Research on the Application of Big Data Technology in Public Opinion Management

Zhou Ying*, Li Pancheng

Jiangxi Police College, Nanchang, Jiangxi, 330029, China

*Corresponding Author

Keywords: Big data, Internet public opinion, Management

Abstract: In the information age, big data has brought new opportunities and challenges to the production, life, network and information security of human society. Social public opinion is formed and spread through certain media, which has an important impact on social order and the image of public organizations, so it is often widely and highly concerned by the society. Therefore, how to improve the network public opinion management mechanism under the background of big data becomes very important. Based on this, based on the basic concept of network public opinion work in the era of big data, the author analyzes its problems, and puts forward solutions to these problems by using big data technology, in order to bring about changes in the application of big data.

1. Introduction

With the arrival of Web2.0 era, big data information is developing in a blowout situation. The data generated by the Internet every day can be calculated in PB scale, accumulating huge network data resources, which also means that China has officially entered the era of big data. Internet public opinion is a brand-new concept, which shows most people's views on something. It not only has the characteristics of Internet platform, but also has the particularity of public opinion.Using big data technology to deal with this amount of data is the need of network public opinion monitoring. According to Chen Xuegang, the influence of network public opinion has been infinitely magnified by the rapid development and wide popularization of network technology [1]. Wei Weihua also believes that when an emergency occurs, online public opinion provides a reference for the relevant government departments, but there is also a risk of causing social problems if the guidance is improper [2].

The application of big data has achieved a deeper analysis and more accurate prediction of social public opinion. Therefore, analyzing the past, grasping the present and predicting the future through the innovative way of big data is conducive to improving the decision-making ability of social public opinion governance and making use of big data and its technology for social public opinion monitoring analysis. This paper discusses the application of big data mining technology in network public opinion monitoring, in order to put forward the countermeasures for the application of big data mining technology in network public opinion monitoring, and contribute to the theoretical and practical application of big data and network public opinion monitoring.

2. Mining and Analysis of Big Data Information

If big data is compared to a mineral deposit, then data mining technology is the key to open this mineral deposit, and only the excavated big data is a meaningful data resource. Different types of network public opinion early warning and emergency plans, etc., through the simulation analysis of public emergencies for many times, through different time, events and public opinion places, the types of media, the basic situation of netizens and the development trend of public opinion, further establish the mining mode and analysis mode [3]. And because of the weakening of traditional supervision methods, all kinds of thoughts, cultures and values can be fully expressed, including

  DOI: 10.25236/ieesasm.2020.085

positive and positive public opinion and negative emotions. Although the quantity and diversity of data are reflected in all aspects, the technical means and network speed can realize instant transmission and data search. It is conducive to exploring the solutions based on big data to improve the decision-making ability of social public opinion governance and create a social public opinion environment.

The ideal mode of Internet public opinion management and control based on big data includes three steps: data collection, public opinion discovery and information source location, as shown in Figure 1.

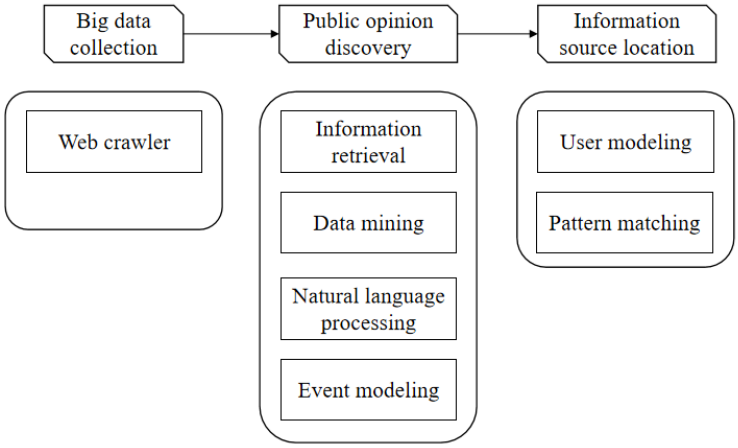

Fig.1 Three Steps of Internet Public Opinion Control

As a big data scene, it is necessary to have massive data as the object of public opinion analysis, and the data acquisition needs to be real-time, otherwise it will be impossible to analyze the hot public opinion. Secondly, analyzing and processing massive data and finding valuable information submerged in noise are important steps in public opinion control.

In the era of big data, we pay attention to speaking with data, and all decisions and actions are based on data. The same is true for network public opinion management under big data, which focuses on mining valuable public opinion information hidden in data for analysis and then carrying out management work. The characteristics of big data are basically consistent with network public opinion. The deep-seated reason lies in that big data and network public opinion analysis have the same goal-timely and accurate prediction. The scenes that need to remove noise from massive unstructured data to obtain information are all big data scenes; Parallel computing technology, distributed storage technology, data mining technology, etc. are all big data technologies that deal with massive data.

3. Applicability Analysis of Big Data and Internet Public Opinion

3.1 The Characteristics of Online Public Opinion Are Consistent with Big Data

The scale of online public opinion information is huge, and it grows exponentially through the big data environment. With the rapid popularization of fast and efficient communication platforms such as Weibo and WeChat in the micro-era, netizens can break the limitations of time and space, publish information anytime and anywhere, and realize people's active choice and two-way feedback of information. People can immediately express their views and positions on events, making the expression of public opinion smoother; At the same time, network public opinion can be pasted quickly, which makes it possible to forward it indefinitely. The openness of the Internet enables news media, self-media and netizens to express their views conveniently and quickly on the Internet. The data of network public opinion monitoring comes from the massive information in the network, and the data of network public opinion is huge [4]. For example, through the collection and retrieval of big data information, the corresponding public opinion keywords are input, and real-time supervision is set. When relevant words appear in the corresponding social platforms and online media, they are immediately recorded and corresponding feedback logs are formed. Because

of its standardized SQL language and integrity constraints, it has good data processing performance and data integrity.

3.2 There is a Strong Demand for Big Data Analysis and Processing Technology to Realize the Monitoring of Network Public Opinion

Because of the low processing efficiency, small amount of information and old monitoring mode, the setting of public opinion early warning scheme does not conform to the corresponding laws, resulting in the situation that there are few public opinion early warning schemes. Big data has the characteristics of low data value density, and the network public opinion monitoring needs to find valuable network public opinion information from massive network data, which also shows the characteristics of low data value density. Make the “data-driven decision-making” model objective and feasible; It is possible to dig deep into the text data of online public opinion to find hidden public opinion, which makes it possible to judge “panoramic” in public opinion guidance [5]; Only on the basis of fully realizing the formation and dissemination rules of speech, can we effectively guide and manage the network public opinion. Through the analysis of social public opinion data, we can monitor, warn and guide and control social public opinion, so as to improve the ability of social public opinion governance and guide the mainstream values and public opinion.

3.3 There is No Difference in the Process of Value Increment

The network public opinion can be quantified by using the big data characteristics of the network public opinion, and the goal of public opinion quantification is to analyze the relevance of public opinion, and then realize public opinion prediction. It is necessary to make full use of big data technology and massive social public opinion information, and study the acquisition and identification, monitoring analysis and early warning of social public opinion information according to the internal mechanism of the generation, development, evolution and decline of social public opinion. For example, by comparing and linking the data collected from website news, forums, blogs and microblogs according to certain indicators, we can analyze the diffusion rate and frequency of public opinion in different groups, so as to know which groups are easier to spread which kind of public opinion. By focusing on information collection in public opinion sensitive network communities, we can fully tap valuable information, analyze the correlation degree of fragmented information, open up cross-domain information fusion channels, realize the information flow among the government, media and the public, and make preparations for effectively guiding online public opinion.

4. Internet Public Opinion Management Based on Big Data

4.1 Big Data Collection

Massive information collected for Internet is the object and foundation of big data analysis. Since Internet public opinion is generally online real-time data, it is necessary to use web crawler technology to comprehensively crawl, store and index the contents of monitored websites and microblogs [6]. In this way, we can not only obtain the corresponding demographic characteristics and behavioral information of the whole network, but also mine the psychological characteristics of netizens by using the text-based emotional analysis methods and techniques in big data technology. Many insightful views have been put forward on the research and solution of the “butterfly effect” of social public opinion in the virtual social field under the big data environment, and the local problems are more likely to get out of control and rapidly evolve into global crises. The big data processing engine Map Reduce technology can complete the timely processing of massive network data, and Map Reduce technology can be subdivided into three levels [7]: A. distributed file system; B. parallel programming model (Map function, Reduce function); C. parallel execution engine. The specific calculation process is shown in Figure 2.

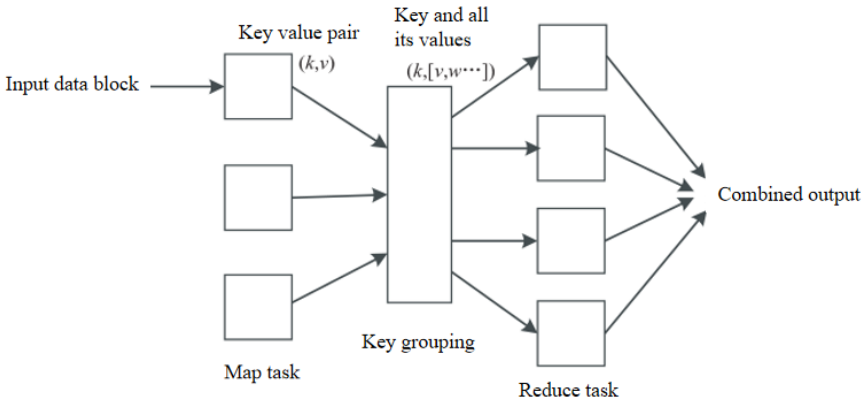

Fig.2 Map Reduce Calculation Process

Big data is not only a huge amount of data generated based on the development of Internet technology, but also a technical means of data storage, mining and analysis, not a pure technical phenomenon. We should not only consider the attributes of big data technology environment, but also the changes of political, economic, technological, social, psychological and policy environment. Most importantly, we can not only get personal psychological files from personal data, but also locate specific people from personal data in reverse. Geographically speaking, it is necessary to pay attention to relevant news in large and medium-sized cities, key areas and other regions; From the industry point of view, it is necessary to pay attention to the relevant news of industry authorities, associations and other institutions. Strengthening supervision of public opinion is conducive to the overall building of public opinion system, the timely control of public opinion and the prevention of further fermentation and escalation of events.

4.2 Search Engine Network Log Data Mining

When an event causes Internet attention, netizens search for event keywords in the search engine to browse related information. At this time, the website server of the search engine records search logs including browsing time, IP address and search terms of netizens. Using big data thinking and big data technology, this paper constructs a public opinion monitoring model from the perspective of public opinion data, and divides the network public opinion monitoring work into a three-step model of “data collection, data processing and data analysis”. Using the task scheduling mechanism of HDFS to improve the scalability and fault tolerance of the network public opinion text mining module system, so as to solve the horizontal expansion problem of large-scale network data analysis and facilitate the information sharing among multiple parts; From the data correlation analysis of the generation, development, evolution and decline of social public opinion, this paper studies the division of labor, cooperation and dynamic adjustment rules among various governance organizations, and studies the operation mechanism of social public opinion governance organizations and the application of policy tools.

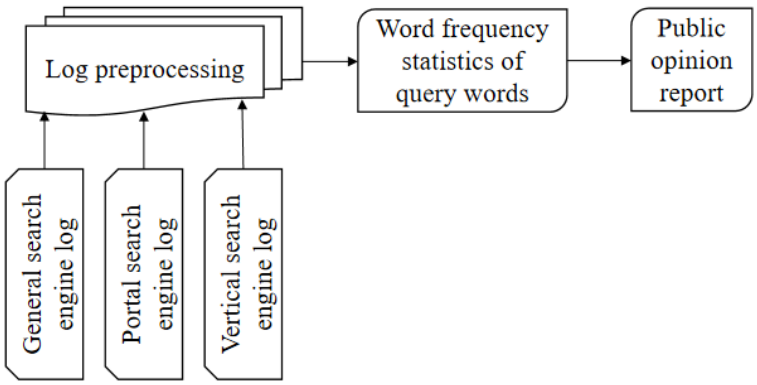

Fig.3 Public Opinion Monitoring Model Based on Search Engine Network Log Data Mining

There are three main steps in online public opinion monitoring using search engine logs. The first step is to preprocess the log, clean up the log, and keep the data needed for public opinion analysis. The second step is keyword statistics. Time or region is the dimension of keyword query frequency statistics. After completing this step, you can find your search situation and find popular unknown keywords for specific query words in specific time or region. The third step is to organize the opinion analysis report and its contents. The ranking of hot spots in a specific time period or area, such as hot topics in a day, hot topics in a week, and a time series diagram of the changing trend of the search amount of specific keywords in a specific time period (Figure 3).

People have social attributes, which make big data have corresponding social properties. In fact, the characteristics of non-structure, timing and diversity of big data are the diverse expressions of Internet users' behavior and speech. In the process of collecting public opinion information, the way of thinking changes from pursuing accuracy to fault-tolerant thinking. In the massive public opinion data, absolute accuracy at the micro level is properly ignored, but it can have stronger insight and execution at the macro level. There is no deep knowledge hidden behind the data, let alone the semantic level of public opinion information, and the system is not highly intelligent. By analyzing and counting the search logs, we can find the hot spots in a certain period of time and a certain area, and then form a trend chart of search volume change and a search ranking list, so as to realize the network public opinion monitoring.

4.3 Analysis of the Subject State of Public Opinion Information

The main body of online public opinion information is human, and its state changes with the change of experience process and attitude of public opinion events, and this change can be followed regularly. A text corpus can be represented by a large-scale sparse TF-IDF matrix, and a dimension reduction method can be applied to represent compressed data. At present, there are two main methods of dimension reduction. Include latent semantic index and probabilistic topic model. Perfecting the big data platform is also a better way to store the differences in the development of online public opinion in different places. A linkage mechanism can be formed through the big data platforms in different places to ensure that there are more than two public opinion early warning schemes and emergency treatment schemes between public emergencies. Big data can gain insight into every netizen's media contact habits and usage preferences, match massive information with personalized needs through algorithms, and achieve accurate information push and efficient distribution. It can be analyzed anytime and anywhere, combining with new data and new algorithms, which is beneficial to timely mining of network public opinion.

SIR is the most classical model for studying the spread of infectious diseases [8]. In which s represents susceptible person, I represents infected person and r represents removed person. The subject of public opinion information can also be divided into three different states: S, I and R, as shown in Figure 4.

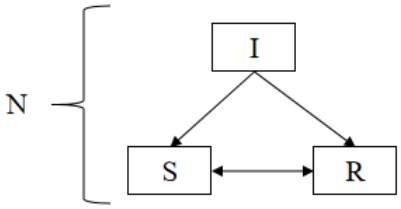

Fig.4 Sir Diffusion Model

According to the above model, we can have a clear understanding of the internal mechanism of network public opinion communication from the level of public opinion subject state, and at the same time, it has an important guiding role in accurately judging the public opinion information subject state, and then implementing different guiding strategies for different state subjects.

In addition, it should be noted that the object of public opinion analysis is human language. Analyzing semantics and emotions from a large number of languages is the difference between public opinion analysis and traditional data mining. Under the condition of preserving text

similarity, a low-level approximation of the original matrix is constructed. A document data block is a collection of a series of elements, but the same element cannot be stored across data blocks. In the system, the input of all Map tasks and the output of Reduce tasks are in the form of key-value pairs. Therefore, when major sudden mass incidents and public incidents occur, public opinion management departments should select media to respond to online public opinion through big data support analysis, grasp the right to speak, and seize the position of online public opinion, which will greatly shorten the time for online public opinion to subside.

5. Conclusion

Internet public opinion has become one of the main factors that undermine social harmony and stability and affect the harmonious development of society. Big data mining technology is the key technology of network public opinion monitoring, and Internet search engine provides a new method for network public opinion monitoring. Big data basic technology is used for platform construction, providing the most basic functional support for computing and storage. In order to further control the development of online public opinion, improve the control ability of public opinion again, and greatly reduce the negative social impact brought by public emergencies. Therefore, we must use big data technology to deeply optimize the existing network public opinion information working mode.

Acknowledgment

Technology research projects in 2019 (Research on the application of public opinion management based on big data technology in public security crime prevention

References

[1] Chen Xuegang. Research on the rapid self-clustering method of microblog public opinion based on big data technology. Journal of Information, vol. 36, no. 005, pp. 113-117, 2017.

[2] Wei Weihua. Research on the Management Mechanism of University Students' Network Public Opinion Based on the Background of Big Data. China Adult Education, no. 17, pp. 69-73, 2017.

[3] Ying Yi, Liu Dingyi, Ren Kai. Research on Public Opinion Analysis and Decision Support System Architecture Based on Big Data Technology. Library and Information Guide, vol. 2, no. 009, pp. 32-36, 2017.

[4] Sun Lingfang, Yin Peipei. Research on the Emotion Strength of Internet Public Opinion Based on Big Data Technology. Computer and Digital Engineering, vol. 046, no. 001, pp. 160-166, 2018.

[5] Yuan Jiang, Liu Tongxiao. Application analysis of civil aviation public opinion monitoring based on big data. Science & Technology Information, 2017no. 15, pp. 12-13.

[6] Diao Shengfu, Feng Guifeng. The application of big data in public opinion monitoring: value, limitations and beyond. Inquiry, vol. 000, no. 003, pp. 83-88, 2018.

[7] Zhang Yajing, Fu Jun, Liu Hao. Research on the realization path of public opinion information collection, research and judgment management system based on big data technology. Consumer Guide, vol. 000, no. 003, pp. 49-50, 2019.

[8] Sun Shouqing. Discussion on the Application of Public Opinion Big Data in News Production. Think Tank Times, vol. 189, no. 21, pp. 262+265, 2019.