoretical research and intelligent development of cyber range.
2. data sources and research method

2.1. data sources
The research data in this article comes from the Web of Science Core Collection database of WOS. We use "artificial intelligence" and "cyber range" as topics for searching. The type of literature is limited to articles and reviews. The time range is set to 2002-2021. Then we got a total of 993 articles.

2.2. research method
This paper uses the literature analysis and visualization software Bibliometric, CiteSpace and VOSviewer to analyze the related literature on the application of artificial intelligence in the cyber range field to obtain related research progress, research hotspots and frontier trends.

Among them, this article imports the downloaded article information into the Bibliometric online platform, and analyzes the total amount of sub-countries and the cooperation relationship between countries in the analysis of total literature and cooperation relationship analysis module. After importing the information of the article in CiteSpace software, format conversion and document deduplication are performed. The node selection is co-citation analysis, the threshold is set to TOP50, the cutting method is selected path search algorithm (Pathfinder), and path network simplification (Pruning Sliced Networks) and merged network simplification (Pruning the Merged Network). In the VOSviewer software, we analyze the key words of the literature. Similarly, we import the literature data into the VOSviewer software, set the minimum number of occurrences of a single keyword to 5 times, get 265 keywords, and generate a keyword density map to show the clustering relationship between keywords.

3. Analysis and results

3.1. Quantitative analysis
Changes in the number of published articles can reflect the development trend of a field. Through the Bibliometric online platform, we have made statistics on the total number of articles in recent years by country, and found that the total number of articles has roughly increased exponentially. The year distribution characteristics of articles derived from the WoS database are shown in Figure 1. According to the time-year distribution and change of the number of articles, this article divides the cyber range's intelligence into three stages: The first stage is the basic construction stage of cyber range (2002-2014), and the number of papers published in this stage is relatively small. The development of theory is slow. The second stage is the practical training stage of cyber range (2015-2019). The number of articles in this stage is increasing exponentially, which is an important transitional stage, which provides a solid research foundation for the intelligent research of cyber range. The third stage is the intelligent development stage of the cyber range (after 2020), with the highest research enthusiasm and the largest number of publications.

![Figure 1. Changes in the number of articles over the years](image)
Through Bibliometric analysis, we found that before 2016, the United States was the main country where articles in this field were published. The amount of publications shows that the United States pays more attention to the research on the intelligentization of the cyber range, the research is more in-depth, and certain results have been achieved. After 2020, China will publish more articles annually than the United States.

In the process of scientific research in this field, academic exchanges have shown a global trend. Many specific results are based on the inheritance of previous experience, and put forward specific optimization schemes for specific problems. The production of new results often requires the collaboration of scientific researchers from various countries. After analyzing the literature, we get the citation map between countries, as shown in Figure 2.

![Bibliographic citations between countries](image)

Figure 2. Bibliographic citations between countries

Among them, the literature in the United States has the highest citation rate, and most of the literature has been cited many times by other countries, which can reflect that the United States has a broad and in-depth research foundation in this research field.

3.2. Co-citation analysis
Co-citation analysis of an article refers to the phenomenon that two references are cited by the same article. By analyzing the clustering and key nodes in the co-citation network, the knowledge structure of a certain research field can be revealed. Using the clustering function of CiteSpace to analyze the co-citation of articles, we can mine the common themes of similar articles[5]. The result of clustering is shown in Figure 3.
After screening the results of the co-citation analysis, we can find that the articles appearing in the center of the clustering results are all based on topics such as "cyber-physical systems", "model", and "internet".

Through analysis, we have found the most cited documents, which can be used as the basic theory of the current research in this field. The widely cited documents in this field are as follows:

| Literature name                                                                 | Published time | Cited the number | The paper source                                      |
|--------------------------------------------------------------------------------|----------------|------------------|-------------------------------------------------------|
| A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems | 2015           | 39               | MANUFACTURING LETTERS                                  |
| Scaling Learning Algorithms Towards AI Cyber-physical systems in manufacturing  | 2015           | 23               | NATURE                                                |
| A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection | 2016           | 22               | CRIP ANN-MANUFACT TECHN                               |
| Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications | 2015           | 14               | IEEE COMMUNICATIONS SURVEYS & TUTORIALS                |
| Intelligent manufacturing in the context of industry 4.0: a review                | 2017           | 13               | IEEE COMMUNICATIONS SURVEYS & TUTORIALS                |

Figure 3. The results of co-citation cluster analysis
The most cited articles in the above table are all about Cyber-physical systems of Industry 4.0, as well as the application of artificial intelligence technology in related fields, including algorithms such as machine learning and data mining. This shows that the future development of this field is inseparable from these basic artificial intelligence algorithms. At the same time, there are related articles on the use of artificial intelligence to conduct security and event analysis in cyberspace.

3.3. Keywords co-occurrence analysis

Popular keywords that appear frequently in articles can also reflect research hotspots in this field. Use VOSviewer to perform keyword co-occurrence analysis on articles, and set the keyword node selection to 5, which can reduce the appearance of less important words, speed up document analysis, and extract more representative popular keywords[6]. After removing duplicate references from CiteSpace and importing them into VOSviewer in the specified format, keywords used for clustering will be generated and a density map of their appearance frequency will be 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 year. Through the time zone clustering function of CiteSpace, the changes in the number of times a keyword is cited can be visualized, as shown in Figure 5.

![Figure 4. The results of keywords co-occurrence cluster analysis](image1)

![Figure 5. Keyword outbreak record](image2)
From the above analysis, we can draw the following three preliminary conclusions:

- The most widely used keywords are Cyber-physical systems, machine learning and security. As a mainstream algorithm, machine learning is an important support for the future development of Cyber-physical systems[7]. At the same time, attention must be paid to security issues in cyberspace.

- In the application of the combination of artificial intelligence and cyber range, there are many kinds of keywords, including behavior, architecture, system, classification, etc. These keywords are the basic problems that artificial intelligence can solve in this field, but the practical problems that artificial intelligence can solve are more extensive[8]. This shows that in the cyber range field, artificial intelligence related technologies have very large room for development.

In general, there are still few existing artificial intelligence technologies that can be applied to the cyber range field[9]. The traditional cyber range still has many difficulties that have not yet applied artificial intelligence technology, and they are constantly advancing each other in verification and implementation[10].

4. Conclusions
In our research, we downloaded 993 articles from the the Web of Science Core Collection database of WOS. Then we use Citespace and VOSviewer to analyze the number of literature, co-citation analysis, and keyword co-occurrence analysis. We can preliminarily draw the following conclusions:

- According to the annual distribution of the number of documents and the cooperative relationship between countries, the number of documents in this field has shown a "blowout" growth in recent years. According to the growth situation, it is foreseeable that artificial intelligence technology will continue to develop in the field of cyber range. China and the United States will become major research countries in this field.

- Judging from the results of co-citation in the literature, in the cyber range field, researchers have begun to combine artificial intelligence technology to develop intelligent applications. There are two main directions for the development of artificial intelligence technology in this field. The first is to use machine learning methods to replace repetitive labor through computing power, simplify business processes, and save time and labor costs. The second is to form a set of empirical strategies through data collection and learning, and to provide the most optimized business processing capabilities in the current state according to intent.

- Judging from the results of co-occurrence of keywords, machine learning methods still dominate, and new algorithms can bring new technological changes to the cyber range field. Articles about cloud computing security, IoT security, and big data security have begun to appear and continue to increase. It can be seen that the future research in this field will be led by these points, focusing on the security research of intelligent applications of the cyber range.

Generally speaking, the application of artificial intelligence technology in the field of cyber range is still in its infancy. The intelligent cyber range can not only effectively support cyber combat capability research and cyber weapon equipment verification in the national military field, but also undertake security assurance, scientific research innovation and security services for important civilian industries.

For future work, we will do some more in-depth and more practical research, such as doing some research on the intelligent generation of network scene topology and the intelligent quantification and presentation of cyber range.

References
[1] Xinhuanet. (2010) Artificial intelligence and big country strategy. http://www.xinhuanet.com/globe/2017-03/ 29/c_136168263.
[2] Guo W., Wang H., Xian M., Ye T.. (2019) Research on "Cyberspace Security Testing and Evaluation" Technology Development Trend. In: CISCE.2019. Haikou, China. pp. 363-367.
[3] He J., Wang H., Xian M., Liu J. (2019) The Application of Big Data in Cyberspace: A Survey. In: CISCE.2019. Haikou, China. pp. 570-574.
[4] Tan Y., Liu J., Wang H., Xian M. (2019) The Development Trend Analysis of 5G Network. In: CISCE.2019. Haikou, China. pp. 382-385.

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

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

[7] Qi, L., Chen, Y., Yuan, Y. et al. (2020) A QoS-aware virtual machine scheduling method for energy conservation in cloud-based cyber-physical systems. World Wide Web 23, 1275–1297.

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

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

[10] Zeng W., Lü J. (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.