Research on International Communication Effects of Chinese Culture Based on Web Search Data

Jingyi Wang[1], Chong Zhang[2]*
xk18wjy@126.com[1], zhangchong@blcu.edu.cn[2*]

Department of Information Science, Beijing Language and Culture University, Beijing, China[1],
Department of Information Science, Beijing Language and Culture University, Beijing, China[2]

Abstract. The study constructs data indicators that showing the effect of international communication of Chinese culture based on Internet search data. Starting from the three-stage communication effect of audience cognition, attitude, and behavior, this study uses Google trend search data to build cultural attention indexes for material, relationship, and concept layer of Chinese culture. The results show that the international communication of Chinese culture has achieved certain effects in worldwide. The attention in different civilization circles is generally increasing, but the communication effects of different layers of culture under different civilization forms are quite different.

Keywords: International communication effect of Chinese culture; network search data; attention index

1   Introduction

In digital age, more and more human activities have been continuously integrated and developed with Internet technology. The Internet has become one of the main sources for people to obtain information. Network behavior data reflects user's degree of attention to certain topics. Massive web search data is an important way to map people's economic and social behavior in social science researches. Network behavior data not only reflects the user's attention to certain topics, but also reflects users' opinions on certain topics. Public attention is about what people think, while public opinion is about how people think. For example, in the process of cultural dissemination, overseas audiences learned through certain channels that China has a traditional festival called the Spring Festival, and he may search for festival customs which related to the Spring Festival. Demand-based search behavior reflects people's attention to certain topics. This paper builds data indicators to show the effect of Chinese cultural communication using Internet search data. This research starts with the three-stage communication effects of audience cognition, attitude and behavior, and uses the search data of Google Trends to construct hierarchical cultural attention indexes for the material, relationship and conceptual layers of Chinese culture. This paper uses Internet data to construct related indexes and monitors the effect of Chinese culture "going in" from the perspective of big data. This data can provide data support and decision-making references for relevant domestic departments and institutions. Main contents of this paper are as follows:

(1) Using Google Trends value as the public attention to Chinese culture to explore whether the spread of Chinese culture has cultural differences. Select Chinese and English benchmark keywords of different layers of culture and expand them to multiple languages.

(2) Expand multilingual benchmark keywords by KeyWordTool and Google related queries. Obtain monthly search data of all multilingual keywords from 2004 to present on Google Trends by python crawler.

(3) Expound the changing of public attention of Chinese culture in different civilization backgrounds. Analyze whether there are differences in its spread.

(4) Based on the above results, provide relevant suggestions for further improving Chinese cultural international communication capabilities.

The use of online behavior data to measure public attention has been gradually developed. The massive data contained in search engines such as Google Trends and Yahoo have been used to measure public attention or predict related events in the fields of economics, medical treatment, and communication. For example, in the medical field: Ginsberg used Google Trends search data to predict epidemics [1], Ripberger used search data related to epidemic diseases to track the spread of infectious diseases [2], Chew C proved the potential of using social media for public health "information meteorology" research [3]. In the field of economics, some researchers used 98 key words in the financial field to predict stock market fluctuations [4]. In the field of communication, Scharkow M proposed that search engines can be used as a powerful supplement to traditional voting methods, representing a new aspect of the public agenda[5]; Chen Yunsong used Google Books big data to gain a 300-year international reputation for Chinese cities Measure the changes and characteristics of the public [6]; Meng Tianguang explored the public search behavior and public policy attention based on the Baidu index, and use search data to explore public issue attention and policy preferences [7]. The emerging research methods in the era of big data provide opportunities and soil for measuring public attention and attitudes. In general, the current domestic use of big data to measure and interpret the effects of Chinese culture in international communication is still lacking. Although a few scholars have tried to evaluate the effect of Chinese culture "going in" from the perspective of the audience, existing evaluation studies have problems such as single data sources, strong subjective factors, and one-sided research perspectives. Based on the studies above, this study uses Internet data to construct related indexes, and monitors and evaluates the effect of Chinese culture "going in" from the perspective of big data. This data can provide data support and decision-making reference for relevant domestic departments and institutions, and assist its actual dissemination effect.

2   DATA COLLECTION AND PROCESSING

2.1   Keyword selecting and processing

Based on the three-layer theory of the concentric circle model of Huang Yunhong [8], the paper divides Chinese culture into material layer, relational layer, and conceptual layer. In order to make the selection of initial keywords more comprehensive, the paper further selects costume culture, food culture, medical culture, architectural culture and scientific and technological culture as the secondary indicators of the material level; selects custom culture, festival culture, economic culture, political culture and legal culture as the secondary indicators of the relationship level; selects art culture, diplomatic culture, religious culture and text culture as the secondary indicators of the conceptual level. On the basis of the secondary indicators, 5-7 benchmark keywords are determined for each secondary indicator and expanded to Russian, Japanese

and Korean. Search users have the characteristics of heterogeneity, and they may pay attention to a certain thing from different perspectives. Therefore, the search keywords should have the characteristics of diversity. It should be a collection that can contain most of the commonly used words related to the benchmark index. To ensure the comprehensive and extensive selection of keywords, the Keyword Tool and Google Trends related queries are used to expand the benchmark keywords. Filter the expanded keywords and delete the words with low search volume, repetition and irrelevant content. The final keyword. Some results are shown in Table 1.

2.2 Keyword data collection

This study uses Python crawler to collect the monthly search data of multilingual keywords from January 2004 to February 2021. The English results are shown in Table 2.

Table 1. Benchmark keyword expansion

Benchmark keywords	Expand keywords
Chinese clothing	China clothing, Chinese clothes, Chinese online clothing, Chinese dress, clothing in china, ancient Chinese clothing, Chinese clothing store, Chinese fashion, Chinese new year clothing, Chinese wear
Chinese traditional clothing	Chinese traditional dress, traditional Chinese clothing male, traditional Chinese clothes, traditional Chinese clothing female
Hanfu	Chinese hanfu, hanfu dress, hanfu clothing, Chinese hanfu dress, male hanfu, hanfu costume, modern hanfu, what is hanfu

Table 2. Monthly search data for some keywords in English

	Chinese clothing	China clothing	Chinese clothes	red qipao
2004/01	22	54	58	0
2004/02	44	39	68	0
2004/03	100	100	100	0
2004/04	47	76	54	0
2004/05	39	35	59	0
2004/06	12	36	30	0
2004/07	41	25	0	0
2004/08	22	44	19	0
2004/09	28	62	62	0
2004/10	36	86	55	0
2004/11	37	44	38	0
2004/12	17	0	9	0

2005/01	37	43	30	0
2005/02	27	82	70	0
2005/03	23	62	40	0
2005/04	47	71	30	0
2005/05	29	71	27	0
2005/06	16	8	14	0
2005/07	13	0	28	0
2005/08	20	30	6	0
2005/09	19	49	24	0
2005/10	33	73	45	0
2005/11	25	41	29	0
2005/12	13	39	22	0

3   Attention index synthesis and analysis

Based on the theory of the three-stage communication effects of cognition, attitude and behavior [9], the attention index synthesized by online search data can reflect the cognition and attention of overseas audiences to Chinese culture. Different countries and nations have certain differences in values, religious beliefs, customs, etc. Such differences are cultural differences, and cultural differences between different civilizations will affect cultural exchanges and transmissions between nations. The American scholar Huntington divided the world civilization into Confucian civilization, Islamic civilization, Hindu civilization, Eastern Orthodox civilization, Japanese civilization, Western Christian civilization, Latin American civilization and possible African civilization in "The Clash of Civilizations and the Reconstruction of World Order". Based on this, this paper selects five countries: the United States, Japan, India, Russia, and South Korea, which are located in the Christian civilization, Japanese civilization, Hindu civilization, Eastern Orthodox civilization, and Confucian civilization. The search data of the representative countries of each civilization is used to synthesize the overall attention index of Chinese culture and the attention index of different layers of culture in the civilization environment.

This part starts with the most common topics of various layers of culture in seventeen years, and examines the attention of different layers of Chinese culture from time and space. Based on the Eviews10.0 operating platform, this paper uses the Hodrick-Prescott filtering to decompose the comprehensive index trend cycle sequence, and then to explore its periodic characteristics and development trends. Furthermore, for the secondary indicators of different layers of culture under different civilizations, the annual index for 17 years from 2004 to 2020 is synthesized, and the differences in the spread of Chinese culture under different civilizations are compared. The index construction formula is shown in (1).

$$\text{Attention} = \sum \text{SearchVolumekeyword}(i) \quad i=1,2,3\ldots n \qquad (1)$$

Attention represents the monthly attention index of a certain culture, and SearchVolumekeyword(i) is the monthly search volume of a keyword in the cultural category.

3.1 Material layer cultural attention index

The material layer is the outermost layer of the cultural system and the externalized embodiment of culture. In this paper, clothing culture, food culture and architectural culture are selected as the representatives of the material layer culture, and the monthly search data of multilingual keywords are used to construct the cultural attention index of the material layer of each civilization. After HP filtering, part of the representative countries of each civilization are decomposed Trend is shown in Fig. 1.
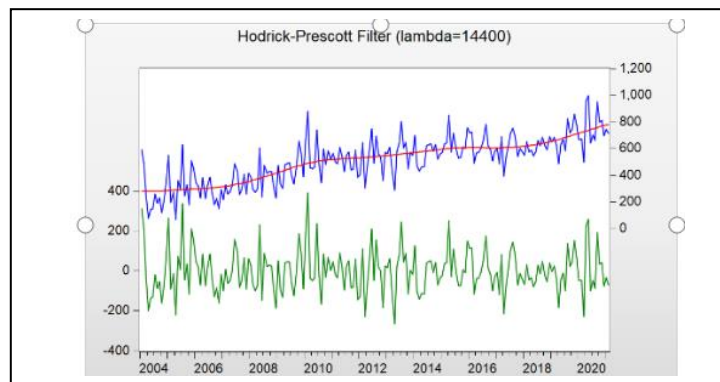
Fig. 1. Russian Material Layer Attention Index
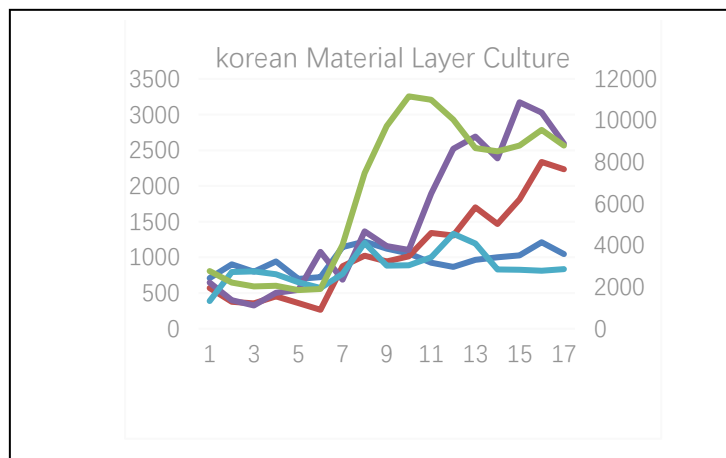
Fig. 2. The attention curve of the secondary indicators of Korea

The monthly search data of multilingual keywords contained in the secondary indicators of the material layer is used to synthesize the attention change curve of each country, and the time span is 17 years. For architectural culture, the basic keywords are the Great Wall, Forbidden City, Terracotta Warriors etc; food culture includes Beijing roast duck, Gongbao chicken, sweet and sour pork, etc; medical culture includes keywords such as Chinese medicine and Chinese medicine; science and technology include Xiaomi, Huawei and Baidu. Fig. 2 shows the attention curve of secondary cultural indicators of South Korea.

3.2　Relational Layer Cultural Attention Index

The relational layer refers to the middle layer of the cultural system, and is the embodiment of culture in terms of institutions and social customs. In this study, festival culture and institutional culture are selected as the representatives of the relational layer culture, and the monthly search data of multilingual keywords is used to construct the cultural attention index of each civilization relational layer. After HP filtering, the trends are decomposed as shown in Fig. 3.

The monthly search data of multilingual keywords contained in the secondary indicators of the relational layer is used to synthesize the attention change curve of each country, and the time span is 17 years. Fig.4 shows the attention curve of secondary cultural indicators of South Korea, and Japan.
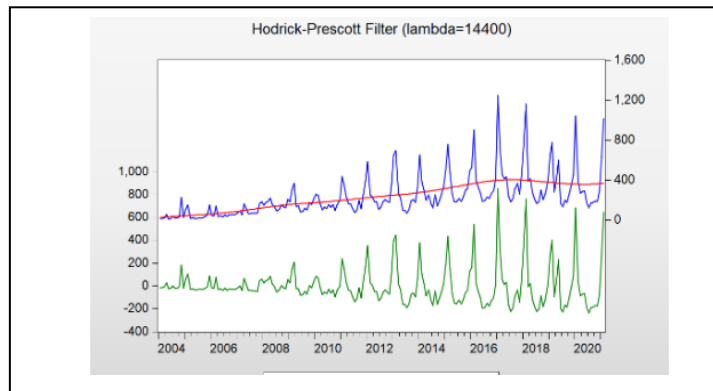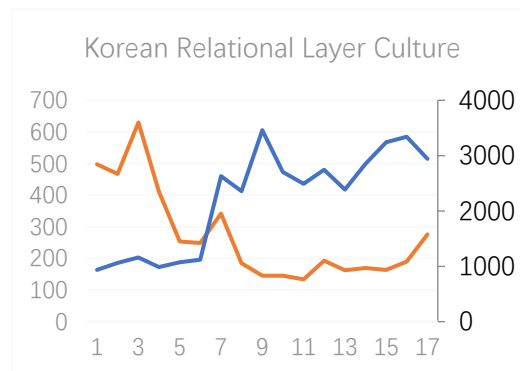

Fig. 3. Russian Relational Layer Attention Index

Fig. 4. The attention curve of the secondary indicators of Korea

3.3　Conceptual layer cultural attention index

The conceptual layer is the innermost layer of the cultural system, and it is the objectification of human cultural mentality and spiritual activities. In this study, religious culture, film and television culture and sports culture are selected as the representatives of the conceptual culture.

The monthly search data of multilingual keywords is used to construct the cultural attention index of each civilization conceptual layer. After HP filtering, the trends are decomposed as shown in Fig.5.
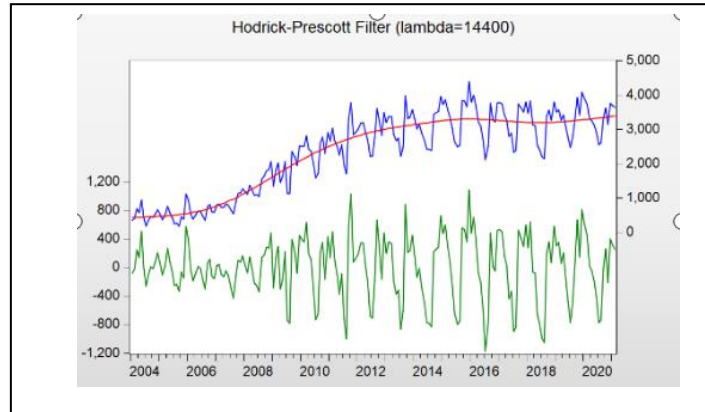

Fig. 5. Russian Conceptual Layer Attention Index

The monthly search data of multilingual keywords contained in the secondary indicators of the conceptual layer is used to synthesize the attention change curve of each country, and the time span is 17 years. Fig.6 shows the attention curve of secondary cultural indicators of South Korea.
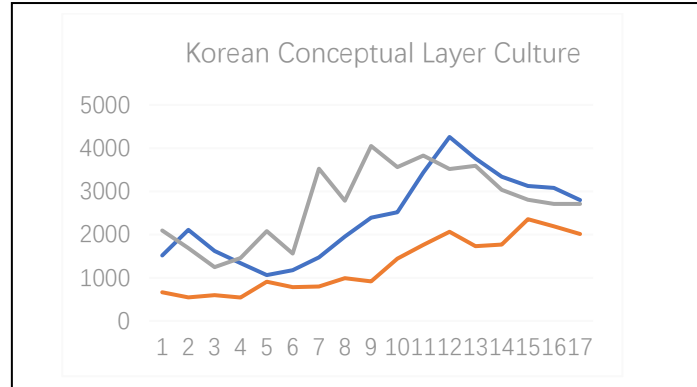

Fig. 6. The attention curve of the secondary indicators of Korea

4   Conclusion

Although Russia and the United States have some geographical or cultural differences with China, two countries' attention to Chinese culture is gradually growing in 17 years. On the other hand, Japan and Korea belong to the Confucian cultural circle, but the attention has not continuously increase. This shows that those countries with similar cultural attributes or geographical proximity do not necessarily have a "natural" interest in Chinese culture, while those with large

cultural differences and distant geographies do not necessarily naturally dislike China. The degree of attention paid by overseas audiences to Chinese culture is a question that combines cultural, political and other complex factors.

Therefore, when promoting the spread of Chinese culture to the world, we should have a target. For example, food culture has a good spread effect in various forms of civilization. Therefore, when Chinese culture goes out, we can increase the promotion of Chinese cuisine and introduce Chinese food to the world through more diversified forms. The major cuisines allow overseas audiences to learn about China from China. In addition to food culture, for film and television works and custom culture, which are closely related to life and do not have a strong political cultural form, communication should be strengthened. Moreover, the communication methods in different countries must have their own characteristics, and fully respect the traditional customs and values of the target country. Therefore, for the conceptual-layer culture, we must strive to find similar and common content with other cultures in the process of dissemination, and use the way of seeking resonance in the process of dissemination by "what others say".

References

[1] Ginsberg, et al. Detecting Influenza Epidemics Using Search Engine Query Data [J]. Nature, 2009, 457: 1012-1014.J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[2] Ripberger, J. T. . "Capturing Curiosity: Using Internet Search Trends to Measure Public Attentiveness", Policy Studies Journal, 39 (2011): 239 - 259.K. Elissa, "Title of paper if known," unpublished.

[3] Chew C, Eysenbach G. Pandemics in the age of Twitter: content analysis of Tweets during the 2009 H1N1 outbreak[J]. PloS one, 2010, 5(11): e14118.

[4] Preis, Tobias, H. S. Moat, and H. E. Stanley. " Quantifying Trading Behavior in Financial Markets Using Google Trends", Scientific Reports, 3(2013): 1684.

[5] Scharkow M, Vogelgesang J. Measuring the public agenda using search engine queries[J]. International Journal of Public Opinion Research, 2011, 23(1): 104-113.

[6] Chen Yunsong, Wu Qingxi, Zhang Yi: "International visibility of Chinese cities in the past three hundred years: description and regression based on big data", "Society" 2015, Issue 5.

[7] Meng Tianguang, Zhao Juan.Internet search behavior and public attention in the era of big data: a dynamic analysis based on the Baidu index from 2011 to 2017 [J].Ocean of Learning,2019(03):41-48.

[8] Huang Yunhong. A Comparative Study of Cultural Hierarchical Structure Models[J].Journal of Guiyang University (Social Science Edition),2013,8(02):7-10.

[9] Luo Xue. On the construction of my country's media's international communication effect evaluation system[J]. Contemporary Television, 2016(10):78-80.