# An Analysis of the Characteristics and Evolution of China's Pension Policy Based on Text Mining

Liang Yang
*Business and Tourism School, Sichuan Agricultural University, China*

Meng Zhao
*Business and Tourism School, Sichuan Agricultural University, China*

Miyan Liao
*Business and Tourism School, Sichuan Agricultural University, China*

Liangqiang Li
*Business and Tourism School, Sichuan Agricultural University, China*, lilq@sicau.edu.cn

Xiaolin Li
*Business and Tourism School, Sichuan Agricultural University, China*

<u>**Full Research Paper**</u>

# An Analysis of the Characteristics and Evolution of China's Pension

# Policy Based on Text Mining

*Liang Yang[1], Meng Zhao[1], Miyan Liao[1], Liangqiang Li[1*], Xiaolin Li[1]*
[1]Business and Tourism School, Sichuan Agricultural University, China

**Abstract:** With the economic and social development, China has changed from the initial demographic dividend to an aging population, and governments at all levels have introduced a series of policies to cope with the negative effects caused by the aging population. This paper collects a total of 167 pension policy documents from 2010 to 2020, and uses text mining techniques to systematically study the characteristics and evolutionary trends of pension policies in terms of issuing agencies and contents. The conclusions drawn in this paper are as follows: First, the Ministry of Civil Affairs and other departments play a major role in the formulation of pension policies. Second, the pension policies formulated by the government during the study cycle are biased toward public service aspects, and the content of medical care, investment, and pensions is mainly indirectly focused on a cyclical basis. The findings of this paper can help the government to check the gaps and improve the existing policy system, and the research methodology can help the government to make more scientific decisions.

Keywords: Pension policy, Policy characteristics, Text mining

## 1. INTRODUCTION

With the continuous development of China's economy and society, and the continuous improvement of medical and health conditions, the issue of population aging is becoming an increasingly prominent issue in social governance. There are three main factors influencing population aging: declining fertility rate, longer life expectancy, and a large number of people entering old age. According to the data released by the National Bureau of Statistics, the number of elderly people aged 60 and above in China reached 254 million at the end of 2019, accounting for 18.1% of the total population, and the number of elderly people aged 65 and above reached 176 million, accounting for 12.6% of the total population, representing a significant increase compared to 7% in 2000. Along with the growth of urban and rural residents' income level, residents' willingness to have children is getting lower and lower, and the trend of population aging in China may get further accelerated.

The rapid growth of population aging may bring a series of social problems in the background. Firstly, the growth of the number of elderly population puts forward higher requirements for local services in medical care, maintenance, welfare, transportation, etc. Secondly, the decrease in the proportion of young and middle-aged population makes the number of labor force decrease, and the aging of the group presents more difficulties for social development [1]. The increase of population aging makes the burden of society and family increase, and how to solve the elderly problem has become a high concern of the society. Meeting the demand for elderly services is the core and key link of the national strategy of actively corresponding to the aging population, as well as an important part of social governance [2]. In addition, due to the weak awareness of family planning and uneven government investment, the problem of rural retirement is becoming increasingly prominent, and there is an extreme imbalance between urban and rural retirement resources [3].

Government departments have played a key leading role in actively responding to the social aging problem. As early as 1980, the Ministry of Justice issued the "Ministry of Justice Notary Lawyer Department on the Issuance of Unclaimed Pension Certificates" to focus on pension-related matters. Since the 18th National

---

* Corresponding author. Email: lilq@sicau.edu.cn

Congress, national ministries and commissions have issued successive policies from different aspects such as finance, medical care and scientific innovation to focus on the development of the elderly and the construction of the pension system. 2019, the General Office of the State Council issued "Opinions on Promoting the Development of Pension Services" systematically proposed 28 measures to systematically solve pension problems, and the pension service policy has achieved a breakthrough in its entirety. For a long time, government departments have mostly formulated domain-related policies from their own perspectives, which eventually led to a segmented and fragmented political system [4]. Although a series of policies have been introduced by various ministries to solve the pension problems, the characteristics presented by the existing pension policies, the deficiencies of the existing policies, and the characteristics of the evolutionary trends of the policies over time are still half understood. Based on this, this paper uses text mining techniques to dig into the characteristics of existing pension policies and their dynamic evolution patterns, explore the shortcomings of existing pension policies, and provide suggestions for government policy formulation.

## 2. PENSION POLICY RESEARCH

Rapid changes in economic conditions and demographic characteristics require that policy design be closely aligned with reality and adjusted in a timely manner in order to advance the development of senior care. The Chinese government foresees an old-age problem arising from the one-child policy in 2020, but is overly optimistic about the outcome and solutions, believing that the problem will be solved through economic development, improved living conditions, and increased social welfare and security [3]. The reality is that China's social pension problem has far exceeded the government's expectations. In recent years, the Chinese government has introduced a series of policies and established a universal social security for the elderly and a social care system for the elderly [5].

Government documents dealing with elder care can be considered as belonging to a continuously evolving public policy narrative, and existing studies have explored existing policies in terms of their characteristics [6]. Zhai et al. argue that the aging of China's population now shows new characteristics such as a significant slowdown in the rate of aging, an increase in the size of empty nesters living alone, and an increase in the literacy level of the elderly, and in this context, the existing elderly care policies need to be updated to promote the development of the elderly care industry [7]. Wang studied the evolution path from the perspective of financial security of elderly services, and found that the financial security responsibility of the existing elderly services showed the characteristics of "less welfare - lack of financial security responsibility - return of financial security responsibility - multiple government, market, society and family Governance" [8]. Yi et al. systematically analyzed the current status of social pensions in China and the existing dilemmas, and they argued that the key factors hindering the sustainable development of social pensions are reflected in two aspects: first, there is no policy to clarify the nature of services and support allocation of relevant resources: second, social pension policies are formulated in a stylized manner rather than based on actual interactive processes [9].

Although many studies have used pension policies as an object to actively suggest suggestions for governmental decision-making, most previous studies are qualitative in nature and few have used quantitative studies to explore the characteristics of pension policies and their evolutionary features. Based on this, this paper intends to use text mining to explore the characteristics of existing pension policies and the analysis of policy evolution, and the findings can help the government make scientific decisions.

## 3. STUDY DESIGN
### 3.1 Data source

In order to ensure the accuracy, relevance and representativeness of the policy text data, this study took the

period from January 2010 to January 2021 as the time interval, and selected the relevant pension policy texts issued by the State Council, the Ministry of Civil Affairs, the Ministry of Finance, the National Health and Welfare Commission, the National Health Insurance Administration, the Ministry of Human Resources and Social Security, and the Ministry of Housing and Urban-Rural Development as the data base. When collecting the data, we followed the principles of authority, openness and relevance, and searched the legal database of Peking University with the title of "pension". After screening the policy texts one by one according to the policy content, excluding the policy texts, approvals, activity notices and other policy texts that have become invalid or partially invalid, we finally sorted out a total of 167 valid policy documents, and the specific adoption situation is shown in Table 1.

**Table 1.    Raw data results and adoption**

| Policy Category | Number of Post | Number of Adoption |
|---|---|---|
| Administrative Laws and Regulations | 17 | 16 |
| Judicial Interpretation | 1 | 0 |
| Regulation | 791 | 149 |
| The Party's Rules and Regulations | 2 | 2 |
| Industry Regulation | 8 | 0 |

### 3.2 Research methods and analytical tools

For policy text research, this paper wants to adopt bibliometric and text mining methods.

(1) Bibliometrics

Bibliometrics takes literature and intelligence as the object of study, and uses mathematics, statistics, intelligence and other disciplines to investigate its structural characteristics, quantitative characteristics, and the laws of change, and to further investigate certain structures and characteristics of science and technology [10]. Generally, the results of quantitative analysis of literature using bibliometric methods are clear and lucid, so in this paper, we use bibliometric methods to analyze the external characteristics of current policies on new coronary pneumonia, and describe and evaluate the existing policies through two levels: chronological order and the subject of issuance.

In conducting the analysis of issuing subjects, the visualization is presented through social network analysis. A social network is a network structure composed of multiple actors, and social network analysis is used to show the connections among actors in the form of a social network diagram [11]. In this paper, we study the collaboration between departments by counting the joint and independent postings of departments and importing the data into Gephi software for visualization.

(2) Text mining

Text mining, or text data mining, refers to the pattern of mining valuable information from a large collection of unstructured texts and can be considered as an extension of data mining or knowledge discovery from structured databases [12] [13]. Text mining techniques enable the discovery of potential connections, patterns, etc., and the extraction of valuable information from a large amount of unstructured text. In the context of this paper, text mining is used to perform cluster analysis, classify different policies, and explore the characteristics of existing policies in terms of content structure. The specific process of text mining in this paper is shown in Figure 1.
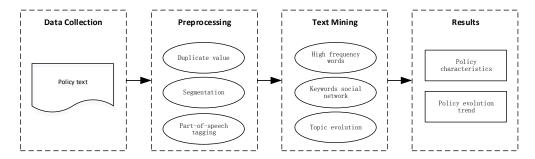
**Figure 1.   Text mining flowchart**

### 3.3 Keyword extraction

Keywords are words that can maximize the information on the topic of the text, and keyword extraction can summarize the text information briefly [14]. In the research environment of this paper, keyword extraction can grasp the topic of the policy text to the maximum extent quickly and provide a basis for content analysis.

In this paper, the TF-IDF algorithm is used for keyword extraction, and the first part TF (Term Frequency) indicates the frequency of words appearing in a text, which is calculated as  (1).

$$TF = \frac{tf}{doc\_lenggth} \tag{1}$$

The second part, IDF (Inverse Document Frequency), is used to measure the general importance of words and is calculated as  (2).

$$IDF = log(\frac{N}{df}) + 1 \tag{2}$$

where,  $tf$   is the number of keyword appearances,  $doc\_length$   is the length of the text, N is the total number of text, $df$ is the number of text containing keywords. The final TF-IDF is the product of TF and IDF.

### 3.4 Latent dirichlet allocation

The Latent Dirichlet distribution, abbreviated as LDA, is a generative probabilistic model for discrete data collections such as text corpora, which assigns topics to documents and generates topic distributions over the words of a given text collection [15]. LDA contains three main layers of structure: words, topics, and documents, as shown in Figure 2.
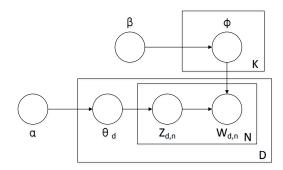


**Figure 2.   LDA structure**

Next we briefly describe the process of LDA generation of documents. First, a document $di$ is selected according to the prior probability $p(di)$; second, the topic distribution $\theta_i$ of document $d_i$ is generated by sampling from the Dirichlet distribution α, and the topic distribution is generated by the Dirichlet distribution with hyperparameter $\alpha$; third, the topic $z_{i,j}$ of the $j$th word of document $d_i$ is generated by sampling from the

polynomial distribution $\theta_i$ of the topic; fourth, the topic $z_{i,j}$ of the Dirichlet distribution $\beta$ to generate the word distribution $\phi_{z_{i,j}}$ corresponding to the topic $z_{i,j}$, the word distribution $\phi_{z_{i,j}}$ is generated by the Dirichlet distribution with parameter $\beta$; finally, the word $\omega_{i,j}$ is sampled from the polynomial distribution $\phi_{z_{i,j}}$ of words to finally generate the word $\omega_{i,j}$ .

## 4 QUANTITATIVE ANALYSIS OF PENSION POLICIES

### 4.1 Number of posts

The last ten years of pension policy releases are shown in Figure 3. According to the table, it is found that the number of pension policies peaked in 2014, the second highest number of pension policies in 2019, and the lowest number of pension policies in 2011.
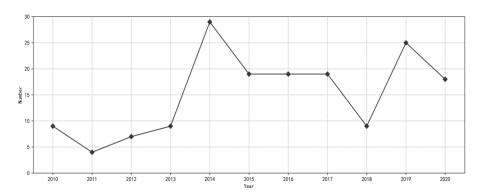


**Figure 3.    Pension policy year distribution**

### 4.2 Institutional characteristics

Based on the collected 167 policy texts, this paper made statistics on the number of policy issuing institutions. A total of 59 departments have formulated and promulgated policies regarding the pension issue. The top 10 institutions with issued policies are shown in Table 2. It can be found that the Ministry of Civil Affairs solves the pension problem by formulating a variety of relief and welfare policies. It plays the role of the founder in maintaining social stability, improving the happiness of the elderly and promoting economic development. In addition, the Ministry of Human Resources and Social Security, the Ministry of Finance, the State Council and other departments have further improved pension policies from different perspectives in order to meet the needs of diversified and multi-level pension services for the elderly.

**Table 2.    Number of institutional policy releases**

| Institution | Number | Institution | Number |
|---|---|---|---|
| Ministry of Civil Affairs (民政部) | 78 | Insurance Regulatory Commission (保监会) | 11 |
| Ministry of Human Resources and Social Security (人力资源与社会保障部) | 49 | National Health and Wellness Commission (国家卫生健康委) | 10 |
| Ministry of Finance (财政部) | 45 | National Health and Family Planning Commission (国家卫生计生委) | 8 |
| State Council (国务院) | 18 | Ministry of Commerce (商务部) | 6 |
| National Development and Reform Commission (国家发改委) | 18 | Ministry of Housing and Urban-Rural Development (住房和城乡建设部) | 5 |

We use Gephi software to describe the social network between sectors. The joint issuance of pension policy departments is shown in Figure 4. The size of the network nodes reflects the number of documents issued by the

departments, and the thickness of the connecting lines between the nodes reflects the closeness of the departments. According to the network diagram, we can find that the Ministry of Civil Affairs has the largest node in the network diagram, and at the same time, the number of nodes connected to it is the most numerous and complex, which indicates that the Ministry of Civil Affairs has issued the largest number of policies in the research cycle of this paper, and at the same time, the policy cooperation with other departments is also the most frequent. However, in terms of the complexity of the relationships of the nodes, the Ministry of Finance, the Ministry of Land and Resources, the Ministry of Housing and Urban-Rural Development, and the Ministry of Human Resources and Social Security jointly issued relatively more policies, which played an important role in the solution of the pension problem. Finally, the National Bureau of Statistics, the National People's Committee, and the National Open Bank have issued relatively few documents, but they have actively cooperated with other departments and jointly issued policies, which have played an active role in promoting the development of China's pensions.
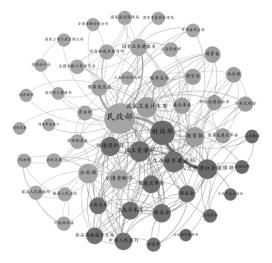


**Figure 4.    Departments jointly issued network**

## 5 PENSION POLICY CONTENT FEATURES

### 5.1 High frequency word analysis

In this paper, we use the jieba library of Python to carry out the word segmentation work, and use the deactivated word list of HIT and the deactivated word bank of Machine Intelligence Laboratory of Sichuan University to get a total of 1897 deactivated words. A total of 167 policy texts were divided into words using the precise mode of jieba library, and according to the research purpose of this paper, only verbs as well as nouns were retained, and a total of 132,160 words and 6,456 dissimilar words were obtained. According to the formula of high and low frequency words proposed by Qinglan Sun: $D=\sqrt{n}$ . ($D$ indicating the critical value of high and low frequency words, $n$ indicating the total number of dissimilar words) [16], the threshold of high and low frequency words in this paper was calculated as 80. Combined with the actual situation, the top 60 high-frequency words are finally selected, as shown in Table 3.

High-frequency words can reflect the focus of a government's policy over time. It can be seen from the high-frequency words that "pension" appears the most frequently, which corresponds to the theme of this study. "Government", "business unit", "the Ministry of Civil Affairs", "agency", "personnel", "enterprise", "department", "organization" and other subjects related to the high frequency words reflects the Chinese government attaches great importance to the elderly pension problem, governments at all levels not only play a good role itself, and promote enterprise pension services better, which is beneficial to improve the people's feeling, happiness and security. The two high-frequency words "senior citizens" and "urban and rural residents" reflect the targets of the

pension policy. The four high-frequency keywords of "society", "region", "community" and "country" reflect the geographical scope that the pension policies formulated by government departments at all levels are adapted to. High-frequency words such as "service", "development", "management", "work", "construction", "support", "establishment", "system", "guarantee", "reform" and "policy" reflect the action arrangements made by government departments to promote the government and enterprises and institutions to improve the quality of old-age service and promote the healthy development of old-age service. "Endowment insurance", "investment", "join", "funds", "offer", "payment", "product", "information", "pension" and other high frequency words reflect the government to ensure that the elderly pension life, in terms of money, products, information services such as content to make detailed regulations, safeguard the rights and interests of the elderly and the life.

**Table 3.    High frequency words (top 60)**

| Number | Word | Frequency | Number | Word | Frequency | Number | Word | Frequency |
|---|---|---|---|---|---|---|---|---|
| 1 | Pension Policy | 5690 | 21 | Payment | 640 | 41 | Unit | 434 |
| 2 | Service | 3339 | 22 | System | 626 | 42 | Household | 433 |
| 3 | Institution | 2745 | 23 | Pilot | 620 | 43 | Way | 427 |
| 4 | Endowment Insurance | 1548 | 24 | Guarantee | 596 | 44 | Treatment | 413 |
| 5 | Elderly | 1423 | 25 | Reform | 588 | 45 | Service Facility | 404 |
| 6 | Development | 1185 | 26 | Department | 544 | 46 | Boost | 403 |
| 7 | Management | 1178 | 27 | Project | 544 | 47 | Business | 390 |
| 8 | Work | 1155 | 28 | Region | 531 | 48 | Life Insurance | 381 |
| 9 | Construction | 1055 | 29 | Policy | 530 | 49 | Complete | 375 |
| 10 | Investment | 990 | 30 | Government | 523 | 50 | Public Institution | 372 |
| 11 | Society | 879 | 31 | Situation | 523 | 51 | Transfer | 366 |
| 12 | Personnel | 856 | 32 | Product | 523 | 52 | Condition | 364 |
| 13 | Insured | 799 | 33 | Information | 501 | 53 | Service Industry | 360 |
| 14 | Community | 791 | 34 | Capital | 498 | 54 | Urban-Rural Residents | 347 |
| 15 | Enterprise | 766 | 35 | Pension | 490 | 55 | Ministry of Civil Affairs | 341 |
| 16 | Support | 736 | 36 | Standard | 481 | 56 | Make | 338 |
| 17 | Correlation | 723 | 37 | Organization | 472 | 57 | Accomplish | 337 |
| 18 | Fund | 711 | 38 | Handle | 449 | 58 | Office | 328 |
| 19 | Provide | 707 | 39 | Implement | 441 | 59 | Country | 327 |
| 20 | Establish | 645 | 40 | Encourage | 436 | 60 | Mechanism | 323 |

### 5.2 Keywords social network

In this paper, we use the TF-IDF method to calculate the word weights in 167 policy texts, take the top 60 keywords in descending order of weights, and generate a keyword co-occurrence matrix, filter the nodes with co-occurrence times lower than 60, and finally use Gephi software to generate a keyword social network graph, and the results are shown in Figure 5, which includes a total of 44 nodes and 1252 edges.

In the network graph of this study, the stronger the intermediary centrality of the nodes, the stronger their role. The results show that the intermediary centrality of "work" is above 90, and it has the largest intermediary centrality, and it plays the most important role in the whole network graph. This work mainly refers to the work activities related to the elderly, which corresponds to the theme of this study. "related", "people", "establishment", "elderly", " development," "institutions," "services," "organization," "management ", "policy", "construction" and other nodes have a large degree of intermediary centrality, indicating that in order to improve the quality of elderly care and elderly services, relevant institutions and organizations strengthen management and improve services according to elderly care policies are important elements throughout most of the elderly care The nodes "human resources", "construction" and "construction" have a high intermediary center. The keywords "human resources", "pension insurance", "project", "home", "ten points" and "ten points" are the most important elements of the policy.

Keywords such as "ten points" are at the margins of the whole study, with a small number of interconnections, and these are not commonly concerned in elderly care policies.
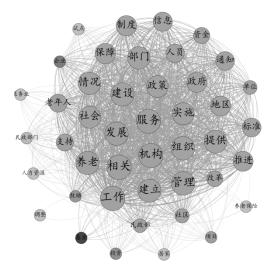


**Figure 5.    Keywords social network**

### 5.3 Policy evolution analysis

In this paper, we use the topic probability model to cluster the policy text content, and according to Figure 6, the coherence score is highest when the number of topics is 6. Therefore, the policy text is divided into six categories, and 10 topic words are taken under each category, and the results of clustering are shown in Table 4 [17]. The first policy category was named as retirement policy based on the key words of funding, community, subsidy, and central finance. The second policy category was named as insurance-type policies based on key words such as pension insurance, participation, contribution, treatment, and institutions. The third policy category is named as investment category policy according to key words such as investment, management, product, fund, and asset. The fourth policy category is named as epidemic prevention and control policy according to key words such as prevention and control, epidemic, area, staff, disinfection, and isolation. Based on the key words of service, institution, development, elderly, and service facilities, the fifth category of policies was named as public service category policies. Based on the key words of life insurance, support, indicator, medical care, and life insurance, the sixth category of policies was named as medical care category policies.
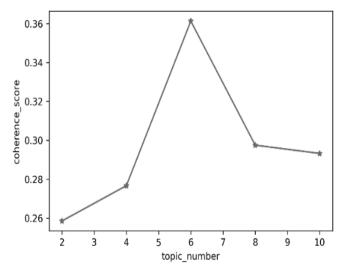


**Figure 6.    Coherence score**

**Table 4.   Policy tobic classification**

| First-level Topic | Retirement | Probability | Insurance | Probability | Investment | Probability | Epidemic prevention and control | Probability | Public Service | Probability | Medical | Probability |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Second-level Topic | Capital | 0.027 | Endowment Insurance | 0.045 | Investment | 0.047 | Prevention and Control | 0.016 | Service | 0.047 | Life Insurance | 0.017 |
| | Community | 0.025 | Insured | 0.027 | Management | 0.045 | Epidemic | 0.014 | Institution | 0.025 | Support | 0.009 |
| | Subsidies | 0.025 | Payment | 0.021 | Product | 0.031 | Region | 0.013 | Development | 0.017 | Indicator | 0.008 |
| | Military | 0.014 | Treatment | 0.014 | Funds | 0.025 | Staff | 0.011 | Elderly | 0.017 | Development | 0.008 |
| | Central Finance | 0.011 | Public Institution | 0.012 | Assets | 0.017 | Disinfection | 0.009 | Society | 0.012 | Type | 0.007 |
| | Retirement | 0.01 | Transfer | 0.012 | Business | 0.011 | Isolation | 0.008 | Construction | 0.012 | Medical | 0.007 |
| | Investment | 0.009 | Urban-Rural Residents | 0.011 | Escrow | 0.009 | Pilot work | 0.007 | Reformation | 0.007 | Premium | 0.007 |
| | Retirement | 0.009 | Personal Account | 0.009 | Account | 0.008 | Region | 0.007 | Government | 0.006 | Group | 0.007 |
| | Pension | 0.009 | Pickup | 0.009 | Combination | 0.007 | Civil Affairs | 0.007 | Service Facilities | 0.006 | Personal Insurance | 0.006 |
| | Ministry of Finance | 0.008 | Social Insurance | 0.008 | Risk | 0.007 | Pneumonia | 0.006 | Service Industry | 0.006 | Product Design | 0.006 |

Based on the results of the clustering of the thematic probability model, the frequency of occurrence of policy contents in different categories in that year is counted on a yearly basis, and the results are normalized in order to compare the results between different categories in the time dimension, and the normalization formula is shown as (3).

$$x^* = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{3}$$

where, $x_{min}$ denotes the minimum value and $x_{max}$ denotes the maximum value.

The policies promulgated by the government often reflect the focus of the whole society at one stage. Based on the evolution of the focus of the policy content, we can find out the patterns, shortcomings and future trends of government policy making, and the evolution trend of the pension policy from 2010 to 2020 is shown in Figure 7.

First of all, from an overall perspective, the number of public service policy contents has been maintained at a high level for most of the time, and as the number of elderly people continues to grow, the number of elderly institutions, elderly activity places, and elderly service nursing staff is stretched to the limit, making it difficult for the needs of the elderly to be met. In 2011, the General Office of the State Council issued the Notice on the Issuance of the Construction Plan for Social Aged Care Service System (2011-2015), which systematically provided guidance for the socialization of public services for the aged as well as the industrialization of a phase of development.

Secondly, the number of epidemic prevention and control type policies increased sharply in 2020 due to the impact of the New Crown Pneumonia epidemic, and government policies from this time onwards focused on restoring elderly services in terms of epidemic services to ensure the life and health safety of the elderly and to

better serve their needs.

Finally, some of the thematic contents, although the overall number is small, the government policy will focus on the topic within a certain period. For insurance policies, some policies have been formulated for targeted solutions in response to the current low level of protection for the elderly and the temporary uncertainty of the treatment mechanism. 2010 and 2014 saw certain increases in related policies, and in 2018, the Ministry of Human Resources and Social Security and the Ministry of Finance formulated the "Guidance on the Establishment of a Normal Adjustment Mechanism for the Treatment of Basic Pension Insurance and Basic Pension for Urban and Rural Residents , the policy is closely focused on the establishment of a guaranteed urban and rural residents' pension insurance as well as pension mechanism. For retirement policies, according to the trend of policy evolution, government policies will mention retirement-related matters for the elderly within 2-3 years. For investment-type policies, in order to strengthen the management of assets for the elderly, the government has indirectly developed some initiatives to address such issues, with a high number of such policies in 2013, 2015, and 2018, with the central government giving full play to its role as a demonstration driver for budgetary investment. For medical policies, the current limited medical resources and independent service system are difficult to meet the medical needs of the elderly, and the government rarely formulates relevant policies directly; instead, comprehensive policies will be mentioned from the medical perspective, and such policy measures still need to be improved.
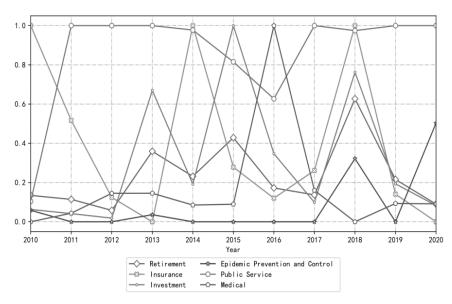


**Figure 7.    Policy evolution trends**

## 6 CONCLUSIONS

In the face of the rapidly growing trend of aging, the government should implement an active aging strategy and formulate an aging policy that matches economic and social development. This paper collects senior care policy texts for a total of 11 years from 2010 to 2020, and uses text mining techniques to study the characteristics of policy texts and their evolutionary features in this stage, in order to suggest suggestions for the government's next stage of policy formulation. The main findings of this paper are as follows.

First, in terms of policy issuing institutions, the number of pension policies issued by national-level departments such as the Ministry of Civil Affairs, the Ministry of Resources and Social Security, and the Ministry of Finance is high. The Ministry of Civil Affairs, the Ministry of Finance, and the Ministry of Housing and Construction are located in the center of the social network, and these departments often jointly participate in the

formulation of the elderly policy to enhance the scientific as well as perfection of the policy from their own perspective. Secondly, from the perspective of policy characteristics, the existing policies focus on urban and rural elderly people from the perspectives of community, medical institutions, and elderly care institutions to pay full attention to pension insurance, elderly care work, infrastructure construction, and old-age security. Finally, from the perspective of policy evolution, the government has been focusing on making policies from the perspective of public services for the elderly, and indirectly making policies for retirement, investment, insurance and medical care to solve the contradiction between supply and demand of elderly services, in addition, influenced by the epidemic, the government has been focusing on epidemic prevention and control for the elderly since 2020.

This paper explores the content of existing pension policy texts through text mining methods and explores the evolution of existing pension policies from the time dimension. This study is valuable in identifying the characteristics of existing policies and the laws of policy evolution, which can help to grasp the current policy structure, improve the existing policies' shortcomings, and ultimately improve the science of policy formulation.

## ACKNOWLEDGEMENT

## REFERENCES

[1] Strunk B C, Ginsburg P B, Banker M I. (2006). The Effect Of Population Aging On Future Hospital Demand. Health Affairs, 25(3): 141-149.

[2] Chen G, Zhao X Y, Suo H Y. (2021). Opportunities and Challenges for the High-Quality Development of Elderly Services in the 14th Five-Year Plan Period. Administration Reform:1-10. https://doi.org/10.14150/j.cnki.1674-7453.20210115.001. (in Chinese)

[3] Song Y. (2014). Losing an only child: the one-child policy and elderly care in China. Reproductive Health Matters, 22(43): 113-124.

[4] Liu R C, Chen Q J. (2016). Path Dependence in Policy Changes: Examination on the Issue of the Pension Equality in China. Academic Journal of Zhongzhou, (11):56-61. (in Chinese)

[5] Zhang Y. (2013). Meeting the ageing challenge: China's social care policy for the elderly. China: development and governance, 2013: 343-349.

[6] Jacobsen F F. (2015). Understanding public elderly care policy in Norway: A narrative analysis of governmental White papers. Journal of aging studies, 34: 199-205.

[7] Zhai Z W, Chen J J, Li L. (2016). Aging in China: General Trends, New Characteristics and Corresponding Policies. Journal of Shandong University(Philosophy and Social Sciences), (03):27-35. (in Chinese)

[8] Wang X J. (2021). Evolutionary Features and Prospects of China's Financial Security Policy for Elderly Services under Population Aging. Economy and Management, 35(01):13-19. (in Chinese)

[9] Yi J, Lu D, Deng Y. (2016). The future of social elderly care in China: from the perspective of service-oriented government. Journal of Service Science and Management, 9(03): 211.

[10] Hicks D, Wouters P, Waltman L, et al. (2015). Bibliometrics: the Leiden Manifesto for research metrics. Nature News, 2015, 520(7548): 429.

[11] Marin A, Wellman B. (2011). Social network analysis: An introduction. The SAGE handbook of social network analysis, 11:25.

[12] Vishal G, Lehal G S. (2009). A Survey of Text Mining Techniques and Applications. Journal of Emerging Technologies in Web Intelligence, 1(1):60-76.

[13] Tan A H. (1999). Text mining: The state of the art and the challenges. Proceedings of the PAKDD 1999 Workshop on Knowledge Discovery from Advanced Databases. sn, 8:65-70.

[14]Wu H C, Luk R W P, Wong K F, et al. (2008). Interpreting TF-IDF term weights as making relevance decisions. ACM Transactions on Information Systems, 26(3):55-59.

[15]Blei D M, Ng A, Jordan M I. (2003). Latent dirichlet allocation. The Journal of Machine Learning Research, 3:993-1022.

[16]Sun Q L. (1992). Boundaries of high-frequency and low-frequency words and word frequency estimation methods. Information Science, (02):28-32. (in Chinese)

[17]O'Callaghan D, Greene D, Carthy J, et al. (2015). An analysis of the coherence of descriptors in topic modeling. Expert Systems with Applications, 42(13):5645-5657.