

## Analysis and Prospect of Smart Energy Based on Co-word Analysis

Kai Wang, Jianyong Li, Changzheng Zhao and Kunpeng He

School of Shandong University of Science and Technology, Shandong 266590, China;

1838975646@qq.com

---

*Abstract: Using the subject of Wisdom (CNKI) as the source of intelligent energy as the data source, using SPSS software and Python3 language, the factor analysis, cluster analysis and multi-dimensional scale analysis of the theme of smart energy are realized. The keywords are divided into three categories by co-word analysis and their research status and development trends are analyzed separately. It is concluded that smart energy is still a new academic direction in China. There are few related literature topics and narrow research fields. Smart energy should be combined with intelligent technology to develop and innovate in order to have better development; smart energy is green development and socialism. The important factors of ecological civilization construction play a vital role in the future development. Smart energy has become the inevitable direction of future development.*

*Keywords: Smart energy; co-word analysis; factor analysis; cluster analysis.*

---

### 1. INTRODUCTION

Smart energy is a relatively new concept that has emerged in recent years. In 2009, the international academic community, including the IBM expert team, proposed that interconnected technology will change the way the entire human world operates, involving the work and life of billions of people. Therefore, the academic community began to propose to build a more intelligent The Smarter Planet proposes the concepts of smart airport, smart bank, smart railway, smart city, smart power, smart grid, and smart energy [1].

At present, the lack of clear and consistent understanding and understanding of the concept of smart energy, the academic community has a different opinion, and has not yet formed a widely recognized authoritative definition. In short, smart energy refers to the form of energy that has self-organizing, self-checking, self-balancing, self-optimizing and other human brain functions to meet system, safety, cleanliness and economic requirements [2].

In this study, CNKI is used as the data source to extract and summarize high-frequency keywords. The co-word analysis method combining factor analysis, cluster analysis and multi-dimensional scale analysis is used to analyze the current research status of domestic smart energy and look forward to its development trend.

## 2. DATA COLLECTION AND PROCESSING

### 2.1 Research methods

Co-occurrence refers to the phenomenon in which the information described in the feature items of the document coexist. The feature items here include the external and internal features of the document, such as title, author, keyword, institution, etc. The “co-occurrence analysis” is the phenomenon of co-occurrence. Quantitative research to reveal the content association of information and the knowledge implied by feature items. The basic principle is to reflect the co-occurrence of lexical pairs or noun phrases in statistical literature to reflect the strength of the association between keywords, and then determine these words. The research hotspots, composition and paradigm of the disciplines or fields represented, the horizontal and vertical analysis of the development process and structural evolution of the subject area. The premise of this method is that the more times the vocabulary appears in the same document, the two the closer the relationship of the topics is. Thus, by counting the frequency of occurrence of the same document in the same document, a common word network composed of these word pairs can be formed. The distance can reflect the close relationship of the subject content [3].

### 2.2 High frequency keyword

The research data source is the CNKI academic literature database, which selects the Chinese academic journal network publication database, the Chinese doctoral thesis full-text library, the Chinese excellent master's thesis full-text database, and the China's important conference papers full-text database as the retrieval database, and retrieves the keyword. The literature for "smart energy". Excluded informal academic literature such as no authors, no keywords, and news reports from the search results, and finally 892 valid articles were obtained. The keywords are extracted and summarized, and then sorted according to the word frequency, and 20 high frequency keywords are obtained as shown in Table 1.

Table 1. High frequency keywords

No.	Key words	Frequency	No.	Key words	Frequency
1	智慧能源	41	11	互联网+	14
2	物联网	31	12	雄安新区	14
3	大数据	28	13	多能互补	14
4	智慧城市	24	14	一带一路	12
5	绿色发展	21	15	信息化	12
6	能源互联网	20	16	创新	12
7	可持续发展	20	17	生态文明	12
8	智能化	15	18	智能电网	11
9	新时代	15	19	可再生能源	10
10	应用	14	20	智慧交通	9

### 2.3 Common word matrix

High-frequency keywords largely reflect research hotspots and development profiles, but only through frequency analysis cannot deeply reflect the internal structure of the field development and the intrinsic correlation between keywords. Therefore, on the basis of high-frequency keywords, it is necessary to establish a co-word matrix, that is, a matrix composed of the number of times that two-two high-frequency keywords appear simultaneously in the same document. Write scripts in Python 3 and save the results as an Excel spreadsheet. Finally, a 20×20 co-word matrix is generated.

Import the SPSS software, use the correlation analysis function in SPSS, and choose to calculate the Pearson correlation coefficient, and convert the common word matrix into the correlation matrix, as shown in Table 2.

Table 2. Correlation Matrix (Partial)

Key words	智慧能源	物联网	大数据	智慧城市	绿色发展	能源互联网	可持续发展
智慧能源	1.000	0.042	0.239	0.062	0.076	0.091	0.199
物联网	0.042	1.000	0.477	0.234	0.259	0.270	0.190
大数据	0.239	0.477	1.000	0.256	0.076	0.286	0.213
智慧城市	0.062	0.234	0.256	1.000	0.316	0.690	0.221
绿色发展	0.076	0.259	0.076	0.316	1.000	0.357	0.000
能源互联网	0.091	0.270	0.286	0.690	0.357	1.000	0.261
可持续发展	0.199	0.190	0.213	0.221	0.000	0.261	1.000

### 3. DATA CALCULATION AND ANALYSIS

In order to further reveal the internal structure and inter-word relationship between high-frequency keywords, factor analysis, cluster analysis and multi-dimensional scale visualization are performed on the correlation matrix.

First, factor analysis is performed on the high frequency keyword correlation matrix. The correlation matrix is imported into SPSS, and the factor analysis in the dimension reduction function is used. The parameters are the main component method, the covariance matrix analysis and the average orthogonal rotation method, and the factor analysis results are obtained (Table 3). The three principal component factors with eigenvalues greater than 1 are retained, and the results will provide a reference for cluster analysis.

Table 3. Factor Analysis (Partial)

ingredient	total	variance %	total %	total	variance %	total %
1	3.326	16.631	16.631	16.631	16.631	0.199
2	2.932	14.662	31.293	14.662	31.293	0.190
3	2.231	11.156	42.449	11.156	42.449	0.213
4	1.908	9.541	51.990			
...	...	...	...			
20	0.000	0.000	100.000			

Cluster analysis of high frequency keyword correlation matrices was performed using SPSS. Using the system clustering analysis method in the classification function, the clustering method is selected as the connection method between groups, the metric is the square geometric distance, and the correlation matrix is clustered and analyzed, as shown in Fig. 1. According to the clustering results shown in the cluster analysis tree, the high frequency keywords are classified into three categories. The first category consists of smart energy, Internet of Things, big data, intelligence, applications, the Belt and Road, information and innovation, denoted by A; the second category consists of smart cities, energy Internet, Internet +, multi-energy complementary, smart grid And renewable energy composition, denoted by B; the third category consists of green development, new era, ecological civilization, intelligent transportation, sustainable development and Xiong'an new district, denoted by C. It can be seen that the clustering results are basically consistent with the factor analysis results.

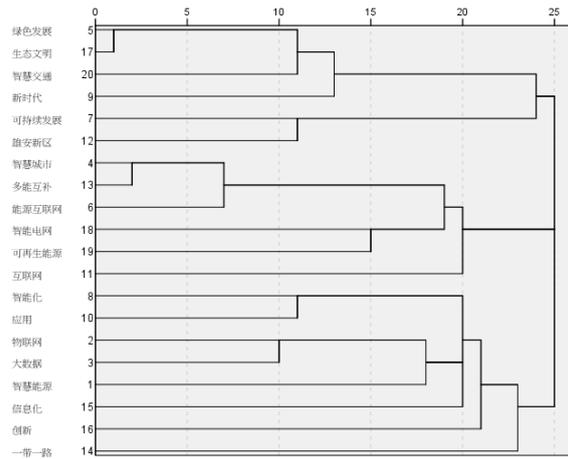


Figure 1. Cluster Analysis

The multidimensional scaling analysis of the correlation matrix is performed using the multidimensional scaling analysis function of the SPSS software, and the multidimensional scaling analysis graph is output as shown in Fig. 2. It can be seen that the multi-dimensional scale analysis results are basically consistent with the factor analysis results and the clustering analysis results, indicating that the analysis process is credible and effective.

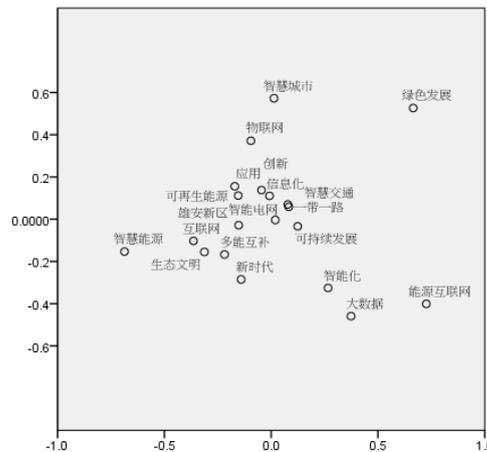


Figure 2. Multidimensional scaling analysis

#### 4. CONCLUSION

(1) It can be seen from the multi-dimensional scale analysis chart that high-frequency keywords are concentrated in the vicinity of keywords such as the smart grid Internet, indicating that in the field of domestic smart energy research, the research direction is relatively narrow, and the research is not yet mature and systematic. According to the statistics of CNKI literature, the concept of smart energy has attracted the attention of domestic academic circles since 2009. In the following years, the degree of attention has risen rapidly and the number of documents has increased rapidly. However, as the field is still a new academic direction in China, there are few related literature topics. Only 41 of the 892 related documents collected are based on smart energy, and the depth of research needs to be further deepened. On the other hand, it also shows that the research direction of smart energy is developing rapidly and has great potential for development.

(2) Class A keywords include smart energy, internet of things, big data, intelligence, applications, the Belt and Road, information and innovation. This type of keyword focuses on emerging intelligent technologies. It shows that smart energy should be combined with intelligent technology, and innovation and development can have better development.

(3) Class B keywords include smart cities, energy internet, internet +, multi-energy complementarily, smart grids and renewable energy. It shows that smart energy has a great connection with smart cities. The development and development of smart energy should be combined with relevant technologies in smart cities to achieve sustainable energy development.

(4) Class C keywords include green development, new era, ecological civilization and smart transportation, sustainable development and Xiong'an New District. It shows that smart energy is an important factor in green development and socialist ecological civilization construction, and plays a vital role in future development. Smart energy has become an inevitable direction for future development.

## **REFERENCES**

- [1] Cheng Siwei. Smart Energy and Ecological Civilization [J]. Guangdong Electric Power, 2013(8):18-18.
- [2] Liu Jianping, Chen Shaoqiang, Liu Tao. Smart Energy: We have been in this 10,000 years [M]. Science and Technology Literature Publishing House, 2013.
- [3] Fu Zhu, Wang Yufen. Research on Some Problems in the Collection Stage of Terms in Co-word Analysis [J]. Journal of the China Society of Scientific Information, 2016, 35(7): 704-713.