

The Design and Application of Public Opinion Monitoring System

Hongfei Long

College of Marxism, Shanghai Ocean University, Shanghai 201306, China

paperiset@163.com

Keywords: Student public opinion, focused crawling, data analysis, opinion monitoring.

Abstract. This paper discusses the formation of student public opinion from the foregoing background, and studies the sources, types and character of the Internet education public opinion, then discusses the strategies for accessing and filtering the sources link and content of the Internet education public opinion based on that, and then designs and realizes the monitoring system of the Internet education public opinion.

Introduction

Education is important for people's livelihood, in relation to the future of our country and the hope of our nation; it has a direct bearing on people's interests and daily life, so it is the focus of the national attention. [1] With the development of network and communication technologies, Internet has accessed to more areas, Internet has accessed to more areas, and been used more widely, Internet became popular and fashionable more and more, Internet is a very fast and convenient way of shaping public opinion for people express ideas and outburst emotion, public opinion online forms very rapidly, and it's scale and influence are not to be overlooked either [2]. Public opinion is a reflection of public thinking, emotion and sentiments. How to get the information of public opinion, how to understand the public's desire for education, that is the great practical significance for the education developing better.

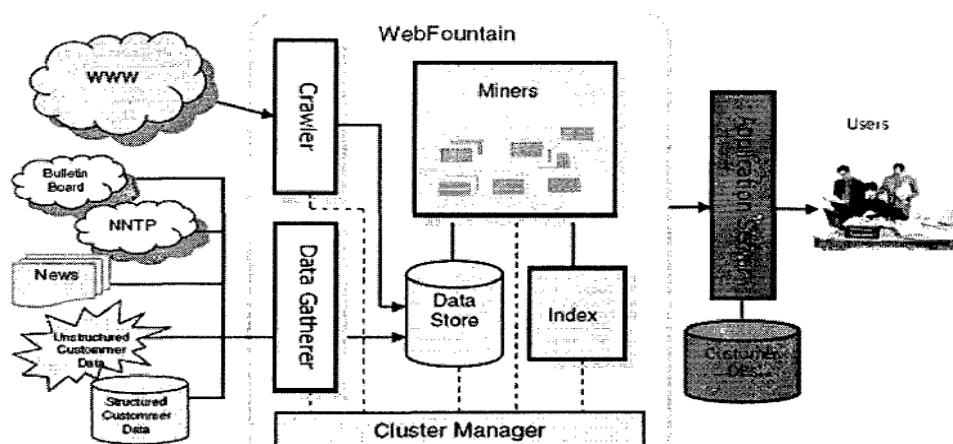


Fig. 1 WebFountain system structure

Fig. 1 is a broad opinion monitoring, and goes by the name of WebFountain. It is a large text analysis tools. The Student Opinion Monitoring System is an automation systems using information technology to monitor public opinion [3]. It is used to replace traditional manual monitoring system, and it can greatly expand the media coverage, and improve the efficiency and quality of public opinion monitoring. It can provide technical support for enterprise decision-making and crisis management. After the first phase of R&D work, The Public Opinion Monitor System has achieved many features, including data high-frequency acquisition, data preprocessing and immediate warning of important and sensitive information. [4] It also realized some analysis functions, including Content classification, positive and negative analysis, topics word frequency analysis, information dissemination analysis. It established a website for employees to browse information.

Evaluation

As the influence of the Internet is increasing today, the negative public opinion may break out within a short time and go viral, which brings serious negative effect to reputation. The traditional scheme of public opinion monitoring for student is to arrange special public opinion monitoring staff searching the negative information on the Internet manually. [5] The solution's weaknesses include the incomplete and not prompt monitoring, the inaccurate analysis, and the high costs. So the student demands on the Internet public opinion monitoring system. The requirements include to automatically collecting the huge amounts of Internet public opinion information with real-time, to identify the key information, to aggregate and analyze the information, and to notify the relevant personnel the analysis results promptly.

The author collected and analyzed the corporate requirements, and divided them into public opinion acquisition, public opinion type judging, sensitive article identification, data statistic and analysis, public opinion browsing, charts display, public opinion health state calculating, public opinion information searching, notification and alarm.

$$w = \frac{(\log C + 1) \log \frac{N}{n}}{\sqrt{\sum \left[(\log C + 1) \log \frac{N}{n} \right]^2}} \quad (1)$$

This formula is mainly to calculate important degree of a w entry in an article. Numerical is the term weights with respect to this article, in the concrete operation in the calculation of the value of TF (*Term Frequency*) is the entry. The number of occurrences is divided by the document length (the number of all words). IDF (*Inverse Document Frequency*) express as:

$$IDF = \log \frac{N}{n} \quad (2)$$

Design and development

Based on the study of domestic and abroad research status about theory and application of public opinion, and learning the characteristics of similar products in the other units which have been running, the feasibility of the network public opinion monitoring system development is analyzed with the basic principles and methods of modern software engineering. Then the needs of various user roles are refined, and core business processing model are established. After completing the system architecture and functionality of the software design, the system functions are developed in Visual Studio 2008 platform with dot net Framework. Finally, the system is tested through scientific and rational test cases that could find potential errors that had been corrected at last, which could ensure the system to meet the require of reliability, ease of use and security.

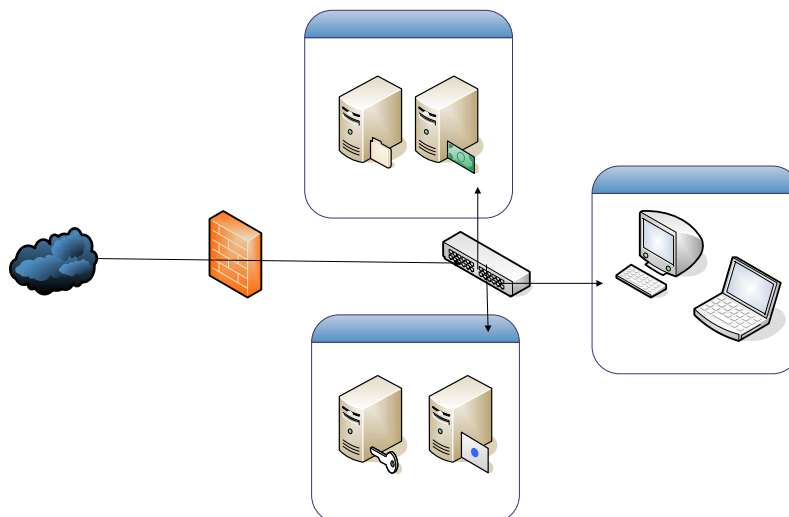


Fig. 2 Public opinion monitoring system structure diagram

Public opinion monitoring system structure diagram is shown in Fig. 2. Key code is following as below:

```
foreach (DataRow dr in drArr)
    {if (dr["Pid"].ToString()=="0")}
{   SqlConnection corm=BC.CreateConnection();
    SqlCommand comm=new SqlCommand("SELECT COUNT(*)FROM tb_user WHERE
    userID={loginName AND userPwd={loginPwd", cone);
    comet. Parameters. Add (new SqlParameter ("C}loginName", SqlDbType. VarChar, 20));
    comm.ParametersE("{}loginName"].Value=loginName;
    comm.Parameters.Add(new SqlParameter("{}loginPwd", SqlDbType.VarChar, 20));
    comm.Parameters["CloginPwd"].Value=loginPwd;
    comet. Connection. Open();
    int i=(int)comm.ExecuteScalar();
    comet. Connection. Close():return i
```

Public opinion has become an information technology research focus at present one of the network, especially the use of information technology networks to carry out excavation and analysis of public opinion has become the top priority of the network event study public opinion. At present, scholars of public opinion on the network and early warning is still in the initial exploration stage. Though the part of public opinion has been established to monitor platform already has a public opinion information retrieval, classification, clustering, filtering, summary and other basic information mining function, but overlooks the public opinion analysis of the core content, information contains ideas and attitudes analysis. In view of this, the writer's emotional characteristics of public opinion as a starting point, will focus on the emotional topic of research tracking technology, will target the web2.0 network environment to achieve set collection, processing, analysis, monitoring in-one platform for the development of public opinion on.

Application of opinion opinion monitoring

Information technology has become released, access and transmission of the main carriers of information. Beyond Web 1.0 user browsing information technology, Web2.0 has allowed users to interact with information on activities, Internet users via the Internet in forums, blog and other online communication platforms on a variety of topics free to express their views. However, public opinion can not be overlooked is the network events but also the attitude of the users, the social stability of a very significant impact on the breadth of its occurrence, transmission speed, always surprising.

This part is the important content of this paper, mainly included monitoring system design and realization of Internet education public opinion, it is divided into two chapters to illustrate a system design and system realization, the monitoring system design of Internet education public opinion, included the system framework design, system workflow design, the system key module analysis and design, etc, and then give the realization design scheme of the system key points; the realization part of monitoring system for Internet education public opinion, mainly introduces systems development tools and development environment, web page crawling realization of the education subject, the revision of Chinese words segmentation, the realization for identifying and accessing the education keywords, as well as the integration and realization of the whole system; The fifth part are the conclusion and outlook, is mainly to conclude this paper work that has done, to provide some suggestions and ideas for the system of the future development and research work.

E.g. Public opinion collection timing diagram is shown in Fig. 3. The Concealment and extensive features of the Internet public opinion limit its applications, as well as network monitoring public opinion need higher requirements. This paper focuses on the design and implementation of the

network public opinion supervision and management system in order to improve the application of the relevant technologies. In this system, we use the public opinion collection, public opinion analysis and tracking, inclination of emotions and public warning system technologies to meet the demands of the client.

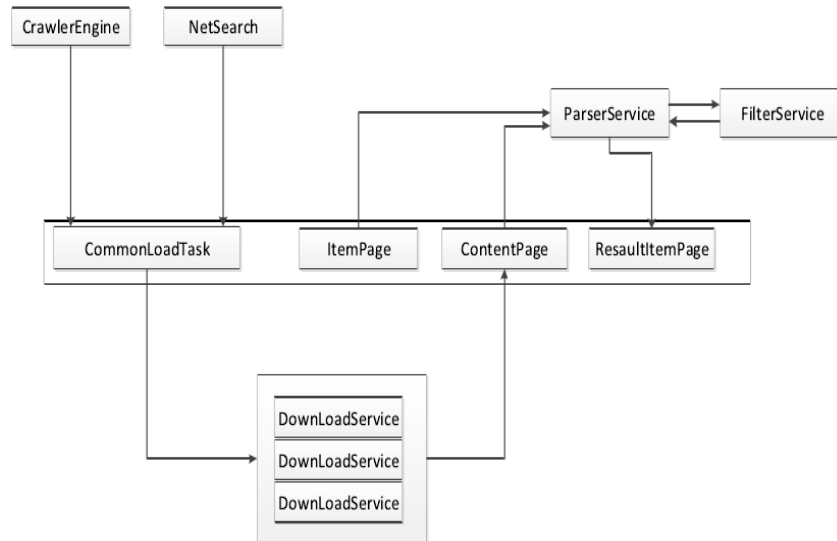


Fig. 3 Public opinion collection timing diagram (from www.google.com)

For the Student is open and virtual, some of the characters of the public opinion such as free, spreading rapidly and hidden, make it difficult to apply. After an in-depth study of the relative technologies of public opinion monitoring, a monitoring frame for Internet public opinion is proposed. Moreover, around forum crawler the key-technology of public opinion monitoring, based on a great deal of analysis of the structures of the forums, a level model for the forums was created, and then a forum crawler was implemented to fetch different forum sites.

To construct the framework for Internet public opinion monitoring, an in-depth of research was done on the relative technologies about it, such as IR, NLP and text mining. And based on this framework, an Internet public opinion monitoring prototype system is established on the distributed programming architecture Hadoop. In order to solve the case that the public opinion will distribute in many sites of different types, and take the breadth and the relevance of public opinion collection both into account, several kinds of ways were used to fetch pages. Finally, the framework uses an ontology-based text mining method to find hot-point.

Conclusions

This paper introduced the implementation process of Public Opinion Monitoring System in detail. Firstly, based on the traditional manual way of network public opinion monitoring, it told the process of collecting requirements of the system. Then, according to the system requirements, it described the design of the system architecture, all the function modules and data storage structure. Finally, it described the detail design and implementation of data collection, data analysis and the website of the system. The architecture design draw lessons from technical experience of well-known enterprises in the industry. Each layer of the system had a good scalability, so as to lay a good foundation for future business development. The paper represented a lot of experience in data collection, data analysis, architecture design and server load.

References

- [1] Information on <http://www.cs.pitt.edu/~swapna/papers/OpinionFinder-extendedabstract.pdf>.

- [2] Guerriero F, Musmanno R. Label correcting methods to solve multicriteria shortest path problems. *Journal of Optimization Theory and Applications*, 2011, 03:519-528.
- [3] Sebastiani, F. Machine learning in automated text categorization. *ACM Computing Surveys*, 2012, 34(1): 1-47.
- [4] Joachims T. Text categorization with support vector machines: learning with many relevant features: *Proceedings of the 10th European Conference on Machine Learning*, 2008: 137-142.
- [5] Studer R., Benjamins V R. Fensel D. *Knowledge Engineering: Principles and Methods*. *Data and Knowledge Engineering*, 2010. 25(122): 161-197.