



The Design and Application of Network Democratic Survey System under the Background of internet plus Democracy

Yangmei Cheng^{1, a}

¹Sichuan Vocational College of Judicial Police, Deyang, Sichuan, China

^ae-mail: 104355944@qq.com

Abstract. At present, as an important way to express people's public opinion, network democracy still has many shortcomings in practical application, process control and scope control. In this paper, network democracy will be the research object, based on Web technology, and text mining technology will be adopted to complete the construction of network democracy investigation system in Java development environment. The design and development of the overall system framework will be completed according to MVC mode, using Spring framework under J2EE specification, and the collection, cleaning and storage of democratic text content on the Internet will be completed in turn by using software programs such as WebCollector, ICTCLAS and MySQL. Then, the mining and analysis of the text content and the visual display of the final results will be completed by algorithm models such as TF-IDF, K-means and Pearson. The system will effectively improve the efficiency of analyzing and summarizing the contents of electronic texts in mass network democracy, improve the accuracy of the analysis of network democracy survey, promote the effective expression of real public opinion according to the survey results, aggregate real demands, master the orientation of public opinion, and provide necessary technical support for the formulation and implementation of public policies.

Keywords: Democracy in internet+; Web technology; Text mining; Machine learning

1 Introduction

In the infinite open and decentralized Internet space, people's information processing and knowledge acquisition have undergone subversive changes, and the cognitive system established through long-term accumulation has gradually been replaced by explosive information overload, so that people's self-consciousness is more freely expressed, their opinions and opinions are more polarized, social thoughts are more diversified, and public opinion hotspots are more value-oriented, and they can become more and more ways to express public opinion demands under different interests. [1]

Compared with the traditional expression of public opinion, the network media has the characteristics of rapid communication, strong interaction, wide influence, openness and transparency, and free expression, and has been continuously innovated and integrated with the current public opinion collection system, forming a new form of "internet plus democracy". On the one hand, the birth of network democracy is attributed to the change of people's thoughts in the stage of social structure change, and on the other hand, it depends on the rapid expansion of the scale of netizen groups in China. On the whole, the rise and development of China's cyber democracy have Chinese characteristics, and its role and influence also have obvious advantages. But the network democracy depends on the network information technology, and its essence is still a kind of virtual data information, which has obvious two-sidedness. In the process of practical application, over-reliance on network information technology will also hinder the expression of public opinion, and its authenticity, participation, representativeness, equality, normative control and institutional binding force all need to arouse our high attention, which is also a major issue before us. For this reason, this paper holds that, in view of the special forms of network democracy and the many challenges it is facing at present, government departments should actively play a guiding role, and actively cite the new generation of network information technology to build a network democracy investigation system and form a set of practical and effective comprehensive application solutions. The system will be based on Web technology, with text mining technology as the core. By collecting, cleaning, storing, analyzing and mining the electronic text content of network democracy, it can obtain effective analysis and induction results, effectively improve the mastery and control of people's public opinion demands, further expand the practice form of network democracy, promote scientific decision-making by government departments, and promote the continuous, healthy and orderly development of socialist democracy.

2 Introduction of key technologies

2.1 Web technology

The Web is a distributed network information service system established in the Internet environment, which can publish, browse, query and other functions of data information through graphical interactive pages. The functions of the Web system will depend on the construction and operation of the Web architecture. The essence of the Web architecture is an interactive process of data information, that is, the processing of the request content and the presentation of the response content are completed under certain pre-conditions. Figure 1 shows the flow chart of the "request/response" processing of the Web architecture.

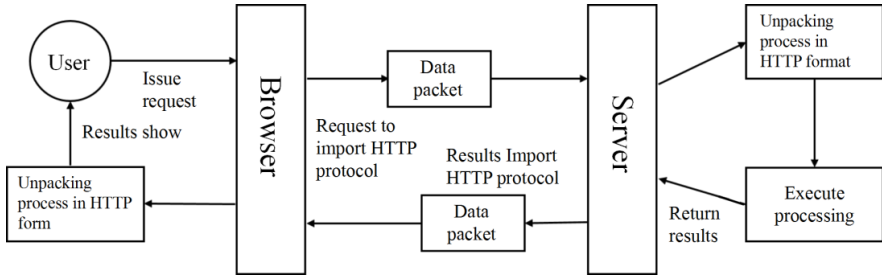


Fig. 1. Request/Response process for the Web (original)

Based on the Web architecture, all technologies applied in the construction process are collectively referred to as Web technologies. In the process of development, Web technology has gradually divided into two categories: client technology and server technology. The client technology aims to complete the design and implementation of user interactive operation pages. The server-side technology is mainly responsible for designing and constructing business execution logic and data processing methods in the server to provide services for the interactive operation of front-end users.

2.2 Text mining

The Text Mining refers to the process of extracting unknown, understandable and finally available knowledge from a large amount of text data, and at the same time using this knowledge to better organize information for future reference. [2] The object of text mining is a large amount of unstructured text information. On the basis of the theories of mathematical statistics and computational linguistics, with the help of a large number of computer hardware settings and software systems, the discovery and analysis of the rules of text appearance, semantic expressions, grammatical relations and other contents are completed.

The text mining is driven by task application, which mainly includes text data collection, preprocessing, feature information extraction, analysis and mining, and result display. The flow chart of text mining is shown in Figure 2. The main contents of text mining include text content extraction, text summarization, text classification, text clustering, association analysis and so on.

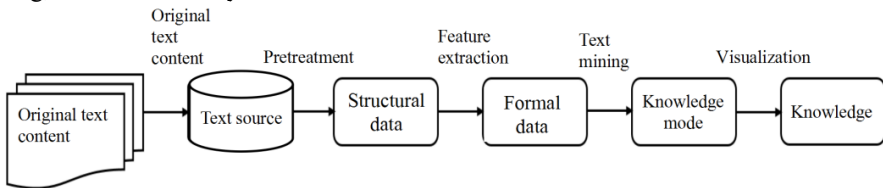


Fig. 2. Illustration of text mining process (original)

2.3 J2EE

The J2EE (Java 2 Platform Enterprise Edition) is an architecture that uses Java platform to simplify the complex problems related to the development, deployment and management of enterprise solutions. It is also a set of enterprise-level distributed application development specifications led by SUN Company. [3]

In the J2EE specification, the design and development of Web applications will be carried out in strict accordance with the layered idea of MVC pattern. The MVC design pattern divides the whole system into three parts: Model, View and Controller. Among them, the model represents data and business logic, is responsible for the processing of each request task, and returns the data results obtained after the processing to the view. The view refers to the functional page that users can see and realize interactive operation, and can receive requests from users, and display the data results returned by the model. As a bridge between the model and the view, the controller can control them to coordinate. On the one hand, it can receive the user request input from the view layer, and on the other hand, it can call the corresponding model to complete the processing of the user request. The MVC design pattern has obvious advantages, with low coupling, high reusability and applicability, and can reduce the difficulty of overall system development, improve work efficiency and realize the agile development of the system.

2.4 Development process

According to the system development requirements and the use requirements of the above key technologies, complete the configuration and deployment of the network democratic investigation system development environment. The overall system design and development environment is divided into two parts. One is to complete the design and development of related algorithm models according to the text mining process; Secondly, under the Java language environment, the integration and encapsulation of various functional modules are realized under the Spring framework, forming a complete Web application, which is deployed on the server side to realize the release of the system.

As an important form of expression of public opinion, online democracy has diversified channels and approaches, including government portals, forums, post bars, chat rooms, Weibo, WeChat, QQ space, personal websites and e-mails. [4] However, its core content belongs to electronic text information, which meets the selection requirements of text mining objects. During the process of text data extraction, the system will use the Handler information mechanism of WebCollector crawler and URL generator to complete the traversal of URL address. The Handle mechanism is more concise and practical than other large crawler frameworks, and the content crawling function can be realized only by declaring customization and sending it to the traversal machine.

In the electronic text information collection, it will be stored in MongoDB, and the system will use ICTCLAS tool to complete the pre-processing of electronic text information content. The main process includes Chinese word segmentation and stop words removal. According to the corpus of probability statistics, ICTCLAS will use HHMM

algorithm to automatically read the original text information, disable the initial quotation of vocabulary, divide Chinese words, filter stop words and output the results.

In the stage of feature extraction, the system will use TF-IDF algorithm to calculate the preprocessed text. The TF refers to the vocabulary that appears most frequently in the text content, that is, the vocabulary frequency, while IDF indicates the weight of a certain vocabulary in different text content, that is, the reverse document frequency. The system uses the API interface of TF-IDF designed and developed by itself to realize the statistics of TF-IDF value of each word. The key implementation code is shown in Figure 3.

```
private List<Map<String,Double>> calTfidf() {
    List<Map<String,Double>> tfidfRes = new ArrayList<Map<String,Double>>();
    for(Map<String,Integer> documentTfMap : documentTfList) {
        Map<String,Double> tfidf = new HashMap<String,Double>();
        for(String word : documentTfMap.keySet()) {
            double tfidf = idfMap.get(word) * documentTfMap.get(word);
            tfidf.put(word, tfidf);
        }
        tfidfRes.add(tfidf);
    }
    return tfidfRes;
}
```

Fig. 3. Code for calculating TF-IDF value of keyword (original)

In the stage of text analysis and mining, the system mainly supports classification mining, cluster extraction and relevance analysis. As for the classification and mining of network democratic electronic text information, the system will select naive Bayes classifier according to the TF-IDF value of the text content to complete the classification processing. But for text clustering mining, the system will use K-means algorithm to complete the clustering process. In addition, Pearson algorithm is used to complete the correlation analysis of online democratic participation.

As for the development of the whole system, Linux is the operating system, CentOS is the version, Java is the basic development environment, JDK version 1.8.0_91 is the development kit, Apache 2.4 is the Web server, IntelliJ IDEA is the Java integrated development tool, and MySQL is the database. After the above software systems are installed and configured one by one, the construction of Web application development environment is completed. The system uses Maven 3 to manage the project structure, divides the whole project into several engineering modules, and completes the design and development of the whole system based on Spring architecture. Through the introduction of the above key technical theories, the overall environment of the system development, the configuration of related software and tools are determined, and the technical feasibility of the overall project of the network democratic investigation system is also clarified.

3 Detailed function realization

The user of the system is supported to complete account registration and identity verification by submitting information, and to log in and use the system with unique iden-

tification information. In order to improve the use security of the system, the user password will be encrypted by hash algorithm, and the user login authentication will be completed by this authentication method. The key code of realizing MD5 encryption algorithm in Java language is shown in Figure 4. [5]

```

public static String getMD5(String message) {
    MessageDigest messageDigest = null;
    StringBuffer md5StrBuff= new StringBuffer();
    try { messageDigest = MessageDigest.getInstance("MD5");
        messageDigest.reset();
        messageDigest.update(message.getBytes("UTF-8"));
        byte[] byteArray = messageDigest.digest();
        for (int i = 0; i < byteArray.length; i++)
            { if (Integer.toHexString(0xFF & byteArray[i]).length() == 1)
                md5StrBuff.append("0").append(Integer.toHexString(0xFF & byteArray[i]));
              else
                md5StrBuff.append(Integer.toHexString(0xFF & byteArray[i])); }
    } catch (Exception e) {
        throw new RuntimeException(); }
    return md5StrBuff.toString().toUpperCase();
}

```

Fig. 4. Key code of MD5 encryption algorithm implementation (original)

3.1 Initiation of investigation

When the user enters the system, he will complete all the operations of the system step by step according to the operation guide. In the survey initiation function module, the system supports users to determine the corresponding survey scope by selecting or inputting survey keywords by themselves. Among them, the scope of investigation can include, but is not limited to, government portals, official forums, Weibo, WeChat and so on. After the user selects, the system will automatically crawl and collect all relevant contents within the scope according to the keywords, and will automatically save them in the database after the completion. After the crawling is completed, the system will also support users to search and view electronic text information.

3.2 Network Democracy Survey

In this function module, the system will support three functions: classification of public opinion expression, clustering of public opinion expression and correlation analysis of online democratic participation. The expression clustering of public opinion means that the system inputs the vector expression of key characteristic words of text information into the classifier model to complete the case classification. After the classification, users will intuitively obtain structured data display, and quickly complete the work of network democracy theme classification, theme detection and public opinion tendency analysis. [6] For the clustering analysis of public opinion expression, the system will also input the vector expression of key words of the text into the clustering algorithm model to complete the clustering of cases. According to the clustering results, users can quickly classify a large number of public opinion expressions without categories, and

classify similar public opinion expression appeal information into one category, so as to automatically generate the core topic words of this category. Through classification and clustering operations, users can quickly grasp and understand the sensitive topics, hot topics and key topics under specific events involved in network democracy, which is convenient for the analysis and formulation of subsequent public decisions. In addition, the correlation analysis can further clarify the corresponding relationship between netizens' identity and online democratic participation, which is convenient for users to accurately control the survey scope in the follow-up online democratic survey.

3.3 Visual display

The visualization class library of D3.js system is used to complete the rendering of the classification results of online democratic electronic texts. The D3.js follows the existing Web standards and can run independently in the client browser. It combines powerful visualization components to drive DOM operations. The D3.js support includes scatter charts, pie charts, ring charts, and can also support the dynamic advanced display of some charts, as shown in Figure 5, which is the network democracy cluster analysis effect chart.

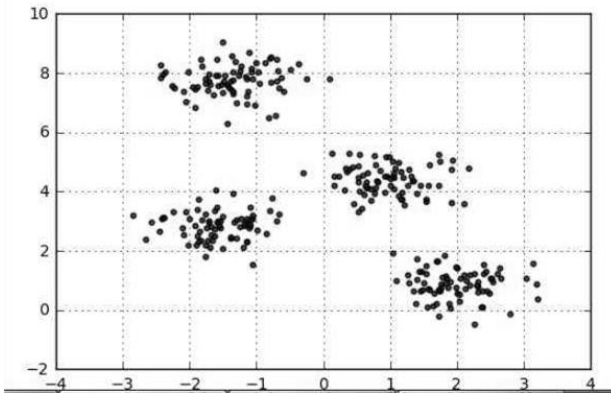


Fig. 5. Cluster analysis results of network democracy (network)

4 Conclusions

This paper takes network democracy as the research object, aims at promoting the scientific formulation and implementation of public decision-making, and builds a network democracy investigation system with the help of the new generation of network information technology and text mining technology. The system will be based on Web technology, and realize the digitalization and informatization construction of online democracy survey under "internet plus Democracy". It will not only help to improve the efficiency of analyzing and summarizing the demands of mass online democracy, but also improve the accuracy of the analysis of online democracy survey. According to the survey results, it will promote the effective expression of real public opinion, aggregate

real demands, and master the guidance of public opinion, which will be a beneficial attempt for the all-round construction of socialist democracy in the new era.

References

1. Chen Lianjun, Ning Xiaoyin. The Connotation, Advantages and Promotion of the Practice of Cyberdemocracy Discourse [J]. Teaching and research of Marxist theory.01:31-40(2022).
2. Li Zhenpeng, Chen Bizhen etc.The Research of Internet Public Opinion Classification Based on Text Mining [J].Journal of Systems Science and Mathematical Sciences.05:813-824(2020).
3. Ke Cai. The Research of Web Application Based on J2EE and MVC Pattern [J]. Software .03:165-167(2020).
4. Li Mingde, Li Meng. Research on Ethical Significance and Realization Path of Network Democratic Participation [J].Journal of Zhejiang Gongshang University.07: 97-106(2021).
5. Zhao Yifan, Bian Liang etc..The Java implementation of salted message digest based on MD5 [J].Software Guide.03:214-220(2018).
6. Chen Caian, Zeng Bigui. The Application of Data Mining Technology in Public Security Public Opinion Analysis [J]. Fujian Computer.11:86-88(2015).

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

