

## An Early Warning System based on Data Mining of Public Opinions in Tourism Social Networks

Jie Su

Guizhou University of Finance and Economics  
Guizhou, Guiyang, China

**Abstract—** Awaiting revised text

**Keywords -** data mining; public opinion; early warning system; tourism emergency

### I. INTRODUCTION

Network opinion is the general situation of attitudes, opinions and viewpoints about social hot events especially crisis by network opinion subject including masses and media employing Internet media, and network opinion has a significant impact on image of the organization, government decision-making and social stability. Researching on the network opinion of tourism crisis is important to the supervisor grasps the changes of network opinion timely, takes effective measures to deal with network opinion crisis, improves the management level of tourism crisis, provides better services for tourists, and maintaining the image of tourism enterprises and destinations. [1]

Using literature research, case study method, content analysis and the mathematical statistics analysis method including frequency statistics, cross tabs analysis, correlation analysis, regression analysis and optimal scaling analysis, this dissertation conducts definition and empirical research of compositional factors and developing mechanism for the network opinion of tourism crisis, and draws the main conclusions. [2]

In recent years, research on online information early warning of unconventional emergencies has become a hot point "scenario-response" turns out to be in the areas of emergency management as well as the basic paradigm. [3] Under this background, this paper tries to theoretically establish a theory model of online information early warning of unconventional emergencies; researching spreading law of the online information based on the consideration of derivative information with the complex network theory to structure key information mining model; based on both models, researching on early warning model and strategies to a reference for responding unconventional emergencies. [4]

With the propelling of information revolution in the end of the 20th century, modern network technique has been widely applied and popularized, and online public media, as the fourth kind of communication media for public opinion after newspaper, broadcast and television, has gradually received attention and entered into the mainstream society's viewpoint for the role of new channel for the collection of public information and distributing centre of public opinion.

[5] Especially in the recent years, with the progressive increase of the quantity of Chinese netizens year by year, more and more netizens use Internet to express their interest appeal, abreast their emotion, participate public affairs, and all kinds of social contradictions producing in the social transformation period are always focusing on the online public platform to form powerful public opinion field at the first time. Meanwhile, the strong characteristics of openness, rapid spread and interaction for Internet make some simple and controllable public opinion incidents in the traditional communication channels and opinion environment turn to unexpected group events and trigger social public crisis, which bring huge pressure and challenge to public administration. Time of events are shown in Figure 1.

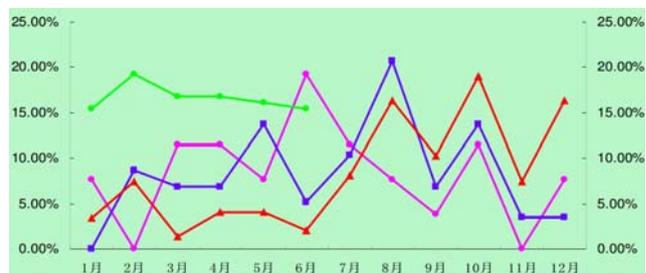


Figure 1. Time of events

### II. METHODOLOGY

As an important function for public administration, online public opinion crisis pre-warning has been paid extensive attention by the public, and the relevant mechanism construction and technical system development has proceeded one after another, governments at all levels pay more attention to the monitoring and pre-warning of public opinion. However, the analysis and judging of online public opinion, as the key link of crisis pre-warning processes, whose relevant research and practice falls behind the practical demand for public crisis pre-warning under web2.0 environment, and can't provide automatic, [6] intelligent and real-time crisis pre-warning support to current dynamic and interactive online public opinion from the aspects of analysis methods and technical support. In this

context, using current technical means and disciplinary methods to innovate methods and technology of online public opinion analysis to provide technical guarantee for crisis pre-warning has important practical significance for maintaining social stability and propelling social transformation. The number of tourism events is shown in Figure 2.

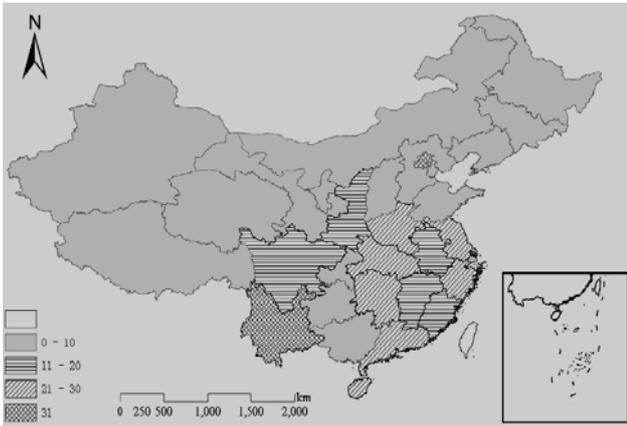


Figure 2. The number of tourism events

Deep research on theoretical key online information early warning model means: first, defining and discriminating the scenario both in decision making and in unconventional emergencies and their semantics as well as researching on the intension and manifestation of unconventional emergencies online information originating from viewpoints posted by subjects such as netizens and media on forums or other network platform post-emergency and in its development, which includes solitary original information and massive derivative information; secondly, researching on the treatment of online information and key online information on the basis of search the relationship between scenario and online information; at last, based on both researching, making theoretical model architecture of unconventional emergencies online information early warning strategies in three dimensions: the overall architecture of models, the one of spreading law of scenario based unconventional emergencies online information, and the one of information based. [7] The research indicates that scenario is the combination of previous, present and future situations, which is partly similar to original and derivative information of the online one. The overall theoretical model architecture would be reasonable from scenario and online information, the spreading and evolution law of online information and the recognition and mining of key information, to key information early warning. Individuals and institutions users are shown in Figure 3.

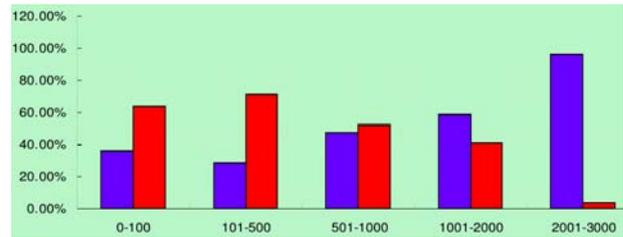


Figure 3. Individuals and institutions users

Deep research on the spreading and evolution of unconventional emergency online information means: first, during its process, the original trajectory varies with the updating and superposition of derivative information. [8] As a result, this paper highlights considering the impact of derivative information to the original one, constructing derivative information based model of the spreading and evolution of unconventional emergency online information, namely in considering the derivative information subject to uniform, normal, chi-square and other continuous probability distribution to establish the computational model and to conduct sampling calculation and simulation; secondly, it is the key point in describing and determining the spreading and evolution law of unconventional emergency online information to effectively recognize key information. Therefore, the fundamental trajectory can be seen to a certain extent by using complex network model to construct unconventional emergency key information mining model with mutual impact factor between derivative and online information. [9] The results suggest that the computational model of the spreading and evolution of unconventional emergency online information model with considering derivative information can effectively recognize the key nodes (key online information); new model of key information mining with considering mutual factors can effectively recognize the key nodes (key online information) in the law of it; to provide basis for early warning.

### III. RESULTS AND DISCUSSION

There are large numbers of network opinion objects about inbound and outbound tourism, junket tour, aviation, tourism planning & development and tourism marketing, etc. [10] There are large numbers of network opinion objects about security issues, resources & environmental protection, government policy management dispute, extravagance & corruption, violation of consumer rights, criminal, national quality, and ethics, etc. (Eq. (1))

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-T)^2}{2}}, 0 \leq x \leq T, T \rightarrow \infty \quad (1)$$

In temporal distribution, network opinion objects are influenced significantly by the development of network new media, public opinion environment, holidays, tourist flows, climate and weather, etc. In spatial distribution, there are large numbers of network opinion objects involved East China, Central China, South China and Southwest China. [11-12] If a region has the higher level of tourism development and the higher network concern degree, it may

involve more network opinion objects. Outside China, East China, Central China, and North China have close relations to the specific content characteristics of network opinion objects. The network opinion objects about railways, consumer prices, leisure & entertainment, other irregularities, junket tour, and extravagance & corruption aren't influenced by spatial factors. (Eq. (2))

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(t-\frac{T}{2})^2}{2\sigma^2}} dt, 0 \leq x \leq T, T \rightarrow \infty \quad (2)$$

About network opinion ontology, firstly, network opinion is influenced significantly by the media agenda-setting. Internet users' discussion about the event is based on the content of media coverage. On the basis of the knowledge of the facts, Internet users have personalized and in-depth web content production such as attitude & assessment, expression of intention, considering other issues and questioning the fundamental reason. Secondly, expression of the idea of not to want to travel there, "vote with their feet", or "boycott" are the main way that Internet users express negative emotions. Thirdly, the network opinion's phenomenon of considering other issues and questioning the fundamental reason is obvious. Through network opinion, tourism events are developed into social events, the negative impact of the event is expanded to the entire tourist destination, and the meaning of individual cases is extended to the level of country and society. Lastly, the network opinion's phenomena of group polarization and network violence are also obvious. Extreme manifestation of group polarization is endless language violence, person pork searches, network spoof and network rumors & slander.

About occurrence mechanism of network opinion, the influencing factors of causing the network opinion of tourism crisis include event involved subject, tourism characteristics, society characteristics, causes events, demands of subject, news value, place of occurrence, and time of occurrence. High frequency event theme is shown in Figure 4.

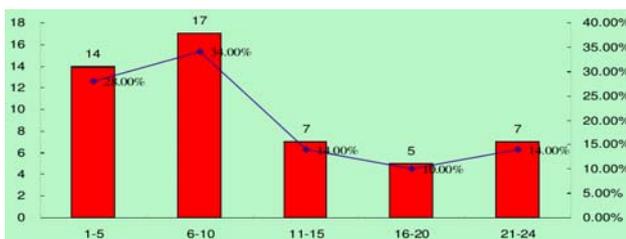


Figure 4. High frequency event theme

About the communication effect of network opinion, approach of posting, content of posts, and number of related posts have strong explanatory power for number of posts forwarded. Approach of posting and location have strong explanatory power for number of posts commented.

About evolution mechanism, the event's evolution of network opinion can be divided into 7 stages including gestation, appear, outbreak, climax, fall, repeated and long tail. Factors which influence evolution of network opinion

include social mentality, event effect, the official words and deeds, media interviews and reports, Internet users' behavior, external stimulus, and response to network opinion. These factors play a boost or counter effect for evolution of network opinion. (Eq. (3))

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\frac{T}{2})^2}{2\sigma^2}}, 0 \leq x \leq T, T \rightarrow \infty \quad (3)$$

In this part, the paper starts from analyzing the connotation of crisis and public crisis, discussing the characteristics and development tendency of current public crisis. Meanwhile, the paper dissects the material sources, concept connotation, characteristics representation, content expression, spread and evolution rule of online public opinion, and takes Deng event as an example to analyze the spread and evolution rule of online public opinion. The traditional media and network media is shown in Figure 5.

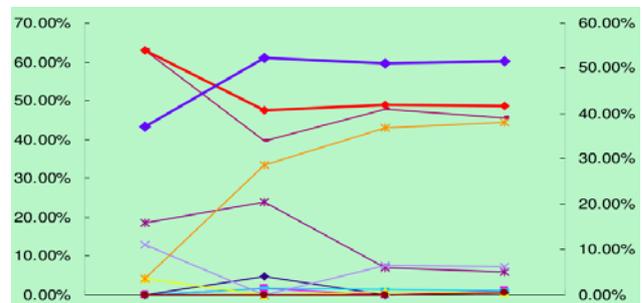


Figure 5. The traditional media and network media

The paper starts from analyzing the rigorous status of public crisis pre-warning for online public opinion, demonstrating the necessity and realistic demand. Meanwhile, it analyzes the processes and functions of public crisis pre-warning, constructing the analysis framework of online public opinion oriented public crisis pre-warning, which constituting information sources layer, technical support layer, analysis method layer, analysis pattern layer, analysis and application layer, covering all the aspects of online public opinion analysis from technical aspect to methodological aspect. In the respective of online public opinion analysis processes, the paper learns from the traditional information analysis processes and puts forward the analysis processes including theme planning, information collecting, information preprocessing, information analysis and crisis pre-warning and processing, and expounds the analysis demands and principles.

The paper mainly discusses the methodology of online public opinion analysis. First, the paper briefly introduces various important methods using in online public opinion analysis and constructs the method system of online public opinion analysis. Then the paper demonstrates content analysis method, web metrics method, and intelligent analysis method relevant to online public opinion analysis in detail. Content analysis and web metrics are the general scientific methods using in social sciences research, and still suitable for analyzing online public opinion. As to intelligent

analysis method, the paper puts forward two different kinds of intelligent analysis methods based on different technologies, one is based on web mining, and another is based on semantic analysis. The analysis method based on web mining is the new application of data mining in online public information analysis, focusing on discovering online public opinion law from the content of web texts, the structure of web links and the usage records of website, including such methods as web text mining, topic clustering, link mining and social online public analysis, and it has become the mainstream method for opinion analysis. The analysis method based on semantic focuses on discovering opinion law from the semantic layer assisted by analyzing the latent semantic structure in web text or ontology and semantic dictionary, including latent semantic analysts and semantic analysts based on external semantic knowledge, etc. Network awareness is shown in Figure 6.

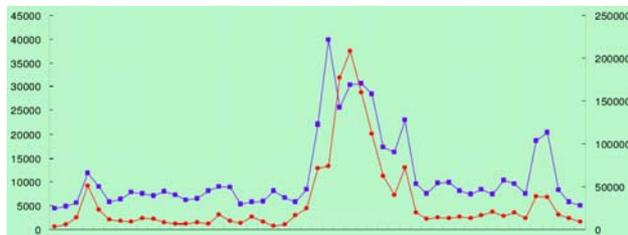


Figure 6. Network awareness

The paper takes four key opinion elements acting in public crisis pre-warning as analysis object, discussing the main analysis patterns from the aspects of topic discovering, event recognizing, individual netizen analysis and group behaviors analysis. As to the analysis of topic discovering pattern, the paper starts from the processes of topic discovering, discussing the method of topic modeling and developing trend. Then the paper emphasizes on hot topic and sensitive topic, the two key factors of triggering public crisis, and discusses the method of opinion discovering and impact assessing. As to the analysis of event pattern, the paper discusses the method of extracting the elements of public crisis event. Based on event modeling, the paper analyzes the processes of emergent event and the assessment of its influence degree and the classification method. As to the analysis pattern of individual netizen, the paper starts from analyzing the constituent and characteristics of netizen in online public opinion crisis, probing how to use text tendency analysis method, link analysis method to analyze the emotion and influence for individual netizen, then discusses the recognizing method for the opinion leader, a special opinion subject in online public opinion crisis. As to the analysis of the behavior characteristics and behavior pattern for Internet population, the paper takes online public community as the main opinion object and analyzes the evolution law and characteristics of the online public community. Meanwhile, the paper discusses the method for recognizing group behavior and predicting their development trend by social online public analysis and trend analysis.

In this part, the paper applies the theory of opinion analysis to the practice and introduces an analysis example

based on TRS opinion management system. This system is developed by TRS Company, and can finish the work of online public opinion analysis and public crisis pre-warning. Then the paper demonstrates the whole processes for TRS OM to carry out information collection, information pre-processing, intelligent mining and analyzing, pre-warning service providing in detail.

Deep research on early warning of key information based unconventional emergency online information means: first, starting form key information, considering nonlinear of the trajectory, imperfection, irregularity of information, and index hardly quantified, and other characteristics, and combing through literature and research, this paper design scientific early warning index system practical and with comprehensive application of multi-level fuzzy evaluation And Trend analysis basedon the DEMATEL method to construct generalized qualitative and quantitative emergency media and early warning model; secondly, taking into account governments, netizens other Agents, this diagram and flowchart, and paper establishes Agent behavior models, causality through simulation, comes up with unconventional emergency online information early warning strategies and guidance, which is of a certain significance.

#### IV. CONCLUSIONS

The originality of this paper consists mainly in two aspects: 1) It proposes the basic compositions of public opinion of tourism in social networks. Public opinion expressed in social networks is the event object. Network opinion subjects are the individuals and organizations which raise issues and express opinion. Network opinion ontology is the information content within the social network. Network opinion media are the technologies and media channels of communicating the information. (2) This project sums up the laws of network opinion generation, development and changes into the mechanism of occurrence, communication and evolution, and puts forward the corresponding content and research propositions to conduct empirical studies.

#### REFERENCES

- [1] Elisabeth Noelle-Neumann. "Turbulences in the Climate of Opinion: Methodological Applications of the Spiral of Silence Theory". *The public opinion quarterly*, vol.41, No.2, pp. 143-158, 2014.
- [2] Knecht T., Weatherford M. S. "Public Opinion and Foreign Policy: The Stages of Presidential Decision Making". *International Studies Quarterly*, vol.50, No.3, pp. 705-727, 2006.
- [3] Campbell J. E. "The fundamentals in US presidential elections: Public opinion, the economy and incumbency in the 2004 presidential election". *Journal of Elections, Public Opinion and Parties*, vol.15, No.1, pp. 73-83, 2015.
- [4] Terhanian G. "Changing Times, Changing Modes: The Future of Public Opinion Polling?". *Journal of Elections, Public Opinion & Parties*, vol.18, No.4, pp. 331-342, 2012.
- [5] Herbst S. "Public Opinion Infrastructures: Meanings, Measures, Media". *Political Communication*, vol.18, No.4, pp. 451-464, 2011.
- [6] Galam S. "Minority Opinion Spreading in Random Geometry". *The European Physical Journal B*, vol.25, No.4, pp. 403-406, 2012.
- [7] Serinoia T. K. "Setting the agenda:the production of opinion at the Sunday Times". *Social Dynamics: A journal of African studies*, vol.36, No.1, pp. 99-111, 2013.

- [8] Habela P. D. "Following the Opinion Leaders? The Dynamics of Influence among Media Opinion, the Public, and Politicians". *Political Communication*, vol.29, No.3, pp. 257-277, 2012.
- [9] Wojdyskia B. W., Riffeb D. "What Kind of Media, and When? Public Opinion about Press Coverage of Politicians' Private Lives". *Journal of Mass Media Ethics*, vol.26, No.3, pp. 206-223, 2011.
- [10] Chena H. T. "Multiple issue publics in the high-choice media environment: media use, online activity, and political knowledge". *Asian Journal of Communication*, vol.22, No.6, pp. 621-641, 2012.