

# Research on International Relations Network based on Text Mining

Gan Tan<sup>1</sup>, Mi Li<sup>2\*</sup>

<sup>1</sup>ZhaoQing university, <sup>2</sup>Hanyang University

\*2021010013@zqu.edu.cn

**Abstract:** With the rapid development of communication and information technology, people's information exchange and dissemination methods are becoming more and more convenient, which not only expands the depth of global information exchange, but also closely links countries in the world together, forming a huge and complex network of international relations. In this context, the formal method of relationship network is used to study international relations and other non-formal issues. Although a lot of research results have been obtained, there are also many research problems. Therefore, some scholars propose to use text mining method to study international relations network. On the basis of understanding the current research status of international relations network and combining the basic theory of information-knowledge-intelligence, this paper deeply discusses the international relations network with text mining as the core. The final experimental results show that the parsing number model and the conditional random field model achieve excellent results.

**Keywords:** Text mining; International relations; Network of relationships; Relationship between characteristics

## 1. Introduction

The research of international relations network refers to the transformation of the relations among countries into a network structure and the identification of the development trend of international relations on the basis of the research of network. With the continuous development of social economy and science and technology, countries in the world have become more and more closely connected, and a diversified network of relations has been formed. Therefore, the study of international relations network has become the main issue for scholars in various countries to discuss. Through the use of formal technical means, it is the basic connotation of the international relations network research to understand the non-formal activities of human beings with complex significance. Since formal technical means can only understand the situation change, and human activities are abstract and non-formal, it is difficult for traditional formal means to deal with abstract content directly, but they can be understood through human sensory perception and thinking organs. [1.2.3]Professor Zhong Yixin put forward the first kind of information conversion principle in his research. Combining with the principle analysis shown in Figure 1 below, it can be seen that the known content of formal understanding is mainly divided into three parts: first, the sensing system can transform external stimuli into formal grammatical information; Secondly, the grammatical information or knowledge base can be used to retrieve the relevant pragmatic information, or the pragmatic information related to the grammatical information can be obtained by performing relevant operations. Finally, as long as both grammatical information and pragmatic information are satisfied, semantic information can be obtained. This theory proves that as long as the system has clear goals, bright knowledge and simple logical deduction ability, then it can complete the transformation from form to content.[4.5.6]

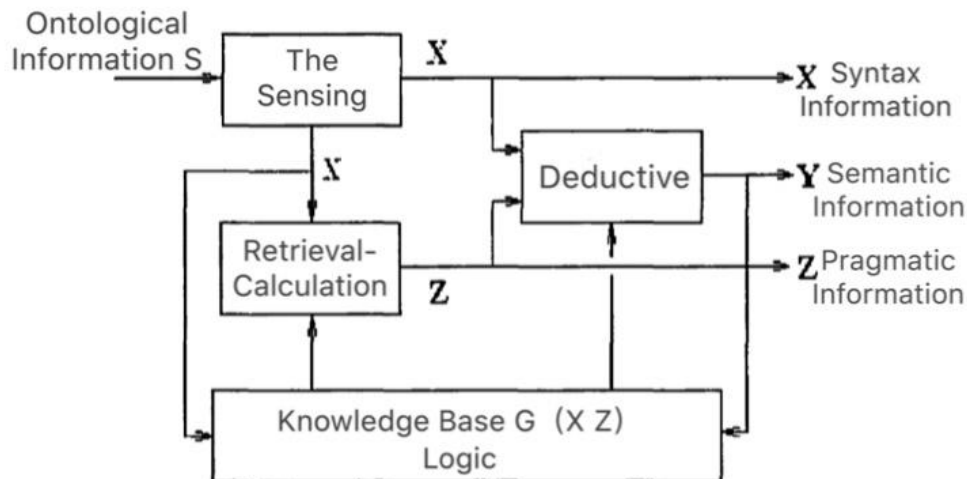


Figure 1 Schematic diagram of the first type of information transformation

In the comprehensive promotion of Internet technology, earth-shaking changes have taken place in human life forms and production activities. As the theory of network technology becomes more and more mature, the interaction between people and the Internet becomes more and more close. The active participation of the public in the Internet makes the information on the Internet explode and grow at an unprecedented speed. The number of web pages in China reached 86.6 billion by the end of December 2011, up 44.3 percent from the previous year, according to a survey by the China Internet Network Information Center. Nowadays, the number and content of web pages on the Internet are characterized by a huge amount of information, which is incomparable to any expert. If these contents can be fully used, a better, more timely and more reliable international relations network can be built. The key content is how to extract valuable content from massive information on the Internet and build a high-quality international relations network. Some scholars have put forward text mining technology in practice. In the traditional sense, people mainly use manual way, annotate and index information, and provide manual services to obtain the required information for users, so as to transform information into knowledge. Although this time-consuming and laborious method can deal with less information, it is difficult to achieve the desired effect when dealing with the explosive growth of Internet information. Therefore, scientists propose to use machines to automatically process massive text information, and clarify the application content of text mining technology. From the perspective of practical research, text mining originated from the traditional sense of data mining, mainly processing unstructured or semi-structured text information, really realize the effective transformation from information to knowledge. If information cannot be distilled into the knowledge that the cognizer understands and applies, then information cannot function. After understanding the unified theory of information, knowledge and intelligence, this paper focuses on how to apply text mining technology to international relations network, fully considering the characteristics of international relations network, and make innovative research on text mining algorithm, so as to mine more valuable information.[7.8.9]

## 2. Method

### 2.1 Text mining technology

This technique, also known as textual knowledge discovery, refers to the discovery and mining of valuable knowledge from text collections. In the theoretical research of modern technology, text mining is a highly borderline cross field. The relevant theoretical knowledge is derived from the traditional data mining, which includes the research content of machine learning, natural language processing, artificial intelligence and other fields. Compared with traditional data mining technology, text mining technology can deal with different objects, but it is difficult to operate

because there is no definite form. Generally speaking, the text mining process is divided into two parts, one is text preprocessing, the other is result mining, as shown in Figure 2 below:

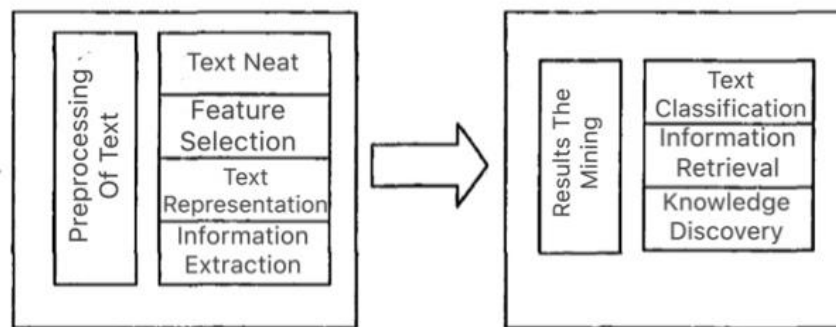


Figure 2 Flowchart of text mining process

## 2.2 Network of International relations

By transforming the relationships of multiple entities in international relations into networks and using network analysis to study the development and change of international relations, we can provide an effective basis for the economic construction and development of various countries. The research method of international relations network is put forward based on the development of communication and information technology. The reason is that in ancient times, due to the limitations of distance and technology, the communication between countries is relatively less, and the influence between each other is not big, so the use of network method to study international relations is not paid attention to. However, in the new era, with the continuous development of Internet technology, countries around the world are closely connected. They should comprehensively study international relations and build a network organization structure with international relations as the core, which can lay the foundation for communication and exchanges among countries. At the same time, the development of international relations network research is closely related to the development of social network analysis.

Nowadays, when studying the network of international relations, the following procedures are mainly carried out: first, determine whether the problem to be solved can be dealt with by network method; Secondly, it is necessary to determine the network framework, whether it is a relational network or a subordinate network, to effectively define the network architecture. Again to search for data information needed to provide research, in the data, determine the method of the point river, the algorithm of the relationship between the measurement; Finally, the network relationship is studied to obtain effective research conclusions.[10.11]

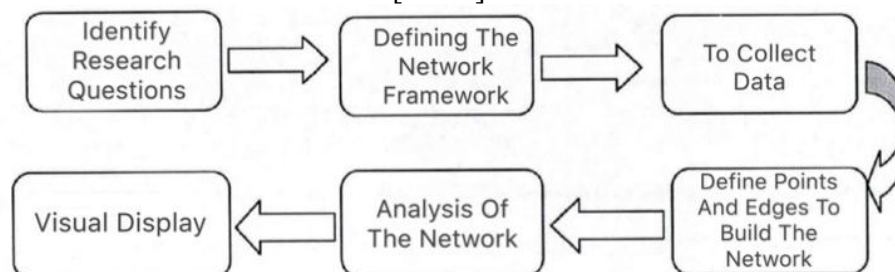


FIG. 3 Flow chart of international relations network

## 2.3 International relations network system analysis based on text mining

On the one hand, since there is no comparison and reference for the construction of international relations network from the text, the initial system scheme is shown in Figure 4 below:

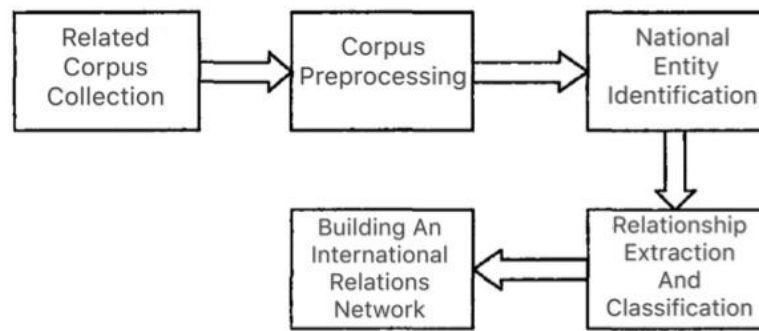


FIG. 4 Schematic design of the system

Based on the above analysis, we can see that the overall operation involves the following points: First, search for corpus information related to international relations to ensure the reliability and standardization of relevant information, which can fully demonstrate international relations; Secondly, the corpus information should be preprocessed to identify the country entities in anticipation and extract the content related to the international relationship network. Thirdly, it is necessary to determine the relationship between countries, extract the corresponding characteristics, and clarify the practical significance of the research work according to the essential characteristics of international relations. Finally, after effective classification according to relationship features, the network is constructed by obtaining points and edges.

On the other hand, according to the analysis of the improvement scheme, compared with the original system proposed above, researchers proposed to use fuzzy reasoning for classification processing, and combined with support vector machine method to complete the relation classification of fuzzy reasoning. At the same time, when understanding and analyzing the text contents, they will be divided into different aspects of international relations for the convenience of users' independent judgment. In addition, the comprehensive way is used to conduct information fusion at the language level, so as to obtain the evaluation results of the nature of relations.

### 3. Result analysis

On the one hand, text classification based on support vector machine. This algorithm follows the principle of structural risk minimization. After comprehensive evaluation of the fitting effect of samples and the complexity of the model, the convex quadratic programming optimization algorithm is used to avoid multi-solution problems. In the experimental analysis, it is assumed that the feature selection algorithm is Chi-test, and the weight calculation method is tf. The comparative analysis is made on the effects of various sum functions in different classification systems, and each sum function is searched to ensure the optimization of parameters. After 2196 news headlines are obtained, 10% of them are selected as the test set, and the remaining 90% are the training set. Finally, the test results as shown in Figure 5 below can be obtained:

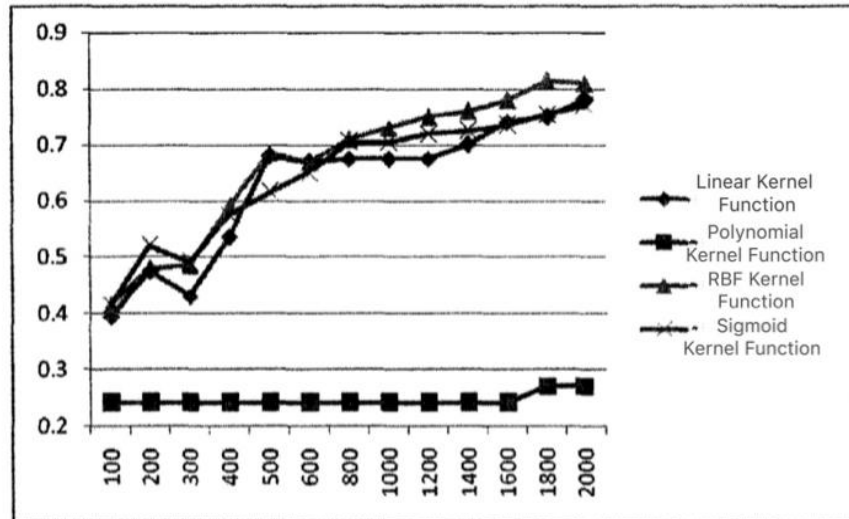


Figure 5 Classification results of news categories

Combined with the analysis of the above figure, it can be seen that the classification accuracy will also improve when the training samples continue to rise. When the number of training reaches 2000, the accuracy curve has not changed greatly, which proves that the actual training has been very sufficient. In each classification system, the best accuracy is close to or above 0.8. At the same time, the performance of linear kernel function and polynomial and function is average, while polynomial and function have serious classification deviation. Therefore, the kernel function and optimal parameters set for each classification system are shown in Table 1 below:

Table 1 Optimal sum function and parameter design results

News category	Semantic credibility	Words and Actions
RBF kernel function	Sigmoid kernel function	RBF kernel function
C=16.0	T=3 C=32 G=0.125	C=22.627416998
G=0.176776695297		G=0.125

On the other hand, fuzzy inference of relation classification. According to the algorithm flow chart shown in FIG. 6 below, in this verification experiment, the application effect of the classification algorithm centering on the relationship between countries is mainly discussed. The specific results are shown in Table 2 below:

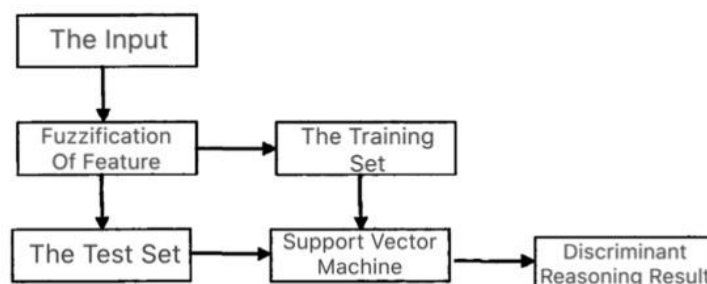


FIG. 6 Flowchart of fuzzy inference algorithm

	Linear kernel function	Polynomial kernel function	RBF kernel function	Sigmoid kernel function
Accuracy before fuzzification	65.4364%	65.0485%	67.767%	53.4277%
Accuracy after fuzzification	72.233%	77.0874%	73.3981%	69.5146%

Table 2 Comparative analysis of classification results

Based on the above analysis, it is found that considering membership as the main feature, more perfect classification results can be obtained. Compared with the data analysis results before fuzzification, it is found that the application of fuzzy inference algorithm can further improve the

operation performance. In the four sum functions proposed in this paper, the actual accuracy increases from 5.6311% to 16.0869%. The results show that the improved classification algorithm based on fuzzy inference relation meets the requirements of international relations network research. Explosive growth trend in the modern network information conditions, will be closely linked to the world, to build international relations network, deep excavation of the text information communication between countries, can help the world solve the problem of economic development and population development, and can guide countries in innovation explore together to achieve sustainable development goals.

#### 4. Conclusion

To sum up, according to the text mining technology proposed in this paper, the research of international relations network can be comprehensively discussed, which can provide new ideas for the current research of related topics. Therefore, on the basis of integrating the theory and application experience of text mining technology, scholars from all over the world should continue to explore the division method of international relations, so as to put forward a more comprehensive and reasonable relationship division system. At the same time, it is necessary to expand the application space of Internet news through the partial limitation of data acquisition by search engines, so as to get more valuable data information in the subsequent research and development, help researchers to design parallel processing algorithms, and improve the efficiency and quality of information acquisition by the system. In addition, it is necessary to strengthen the training of professional and technical personnel, change the traditional research ideas of international relations network, and fully demonstrate the application value of text mining technology. Only in this way can countries in the world develop good relations.

#### Reference

- [1] Juanjuan Liu, Longyue Liang, Xuanye CAI. Research on credit fraud based on text Mining technology [J]. Intelligent Computer and Applications, 2022, 12(7):8.
- [2] Zhenpeng Li, Bizhen Chen, Jingyu Luo. Research on online Public opinion classification based on text Mining [J]. Journal of Systems Science and Mathematics, 2020, 40(5):14.
- [3] Jiangyue Zhou, Laixin Zhang, Qingyang Feng. Review Description Analysis of video card based on text mining [J]. China Broadband, 2021(5):1.
- [4] Haituo Qie, Zhijuan Zhang. Research on quantitative evaluation of Chinese science and technology innovation policy based on text mining [J]. Global Outlook on Science and Technology Economy, 2020, 35(6):9.
- [5] Tong Liu, Shaoyan Li, Xueping Gu, et al. Research on auxiliary Decision method of power grid maintenance work ticket based on text Mining [J]. Electrical Measurement and Instrumentation, 2020, v.57; No.727(02):44-50.
- [6] Fangyu Liu, Jing Zhao, Ze Li, et al. Construction and empirical study of Beijing family doctor evaluation system based on text mining method [J]. Chin J General Practice, 2020, 23(25):4.
- [7] Mengxin Liu, Yuewen Li. Cognitive sentiment analysis of social public towards artificial intelligence based on text mining [J]. Progress in Applied Mathematics, 2022, 11(9):8.
- [8] Linjie Jiang, Ling CAI, Jun Zhao. Design of group Complaint Information Analysis System based on Text Mining [J]. Information and Computer, 2022(012):034.
- [9] Ya Xi Yang, Rong Ji Mu, Huan Zhang, et al. Research on industrial poverty alleviation policy based on text mining [J]. Inner Mongolia Science, Technology and Economy, 2021(11):4.
- [10] Nuoyi Chen, Jing Shan, Jiaying Wang, et al. Sentiment analysis and research of User comments in Hongmeng System based on text mining [J]. Science and Technology Information, 2021, 19(16):4.
- [11] Yuxuan Xin, Xiaodong Wang. Research on Sentiment Analysis of Movie Reviews based on text Mining. Journal of Mudanjiang Normal University: Natural Science Edition, 2021(1):4.

- [12] Shihao Lu, Yun Zhu, Xiwen Lu, et al. Research on Fault Diagnosis Method of Hydropower Station Based on Text Mining Technology [J]. Large Electric Machine Technology, 2022(000-004).
- [13] Xiaoyu Meng. Analysis of research hotspots in artificial intelligence based on text mining technology [J]. Statistics and Application, 2022, 11(1):10.
- [14] Feifei Zhou, Guanghai Zhang. Research on perceived needs of museum tourists based on text mining [J]. Journal of Jingdezhen University, 2022, 37(2):5.
- [15] Li Yuan. A comparative study of posts in Internet enterprises based on text mining [J]. Modern Information Science and Technology, 2022, 6(7):5.