A Survey of Internet Public Opinion and Internet New Words

KEFEI CHENG, XIANZHI WEN and KE ZHOU

ABSTRACT

The rapid development of Internet new words has had a profound impact on traditional language and everyday expressions, which has gradually attracted the attention of the whole society. Internet new words often come from social hot events, and their emergence and use reflect people's attitude towards hot events. Therefore, it can be said that the outbreak of Internet public opinion is often closely related to the rapid update of new network words. To this end, this paper summarizes the research methods of Internet public opinion and Internet new words, analyzes the relationship between Internet new words and Internet public opinion, and puts forward several challenges of Internet new words for the control and management of Internet public opinion.

KEYWORDS
Internet public opinion; Internet public opinion analysis; Internet public opinion management; Internet new words; new word recognition.

INTRODUCTION

With the development of the times and the progress of network technology, the emergence of Internet new words has become an inevitable language phenomenon. Internet new words refer to a kind of new Internet words and phrases. It aims to present boundless and various content with unlimited variety of formation. According to various unique combinations of distributed and meaningless speeches, some new morphemes can be generated. These morphemes compose the neology and can reflect the Internet public opinion [1]. The Internet new words are usually divided into several categories: homophonic type, phonetic abbreviation type, pictographic action type, English abbreviation type, digital type, industry special type, and etc. Such as “Blue thin mushrooms,” “old driver”,” dog band” and so on. These Internet new words represent the hot event, but also reflects people’s attitude towards the hot spots and new things.

Research and facts have proved that social networks are the most active place for Internet public opinion, however, the outbreak of Internet public opinion is often associated with fast-updating Internet new words, because Internet new words are...
always generated in social hot events [1]. The generation rules of the Internet new words can help the government respond to the social problems timely and make public opinion guidance work quickly and even ahead of time. By studying the relationship between Internet new words and Internet public opinion, studying the emotions and public expressions of Internet new words, we can better understand the social significance of network vocabulary changes, timely understand the central issue of social contradictions and the hot point of public opinion, and understand the social mentality of the denizens.

This paper firstly summarizes the research emphases and related research methods of Internet public opinion and Internet new words. A summary of Internet public opinion research mainly from the Internet public opinion analysis method, the Internet public opinion dissemination evolution method, the network public opinion monitoring and early warning method and the Internet public opinion management control method four aspects. The summary of Internet new words is divided into two aspects: natural linguistics and computer information technology. Finally, the paper discusses the challenge of Internet new words for Internet public opinion management.

INTERNET PUBLIC OPINION AND ITS KEY TECHNOLOGY

Internet public opinion refers to the collection of people's cognition, attitude, emotion and behavioral tendency for various events on the Internet [1]. Internet public opinion analysis is a composite subject, and its research covers many fields, such as journalism, communication, sociology, politics, management, information technology, and etc.

Scholars have studied the Internet public opinion from various aspects and published some academic achievements. The main research areas of it can be summarized as four aspects: Internet public opinion analysis, Internet public opinion dissemination and evolution research, Internet public opinion monitoring and early warning research and public opinion management control research. Also, the main methods and techniques of public opinion research include: Internet public opinion information collection and extraction, text data mining, content analysis, network text orientation analysis technology, topic discovery and tracking technology.

Internet Public Opinion Analysis Methods

There are different methods for the analysis of Internet public opinion, including the contents of the Internet public opinion data mining [2], public opinion emotional tendencies [3], and public opinion heat analysis [4] and so on. The object of public opinion analysis includes a variety of categories, which are mainly concentrated in the natural disaster public opinion analysis [5], college public opinion analysis [6], and social mass emergencies public opinion analysis [7] these three aspects.

There are many methods of public opinion analysis, and the collection of public opinion data is the prerequisite of network public opinion analysis. Therefore, the key technologies involved in network public information mining are summarized. Internet public opinion information mining is generally divided into the following process to achieve, as shown in Fig. 1.
Information collection is the premise of public opinion analysis. It mainly collects data from the target site through data acquisition tool. At present, the sources of Internet public opinion data are mainly the social networking platforms such as microblog, news website, BBS. The data collection method is mainly to crawl web content through web crawler. From the function and structure of web crawlers, which are generally divided into generic, dedicated and incremental web crawlers [2]. The core of web crawler is web analysis algorithm, which can be roughly divided into three categories: algorithm based on the network topology, algorithm based on web content and algorithm based on user’s access behavior. In the algorithm based on network topology, PageRank algorithm [8] is the most typical one of all the algorithms based on network topology. It calculates the link degree of web pages recursively to get the quantitative evaluation of the importance degree of each web page. The algorithm based on web content is to analyze the theme relevance of the text content information such as Web page text content, URL string, anchor text, context of anchor text, text feature of parent text, and then decide the reptile crawling strategy. Its typical algorithms include Best First Search, Fish Search, and Shark Search and so on. The algorithm based on the user's access behavior is mainly through the user’s access log analysis, to develop a crawler crawling strategy. There are also crawling strategies based on statistical models, genetic algorithms and Context Graph.

All kinds of webpage analysis algorithms have their own advantages and disadvantages. In the crawling process, the network topology based algorithm ignores the relevance of the subject content, and is prone to "theme drift phenomenon". The algorithm based on web content cannot accurately and effectively predict the value of the linked theme. The algorithm based on the user's access behavior starts completely from the user's access, ignores the differences among individuals and makes it difficult to obtain accurate data.

Information processing mainly deals with the collected text. The collected data is first cleaned, the purpose of which is to filter the noise. Then, a text representation vector model is established through text segmentation and feature selection to form a computer tractable form. Finally, the classification and clustering techniques are used to obtain data that can be used for analysis.

The more commonly used of word segmentation method is based on dictionary matching, based on understanding and based on statistics. The text representation method is diverse, the more common methods are vector space model, probability model, Boolean model and other models. Vector space model is the most commonly used.
Internet public opinion dissemination and evolution methods

For the study of the dissemination and evolution of Internet public opinion, this paper mainly from two aspects: qualitative research and quantitative research. Through sorting out, the qualitative research on dissemination and evolution of Internet public opinion focuses on the hot spot law of public opinion, the evolution characteristics of public opinion, etc. In 2016, scholar Steinfeld [9] argued that network users' comments in complex networks had a very important impact on the dissemination of public opinion information. Huang [10] believed that if we analyze the user's retrieval behavior, we can predict the retrieval steps and goals, and then control the dissemination process of public opinion information. However, scholar Kim [11] believed that social networks can reveal controversial social problems through public opinion dissemination. Through public opinion information diffusion theory, it can reveal the law of public opinion information dissemination.

The quantitative research on the dissemination and evolution of public opinion is mainly based on the establishment of mathematical models to analyze the empirical data. The construction of mathematical model is mainly from the public opinion information itself and the Internet public opinion subject namely the netizen group. The literature [12] supposed the public opinion information quality as the two-order reflection type variable, and built the dual-pathway model of the factors of influencing public opinion dissemination behavior of the audience. Besides, it also combines the research about the key factors of the public opinion dissemination to analyze the mechanism of it. Zhu [13] by simulation methods to investigate the netizen individual behavior in the role of Internet public opinion propagation, focuses on the contents and the netizen individual properties affect the spread of the evolution of public opinion.

Different scholars have conducted quantitative research and qualitative research on the dissemination and evolution of Internet public opinion, whether it is the qualitative research of public opinion law and evolutionary characteristics, or quantitative research of dissemination mode through mathematical model. It is too simple to fully show the distribution and evolution of public opinion. Especially in the context of big data, the growing public opinion information is complicated. It is not enough to conduct research only through qualitative or quantitative methods. It is better to combine quantitative analysis based on qualitative research on the premise of mining big data public opinion information, and conduct the modeling of public opinion dissemination and evolution to show the law of public opinion information dissemination.

Internet public opinion monitoring and early warning methods

The monitoring and early warning of Internet public opinion is to, on the basis of mining the important factors among the emergence, spread, development and disappearance of Internet public opinion, continuously and dynamically measure, monitor and collect the information of these factors. Then evaluating the risk level of the current public opinion and predicting its development tendency using some comprehensive analysis technology such as computer technology and manual work, in order to give a level prediction report timely [14]. For public opinion monitoring and early warning research, its research direction is a single, mainly focused on the following three directions: Internet public opinion monitoring and early warning index
system construction, the public opinion data mining and analysis methods, and the Internet public opinion dissemination and crisis treatment countermeasures. The construction of index system is the basis of monitoring and early warning the risk of Internet public opinion, and whether the index system is scientific has a direct impact on the correctness of the judgment result of Internet public opinion crisis. Therefore, we should ensure that the index system is properly constructed. The data mining method of public opinion has been summarized in the previous section, here we mainly introduce the public opinion index system, and the main methods of index construction shown in table 1.

As can be seen from table 2, Internet public opinion index can be divided into three categories: generic Internet public opinion index, specific event type Internet public opinion index, and specific domain platform network public opinion indicators. In general Internet public opinion index construction, the researchers mainly by public opinion circulation, public opinion factor, public opinion trend as the foundation, combined with social and technical indicators [15], the dynamic factors such as the change rate of release quantity and comment quantity [16], and frequency information such as click, transfer and reply [17] to build public opinion index. Specific event type Internet public opinion index construction, the researchers use the analytic hierarchy process [18], BP neural network method [19] to build public opinion index. In the construction of specific field platform Internet public opinion index, Internet media as public opinion carrier has its own characteristics in the efficiency and influence of information dissemination. Researchers mainly build public opinion index such as micro-blog [20], video website [21] and other platforms. Li [21] built a security monitoring index system for Internet public opinion, which is based on diffusion degree, attention degree, attitude orientation degree and excitation degree as the first level index. This is the first study of the Internet public opinion index system for specific manifestations.

### TABLE 1. PUBLIC OPINION INDEX CONSTRUCTION.

| Index Classification                  | Author | Index Construction Elements                                                                 |
|---------------------------------------|--------|---------------------------------------------------------------------------------------------|
| **Generic Internet public opinion index** | Dai    | public opinion flow index/ public opinion factor index/ public opinion trend index           |
|                                       | Sun    | the change rate of release quantity/ the change rate of comment quantity/ the variable fuzzy model theory |
|                                       | Wang   | the frequency information of click, reply, transfer/ the main dimension of heat, intensity, inclination, growth |
| **Specific event type Internet public opinion index** | Zeng   | police source, police warning and police situation                                           |
|                                       | Zhang  | BP neural network method/ event explosion force/ netizen force/ Internet media influence/ government guidance force |
| **Specific domain platform Internet public opinion index** | Gao    | I-Space framework/ public opinion subject/ public opinion information/ public opinion dissemination/ public opinion audience |
|                                       | Li     | diffusion degree/ attention degree/ attitude orientation degree/ excitation degree           |
Internet public opinion management and control methods

For the Internet public opinion management and control, there are typical management models overseas at present. From the perspective of management and sociology, the content of Internet public opinion can be managed according to the law by improving laws and regulations, and regulating the content of public opinion. Several of administrative supervision means and special actions for restraining cyberspace should be formulated and carried out to manage and lead the behavior pattern of public for surfing the Internet. The self-discipline of network Medias and users should be actively guided and encouraged by the means of building private institution. Network information should be controlled and filtered forcibly by nation and government through technological means [22].

From the perspective of computer and information technology, Internet public opinion management is usually conducted after the monitoring and analysis of the Internet public opinion. By investigating how to use the search engine, natural language processing, social network analysis, machine learning and other information processing technologies, the Internet public opinion and its law can be found and analyzed automatically by the network monitoring and analysis software or system tools. According to the requirements of each department for relevant Internet public opinion management, public opinion bulletin and newspaper will be formed to help decision-making level grasp public relation, provide gist to make the right decision, and manage the Internet public opinion expediently. In response to the public opinion management, this work mainly studies the guiding strategy of public opinion [23], coping method [24] and governance model. Scholar Maccoun [25] thoughted that the combination of network and the dissemination analysis of public opinion information can help government publish national policy. In the opinion of scholar Aklin [26], the results of the Citizens' Satisfaction Survey show that the government has a negative impact on the wrongful interference of public opinion information. Reference [27] thinks that the Internet public opinion information supervision and social affairs have an important influence on the evolution of public opinion based on the results of numerical simulation.

However, the study of Internet public opinion management and control is still in an early stage. The public opinion control solution only can solve the problem of public opinion initially, has not yet reach the level of efficiency and strength at present. Internet public opinion will be managed and controlled effectively by summarizing and researching to provide readers some enlightenment.

Conclusion

This work mainly studies the fields of Internet public opinion analysis method, public opinion spread, and public opinion monitoring and public opinion management recent years. Though, Internet public opinion analysis made a lot of results still have many problems need to be resolved.

Internet public opinion analysis method is complicated, such as the Internet public opinion topic text tendency analysis, public opinion heat analysis, and emotional tendency analysis are mature. However, the exhibition of the public opinion analysis results should be improved. Advantages of information can be converted to decision-making advantages by the visualization technology, and a complete Internet public
opinion visual system of emergencies can be established to provide technical support
to the decision makers in early warning and disposal of emergencies.

The internet vocabulary produced in emergencies can form massive Internet public
opinion information easily. However, how to identify the new vocabulary on Internet,
determine the emotional tendencies of the new Internet words, and convert
information advantage to decision advantage timely and accurately remain to be
studied.

INTERNET NEW WORD AND ITS MAIN TECHNIQUES

In recent years, with the rapid development of social networks, Internet new words
always appear in the social hot events, therefore, the outbreak of Internet public
opinion is often associated with the rapid update of new Internet words. Obviously,
Internet new words have become an important part of research about Internet public
opinion information. Predecessors research on Internet new words have done a lot of
work from different fields and perspectives. This paper reviews the research
techniques of Internet new words from two perspectives: natural linguistics and
computer information technology.

Natural linguistics research perspective of Internet new words

From the perspective of natural linguistics, researchers have made corresponding
academic research on the word formation pattern and word meaning evolution of
Internet new words [28-30], the reasons of emergence and prevalence of Internet new
words [31-33]. Table 2 summarizes the main research contents and research methods.

According to the study, concerning word meaning evolution of Internet new
words, Jiang [30] expatiated the mechanism of word meaning evolution about Internet
new words from the perspective of analogy mechanism and emotion analysis. In 2016,
Zhou [31] used the qualitative comparative analysis (QCA) method to analyze the
micro mechanism of network catchwords with high heat transmission, and extracted
the six micro factors that influence network catchwords. Even focusing on the study of
the transmission mechanism of Internet new words, it is often more macro speculation
than the micro-discussion. For example, some studies have analyzed the process of
constructing and disseminating of network catchwords from specific perspectives such
as memetics and the concept of social psychology, or to explore the generation
mechanism and normative strategy of catchwords in the context of new media [34].

| Main research contents            | Main technologies                          | Author       |
|-----------------------------------|--------------------------------------------|--------------|
| Word formation pattern/Word meaning evolution | Memetics                                   | Zhou [28]    |
|                                   | Cognitive motivation                        | Zhang [29]   |
|                                   | Analogy mechanism/Emotion analysis method   | Jiang [30]   |
| Dissemination evolution/Prevalent reason | Qualitative comparative analysis (QCA)     | Zhou [31]    |
|                                   | Cognitive motivation analysis               | Yu [32]      |
|                                   | Quantitative analysis                       | Ma [33]      |
TABLE 3. INTERNET NEW WORDS RECOGNITION METHOD.

| Research Contents                                                      | Research Methods       | Author     | Main Features                                                                 |
|----------------------------------------------------------------------|------------------------|------------|------------------------------------------------------------------------------|
| the extraction of candidate new words and the filtering of garbage strings | Statistical -based methods | Su[35]/Sun[36]/Sarna[37] | Entry composition/Feature information/ Mutual information                     |
| the combination of statistical -based and rule-based methods          | Rule-based methods      | Huang[38]/Zheng[39] | Template construction/Word-formation principles/Semantic information/ POS information |
| the Part-of-Speech(POS) Tagging Method of Internet new words           | the combination of statistical -based and rule-based methods | Huo[40]/Zhou[41]/Ya o[42] | Combination of lexical features and statistical features                      |
|                                                                      |                         | Xu[46]/Goh[447] | Dictionaries/Rules/Statistical data                                            |

Computer information technology perspective of Internet new words

In addition to the study of the word formation, meaning evolution, transmission evolution and popularity of Internet new words from the perspective of natural linguistics, scholars have also used computer technologies to analyze Internet new words. Among them, the most researched and widely used is the network new word recognition technology.

In general, the research on internet new word recognition is mainly carried out in two aspects: (a) the extraction of candidate new words and the filtering of garbage strings. (b) the Part-of-Speech(POS) Tagging Method of Internet new words. At present, scholars are more research on the first aspect, and further research is needed on the part-of-speech tagging of new words. There are three kinds of research methods for the recognition of Internet new words: (a) Methods based on statistics; (b) Methods based on rules; (c) Methods based on the combination of statistics and rules. The induction is shown in table 3.

The extraction of candidate new words and the filtering of garbage strings.

STATISTICAL-BASED METHODS.

Concerning the disadvantage of statistics method and micro-blog text features, Su[35] proposed an improved method to extract micro-blog text neologisms, by means of the method of increasing and decreasing weight to improve the entropy filter useless strings, and ultimately improve the recognition performance. Sun[36] proposed a new method of Internet Chinese word discovery based on information dissemination, which solved the problem of short life cycle of existing method identification results and could not be used again soon. This method combines the popularity and
persistence of new words, and defines three statistical information, such as user coverage, topic coverage rate and new word life cycle. The experimental results show that when the three statistics are combined together, the accuracy is improved by 16%. Each statistic considered by this method can improve the accuracy. Considering the combination of the three statistics, not only a statistic is considered, we can get a more accurate estimation rate. In 2016, Sarna [37] proposed a probabilistic method for automatically extracting new words from social media. This method implements new word extraction and applies these new words to various categories of monitoring.

RULE-BASED METHODS.

Such as Huang [38] according to the Part of Speech, the rules are divided into positive rules and negative rules to identify new words. But, Zheng [39] founded new words totally according to the rules. They established the rules library according to the Chinese word-formation method, and the new words were determined by calling "mutex string" filtering rules and word formation rules. The automatic recognition experiment system of Internet new words is designed and implemented. However, with the emergence of new rules, it is difficult to sum up all the rules of a new word. With the passage of time, the system based on certain time rules will soon be out of date.

AN ASSOCIATIVE METHODS BASED ON STATISTICS AND RULES.

The advantages of rule-based methods are high accuracy and pertinence, but it is difficult to write and maintain rules by hand. The advantages of statistical-based methods are flexibility, adaptability and portability, but large-scale corpus is needed for model training, and because of less language knowledge, there are sparse data and low accuracy problems. At present, most researchers use the associative methods based on statistics and rules in order to play their respective advantages, improve the effect of new word discovery [40-42]. For example, Huo [40] putted forward an iterative context entropy algorithm which introduces the relevance information of words, and introduces the method combining lexical features and statistical features, but this method relies on the word segmentation system, and it choice word frequency as statistical features, it is easy to ignore low-frequency words.

Other methods.

In addition to using the above three methods to identify new words, researchers have begun using big data technology to solve new word recognition problems [44, 45]. In 2016, Xiao [45] proposed a new word filtering algorithm for micro-blog content, and proposed improved mutual information and relative adjacency entropy method. In the case of recognition of new word in massive corpus produced from huge amounts of multiple-feature data, this method utilizes the Map/Reduce parallel computing model to extract three features which are the enhancement of mutual information, the relative adjacency entropy and the document frequency of the context, then it constructs feature vectors of candidate words according to these features. Above feature vectors will be used to train a SVM model with training corpus in order to identify new words. The method shows that parallelization reduces the whole time of new words recognition.
The Part-of-Speech (POS) Tagging Method of Internet new words.

Part-of-Speech (POS) tagging of new words is mainly based on dictionary, rules or statistical data to determine the POS of the marked words. Different from common POS tagging, the main difficulty of new words POS tagging is that there is no support for dictionary and data statistics. At present, the research methods of new word POS tagging are mainly based on the combination of statistics and statistics and rules.

Xu [46] extracted the positive and negative samples from training corpus which was handled by segmentation and POS Tagging according to the dictionary, and then combine the constraint conditions of the support vector machine classification to recognize new words. This method can improve the accuracy of word recognition and recall. Goh [47] used HMM to mark the words, then used the result of the tag to train the SVM model to get the character based annotation, and used the character sequence of the lexical blocks to detect the unknown words. And he combined with the detection of unknown words and the initial segmentation results, the final word segmentation results are obtained.

INTERNET NEW WORDS AND INTERNET PUBLIC OPINION MANAGEMENT

The challenge of Internet new words to the control and management of Internet public opinion

THE RAPID EMERGENCE OF A LARGE NUMBER OF NEW WORDS ON THE INTERNET AND THE IDENTIFICATION OF INTERNET NEW WORDS IS A BIG CHALLENGE.

One of the basic tasks in network public opinion management and control is to quickly identify new topics, hot topics, emergencies and the like from a large amount of network corpus data. However, these new topics, hot topics, keywords of emergencies are often some new words. Usually these new words are not included in the machine dictionary, but usually is one of the important basis for network public opinion monitoring system clustering. If these words are not well identified, they often lead to poor clustering results and excessive result sets. Therefore, the identification of new words on the Internet is a major challenge of network public opinion management.

IN THE CONTEXT OF BIG DATA, THE RAPID INNOVATION OF NETIZENS LEADS TO THE DIFFICULTY OF PUBLIC OPINION FINDING.

Netizens are always able to create the latest words in the quickest and the shortest time. These words are often not in regular grammatical structure and lexical collocations, so as to avoid the supervision of public opinion, so that Internet users can speak quickly and freely. Netizens utilize unique Internet new words involved in the discussion of social hot events to reach a whirlpool. The rapid formation of public opinion on the network results in enormous pressure of public opinion, making it difficult for regulatory authorities to identify, which can delay the effective control of the best timing.
INTERNET NEW WORDS TECHNOLOGY TO SOLVE THE PROBLEM OF
INTERNET PUBLIC OPINION CONTROL

The rapid emergence of Internet new words, the analysis of the Internet public opinion, and the difficulty of trying to find a way to do that, is to be able to identify and judge the new words from the following three aspects, so as to provide support for public opinion management and control.

By using big data technology, through distributed platform (Hadoop, Map Reduce, etc.), the network public opinion information that contains Internet new words generated by netizen has been processed, thus improving the detection speed of Internet new words and accurately judge the dynamics of public opinion.

At present, the automatic detection technology of sensitive words is based on the filtering technology of sensitive dictionary information, which can be detected in time, such as sensitive words and other harmful information. However, for the Internet new words of netizen innovation, there is no perfect dictionary of Internet new words, and the sensitive information contained in it cannot be accurately judged. We optimize the Internet new words recognition algorithm constantly, and then build an internet new words dictionary as the basis of analyzing the public opinion.

The emergence and evolution of new Internet words can be analyzed quantitatively, to find out the emotional meaning contained in new Internet words and to visualize their analysis results, so as to reflect the evolving trend of public opinion and to better control the public opinion.

SUMMARY AND PROSPECT

From the analysis above, after many years of efforts, Internet public opinion and Internet new words research has made great progress, but there are still many problems and challenges:

Research on Internet public opinion.

Large-scale data information interferes with the control of public opinion. Lack of systematic public opinion monitoring and management system.

Research on Internet new words.

Research on the evolution of Internet new words is only described qualitatively by the written language, not to choose the right influence factors and to make a quantitative analysis of its evolution through the data is more scientific and accurate.

There are some limitations in the recognition of Internet new words, and the recognition effect needs to be improved. From the research results of Internet new word recognition, the accuracy rate, recall rate and F value are still low. There are many problems in the new word recognition. For example, the definition of new words is not unified, the subjectivity of new words is judged artificially, the complexity of garbage string filtering, etc.
ACKNOWLEDGMENTS

This paper was sponsored by the Management and Early Warning Mechanism of Group Events in Cyberspace of The National Social Science Fund of China (17XFX013). The valuable comments and helpful suggestions in the referees are appreciated.

REFERENCES

1. Lv P. A Study of New Words and Public Opinions on the Internet[J]. Journal of Xihua University, 2010, 29(1):102-106.
2. Chen Z. An Improved Text Categorisation Algorithm Based on Centroid[J]. Computer Applications & Software, 2013, 30(1):43-36.
3. Wang Z, Tong V J C, Chin H C. Enhancing Machine-Learning Methods for Sentiment Classification of Web Data[C]// Asia Information Retrieval Symposium. Springer International Publishing, 2014:394-405.
4. Cao X, Zhang X, Liu L. Research on Internet Public Opinion Heat based on the Response Level of Emergencies[J]. Chinese Journal of Management Science, 2014, 22(3):82-89.
5. Lv X, Chen D. Geospatial Grid Display of Internet Public Opinion Information about Natural Disasters[C]// International Conference on Network and Information Systems for Computers. IEEE, 2015:399-403.
6. Chen F, Huang J. Network Public Opinion Propagation Model in the Emergency[J]. Information Science, 2015, 33(12):8-12.
7. Jin S, Xu Y. The Design of Public Opinion Analysis System Based on Topic Events[J]. Advanced Materials Research, 2014, 926-930:2233-2236.
8. Sun H, Wei Y. A Note on the PageRank Algorithm[J]Applied Mathematics and Computation, 2006, 79(2):799 -806.
9. Steinfeld N, Samuel-Azran T, Lev-On A. User comments and public opinion: Findings from an eye-tracking experiment[J]. Computers in Human Behavior, 2016, 61:63-72.
10. Huang B, Yu G. Research and application of public opinion retrieval based on user behavior modeling[J]. Neurocomputing, 2015, 167: 596-603.
11. Kim K, Baek Y M, Kim N. Online news diffusion dynamics and public opinion formation: A case study of the controversy over judges’ personal opinion expression on SNS in Korea[J]. The Social Science Journal, 2015, 52(2): 205-216.
12. Zeng Q, Cheng X, Zhou X, et al. Study of Public Opinion Dissemination Mechanism on Social Networks Based on Dual-pathway Model[J]. Information Science, 2017(6):29-33.
13. Zhu Y, Zhang C. An Application of Impact Model to Simulating the Evolution and Dissemination of Network Public Opinion[J]. Journal of Intelligence, 2015(2):28-36.
14. Wu S, Li S. A Study on Early Warning Mechanism of Internet Public Opinion[J]. Journal of Chinese Peoples Public Security University, 2008,14(3):38-42.
15. Dai Y. Research into Information Mining and Evaluation Index System Based on the Security of Public Opinion on the Internet[J]. Information Studies Theory & Application, 2008,31(6):75-78.
16. Sun L, Zhou J, Xu H, et al. The Concept of Network Public Opinion Crisis Analysis and Index Set[J]. Journal of Modern Information, 2014, 34 (11):25-28.
17. Wang Q. On the Construction of Internet Public Opinion Index System for Monitoring and Early Warning[J]. Library & Information Service, 2011,55(8):54-57.
18. Zeng X. Construction of the Network Opinion Emergencies Early Warning Index System[J]. Information Studies Theory & Application, 2010, 33(1):77-80.
19. Zhang Y, Qi J, Fang B. The Indicator System Based on BP Neural Network Model for Network-mediated Public Opinion on Unexpected Emergency[J]. China Communications, 2011, 8(2):42-51.
20. Gao C. Research on Public Opinion Monitoring Index-system in Micro-blogging[J]. Journal of Intelligence,2011,30(9):66-70.
21. Li J. Study on the Safety Monitoring Index System and Monitoring Method of Network Public Opinion[J]. Journal of Xi’an University of Finance and Economics,2013,26(3):116-122.
22. Chinese cyberspace Research Institute. Overseas Internet bad information supervision [M]. Beijing: Law Press, 2016.
23. Zhang Y. Causes and control strategies to the network public opinion of emergencies: Psychological analysis of their subjects [J]. Journal of Intelligence, 2012, 31(4): 54–57.
24. Wang G, Feng W, Wang Y. Research on response to public opinion based on the classification of network public opinion [J]. Journal of Intelligence, 2013, 32(5): 1–4.
25. Maccoun R J, Martin K D. Legal Issues: Public Opinion[J]. International Encyclopedia of the Social & Behavioral Sciences, 2015:747-752.
26. Aklin M, Bayer P, Harish S P, et al. The political economy of energy access: Survey evidence from India on state intervention and public opinion ☆ [J]. Energy Research & Social Science, 2015, 10:250-258.
27. Liu D, Wang W, Li H. Evolutionary Mechanism and Information Supervision of Public Opinions in Internet Emergency ☆ [J]. Procedia Computer Science, 2013, 17:973-980.
28. Zhou Q, Tan D. The Formation Features of Cyber Catchwords from the Memetic Perspective[J]. Journal of Hunan University of Science & Technology, 2013,16(5):139-141.
29. Zhang Y. Variations of Meaning in Cyber language[J]. Applied Linguistics.2014,No.92(4):108-115.
30. Jiang T. On the Role of Analogy Mechanism in Meaning Evolution of Chinese New Words[J]. Chinese Studies, 2015, 04(3):89-94.
31. Zhou J, Wang M. How Internet Buzzwords Spread in China: Crisp-set Qualitative Comparative(csQCA) Analysis of 12 Public Events[J]. Chinese Journal of Journalism & Communication, 2016(4):26-46.
32. Yu Z. The Generation of New Words Prevailing on the Internet and Its Revelation to the Innovation Environment Cultivation[J]. Energy Procedia, 2012, 17(Part A):717-722.
33. Ma P. A Study on Popular Language Known and Analysis during 2006-2011[D]. Guangzhou University, 2012.
34. Li L. Generation and Transmission Mechanism of the Popular Language in the Context of New Media and its Standard Strategy[J]. Xue Hai,2014(6):85-90.
35. Su Q, Liu B. Chinese new word extraction from MicroBlog data[C]// International Conference on Machine Learning and Cybernetics. IEEE, 2013:1874-1879.
36. Sun L, Zhou Y, Guan X. A method of discovering new chinese words from Internet based on information propagation[J]. Hsi-An Chiao Tung Ta Hsueh/Journal of Xi'an Jiaotong University, 2015.
37. Sarna G, Bhatia M P S. A probabilistic approach to automatically extract new words from social media[C]// ieee/acm International Conference on Advances in Social Networks Analysis and Mining. ACM, 2016:719-725.
38. Huang X, Li R. Discovery method of new words in blog contents[J]. Modern Electronics Technique, 2013, 36(2):144-146.
39. Zheng J, Li W. A Study on Automatic Identification for Internet New Words According to Word-Building Rule[J]. Journal of Shanxi University, 2002, 25(2):115-119.
40. Huo S, Zhang M, Liu Y, et al. New Words Discovery in Microblog Content[J]. Pattern Recognition & Artificial Intelligence, 2014, 27(2):141-145.
41. Zhou Q, Chen Y. New Words Recognition Algorithm and Application Based on Micro-Blog Hot[C]// International Conference on Measuring Technology & Mechatronics Automation. IEEE, 2015:698-700.
42. Yao R, Xu G, Song J. Micro-blog new word discovery method based on improved mutual information and branch entropy[J]. Journal of Computer Applications, 2016, 36(10): 2772-2776.
43. Zheng J, Li W. A Study on Automatic Identification for Internet New Words According to Word-Building Rule[J]. Journal of Shanxi University, 2002, 25(2):115-119.
44. Nugumanova A, Novosselov A, Baiburin Y, et al. Automatic keywords extraction from the domain texts: Implementation of the algorithm based on the MapReduce model[C]// International Conference on Current Trends in Information Technology. IEEE, 2013:186-189.
45. Xiao C, Gan J, Wen B, et al. Research on Micro-blog New Word Recognition Based on MapReduce[C]// Bio-Inspired Computing - Theories and Applications. Springer, Singapore, 2016:379-387.
46. Xu Y, Gu H. New Word Recognition Based on Support Vector Machines and Constraints[C]// International Conference on Information Science and Control Engineering. IEEE, 2015:341-344.
47. Goh C L, Asahara M, Matsumoto Y. Machine Learning-based Methods to Chinese Unknown Word Detection and POS Tag Guessing [J]. Journal of Chinese Language and Computing, 2006, 16(4):185-206.