A dynamic network analysis method for identifying emerging technology

A new round of scientific and technological revolution and industrial transformation is accelerating, and emerging technologies, such as artificial intelligence and block chain, are constantly emerging, which have a profound impact on economic and social development. Identifying emerging technologies becomes an important issue which affects the development strategies of countries, companies and research institutions. However, due to the uncertainty, ambiguity and complexity of emerging technologies, the identification process is full of high difficulty and low accuracy.

Many researchers in bibliometrics filed have paid great attentions to the identification of emerging technologies, and the methodologies in previous studies could be divided into three categories: citation analysis, co-words analysis and the hybrid pattern of them. Recently, network analysis has become the mainstream because of its advantages at quantifying large-scale and complex text data, mining valuable information hidden in the topology structure and presenting dynamic evolution of networks. Link prediction method in complex network provides a new solution to the identification of emerging technologies, which could calculate the possibility of links between unconnected nodes and predict the future evolution trend of networks, and it enables to solve the problems of sparse data of new nodes in network. However, most literatures are limited to the simple application of existing models and prediction indicators, lacking effective combination with specific network characteristics, and the prediction accuracy is difficult to achieve expectations.

In this vein, this paper proposes an emerging technology identification method based on dynamic complex network analysis, which predicts the future evolution trend of dynamic weighted Co-word network by link prediction, so as to reveal the future development trend of technology and solve the problems of fuzziness and uncertainty caused by the unclear technical characteristics of emerging technologies. At the same time, this paper combines machine learning algorithm to learn network multiple information to improve link prediction accuracy. Firstly, based on link prediction index, the local information of nodes, network path information and dynamic random walk information in dynamic weighted Co-word network are obtained. Then, the back propagation neural network algorithm is used to realize the non-linear fitting of the index to predict the future evolution of dynamic network. Finally, according to the definition of emerging technology, an indicator system based on predicted dynamic network is established to identify emerging technologies, which contains multi-dimensional factors—novelty, growth, coherence and impact.

The method includes three steps：

1）　**Step1: Data acquisition and dynamic network construction**

Focusing on a specific technical field, literature data is obtained from Web of Science (WOS), and technical terms are extracted and cleaned by using "Term clumping" topic cluster method. Then, the cleaned technical data set is divided into several sub-datasets according to the time sequences, and a

dynamic weighted Co-word network is constructed.

2) Step2: Link prediction with dynamic network

Link prediction method is used to obtain the local, path and random walk characteristics of dynamic network nodes, and BP neural network is used to train the link prediction model to achieve the non-linear fitting of these three indicators. The trained model is used to predict the future evolution of the network, and then reveal the future trend of the technology field. Specific steps include link prediction index selection, link prediction model training and future network prediction.

3) Step3: Identification of emerging technology

Based on the predicted dynamic network, the technical terms are clustered into multiple technical topics by using community discovery algorithm. An indicator system based on predicted dynamic network is established to identify emerging technologies, which contains multi-dimensional factors—novelty, growth, coherence and impact. Specific steps can be divided into technology topic acquisition and emerging technology topic recognition.

An empirical study on information science is conducted to demonstrate the reliability of the methodology. This paper downloaded 8285 journal papers from 9 journals in the field of information science from WOS database from 2009 to 2016. Nine journals are: Scientometrics, JASIST, Information Research, Journal of Informetrics, Information Processing & Management, Journal of Documents, Journal of Information Science, Library & Information Science Research, Research Evaluation.

The innovations of this research are as follows:

1) Dynamic network analysis method is introduced in the process of emerging technology identification. Compared with traditional static network analysis method, our study considers the development and changes of technology topics.

2) For link prediction method, based on dynamic weighted network, this paper synthetically considers three kinds of network structure information: local network, path (global) network and dynamic random walk network. Combining with machine learning method, the prediction model is trained. Compared with single prediction index, the prediction accuracy is significantly improved.