Multi-Dimensional Analysis and Visualization of Social Attributes in Weibo Public Opinion: A Case Study of a Vaccine Incident Postprint

Authors: Mou Dongmei, Shao Qi, Han Nannan, Wang Ping, Jin Shan, Jin Chunyan

Date: 2023-04-01T16:15:49+00:00

Abstract

[Purpose/Significance] This study integrates sociological perspectives with informatics methodologies to align and fuse the social attributes and externalized manifestations of Weibo public opinion, thereby offering a novel framework for interpreting the social issues underlying public opinion phenomena. [Method/Process] Following a review of current research on public opinion, we construct a model of social attributes and externalized manifestations of Weibo public opinion. Centered on public opinion, the model elucidates the internal logic among three social attributes—crowd, content, and emotion—and three externalized manifestations: opinion leaders, events, and sentiment. An empirical study and visualization are conducted using a vaccine incident as a case example. [Results/Conclusion] The empirical findings validate the effectiveness and operability of the model, demonstrating that from a sociological perspective, leveraging the advantages of quantitative and visual research methods in informatics can comprehensively interpret the externalized manifestations of different social attributes of public opinion events, thereby facilitating deep exploration of the essential issues underlying such events.

Full Text

Preamble

Research on Multi-Dimensional Social Attribute Analysis and Visualization of Weibo Public Opinion—Taking a Vaccine Incident as an Example

Authors: Mou Dongmei[1],[2], Shao Qi[2], Han Nannan[3], Wang Ping[2], Jin Shan[2], Jin Chunyan[2]
[1]The First Hospital of Jilin University, Changchun 130021

[2]School of Public Health, Jilin University, Changchun 130021
[3]School of Information Management, Wuhan University, Wuhan 430072

Abstract: [Purpose/Significance] By integrating sociological perspectives with information science methodologies, this study aligns the social attributes of Weibo public opinion with their external manifestations, offering a novel framework for interpreting the societal issues underlying public opinion events. [Method/Process] After reviewing current research on public opinion, we construct a model linking social attributes to external manifestations of Weibo public opinion. Centered on public opinion itself, the model explicates the internal logic among three social attributes (crowd, content, sentiment) and three external manifestations (opinion leaders, events, emotions). We demonstrate this through an empirical case study of a vaccine incident with visual presentations. [Result/Conclusion] The empirical findings validate the effectiveness and operability of our model, demonstrating that a sociologically-grounded approach combined with quantitative and visualization methods from information science can comprehensively interpret the external manifestations of different social attributes in public opinion events and uncover their essential nature.

Keywords: online public opinion; visual analysis; knowledge graph
Classification Number: G250
DOI: 10.13266/j.issn.0252-3116.2020.03.012

The internet has transformed the media landscape for social public opinion, fundamentally changing how people acquire, disseminate, and discuss information. As public opinion increasingly spreads through digital channels, Weibo has emerged as a primary social networking platform and a critical pathway for the propagation of online public opinion hotspots. The online society extends from the real world, and online public opinion reflects real-world problems, serving as a mirror of complex societal contradictions across different levels [1]. Effective management of online public opinion requires a comprehensive and intuitive understanding of its dynamics. This study adopts a sociological perspective while leveraging quantitative and visualization methods from information science to reveal the multi-dimensional social attributes of Weibo public opinion, providing a new analytical lens and conceptual framework for interpreting the societal issues underlying these phenomena.

2 Research Status

Online public opinion refers to the collective expression of public discourse, emotions, and attitudes toward social issues via the internet [2]. In the field of information science, scholars have conducted diverse research on online public opinion in recent years. Ding Xuejun examined the status of online public opinion regarding public health emergencies [3]; Li Fei analyzed characteristics and patterns in public opinion dissemination using social network analysis [4]; M.J. Lee et al. [5] analyzed users' willingness to express opinions based on Weibo

comment content; S. Blinder [6] conducted simulation studies based on user posts, comments, and tags to reveal internal and external drivers of public opinion propagation; and Jia Hongyu [7] constructed network relationship graphs using forwarding and commenting data between users. In visualization research, most studies have focused on building knowledge graphs through bibliometric analysis of literature from databases like CNKI. Scholars such as Huang Wei [8], Zhang Baosheng [9], and Su Nan [10] employed citation analysis tools and Ucinet to construct scientific knowledge maps analyzing research trends in online public opinion, while Xu Xing [11] built parallel coordinate visualizations based on comment and forwarding counts.

However, current visualization research on online public opinion data either uses bibliometric data or merely analyzes numerical indicators such as forwarding and comment volumes, rather than conducting in-depth mining of textual content from Weibo posts. In sociological research, online public opinion has shifted from "offline" to "online," representing a prominent reflection of various social contradictions [12]. Liu Quan [13] studied the evolution characteristics of coupled online-offline networks based on individual social attributes and proposed a public opinion evolution model; Luo Chuying [14] analyzed irrational online public opinion from a social psychology perspective and proposed corresponding regulatory mechanisms; and Song Xiangli et al. [15] conceptualized online public opinion as a projection of public social attitudes, thoughts, values, and evaluations on the internet, interpreting it through subject-object components. Despite these contributions, comprehensive multi-dimensional analysis of public opinion events from sociological perspectives remains underdeveloped, limiting our ability to grasp the influence of Weibo public opinion across multiple social attributes.

In summary, while online public opinion research has become a priority, few studies integrate sociological perspectives with the social attributes inherent in the development of online public opinion to deeply explain its essential nature. This study combines three social attributes of Weibo public opinion—content, sentiment, and crowd—with three external manifestations—events, emotions, and opinion leaders. Grounded in sociological theory and employing information science methodologies, we analyze the "people-event-emotion" dimensions of online public opinion through a case study of a vaccine incident. We address three key questions: (1) How to interpret the three social attributes of Weibo public opinion from a macro sociological perspective; (2) How to use empirical analysis to reveal the characteristics of different social attributes in a vaccine incident and visualize them; and (3) How to leverage quantitative and visualization methods from information science to map the full picture of public opinion and depict its evolution.

3 Integration of Sociological Perspective and Information Science Methods

Multi-source heterogeneous online public opinion data contains vast amounts of information. Deep mining of the societal issues, developmental context, trends, and evolution patterns behind public opinion, combined with visualization for presentation and analysis, is crucial for understanding the social essence and evolution patterns of online public opinion at a macro level and for coupling online society with offline reality. Xie Lulu [16] proposed four key factors of online public opinion from a sociological perspective: the emergence of opinion leaders, deepening of discourse symbols, herd behavior, and the function of mainstream media. Mainstream media and opinion leaders both reflect crowd attributes, while discourse symbols and herd behavior manifest as expressions of event content on online platforms. However, the fermentation and dissemination of online public opinion involves not only interaction between crowds and events but also emotional diffusion that ignites public opinion evolution. Therefore, we synthesize, integrate, and supplement previous research to construct a model of social attributes and external manifestations of Weibo public opinion. The model centers on public opinion, explaining the internal logic among three social attributes (crowd, content, sentiment) and three external manifestations (opinion leaders, events, emotions), forming a stable structure. Internal attributes provide sociological dimensions for analysis, while external manifestations employ information science methods—including opinion leader evaluation systems, social network analysis, and sentiment classification standards—to visualize these attributes.

The model comprises three features: the crowd influence intensity of opinion leaders, the content attributes of events, and the sentiment attributes of public opinion emotions. It analyzes hot public opinion events from the perspectives of people, events, and emotions to visualize trends and influence scope, aiming to uncover the intensity and driving forces of propagation. Referencing relevant weight evaluation index systems [17] and employing social network analysis and sentiment analysis methods, we construct visual knowledge graphs of opinion leaders, social network relationship diagrams, and emotional trend charts to visualize the social attributes of hot public opinion events across these three dimensions.

3.1 Crowd Attribute of Public Opinion: Opinion Leaders

In sociology, scholars [1] argue that the virtuality and anonymity of online environments weaken individual responsibility and reduce external constraints, leading to "deindividuation" where individual behavior separates from attitudes. This phenomenon causes individuals to conform to group norms while continuously interacting in online society. During public opinion propagation and evolution, different groups exhibit significantly varying influence. Opinion leaders provide information, clues, and attitudinal tendencies to other groups during public opinion dissemination, possessing social integration and mobilization ca-

pabilities. Identifying and quantifying the influence of opinion leaders is crucial for monitoring and guiding discourse direction. Unlike real society, online opinion leaders are not necessarily individuals with social status but rather those who can timely provide information and guide public opinion trends through their statements. Leveraging quantitative methods from information science and referencing Weibo opinion leader evaluation systems [17], we determine weights for parameters including comment counts, forwarding counts, likes, and follower numbers to quantitatively identify opinion leaders on specific topics.

3.2 Content Attribute of Public Opinion: Events

In sociology, the information content disseminated during the development of online public opinion constitutes its basic form and primary substance [1]. For complex and rapidly spreading public opinion content, we must both grasp the overall thematic distribution, extract specific details, and comprehensively understand the scope of topics and discussion content to identify the overall distribution of public attention. Simultaneously, we must track the evolutionary process of event content by refining time granularity to discover how content evolves across time nodes, revealing shifts in public attention focus behind these changes. Social network analysis, widely used in information science, treats each entity as a node and connections between nodes as edges representing relationships, where nodes mutually influence each other with synergistic and dependent relationships, and edges represent propagation paths [18]. Using tools such as Ucinet, Netdraw, and Gephi, we construct knowledge graphs of public opinion content attributes to analyze overall topic distribution characteristics, parse relationships between keywords, and investigate content attributes during evolution using metrics like degree centrality.

3.3 Sentiment Attribute of Public Opinion: Emotions

Sociological research considers emotion a key driver of contemporary public opinion evolution [19], spreading from individual to group levels. Intergroup emotion theory in sociology suggests that group emotions help reconcile intragroup and intergroup attitudes and behaviors, and emotional contagion among netizens further drives public opinion development. Similarly, information science recognizes that user emotional expression significantly impacts public opinion propagation and evolution, accelerating dissemination speed and rapidly infecting others. The ROSTCM6 content mining system provides standardized sentiment classification criteria, statistics, and algorithms [20], enabling judgment of users' emotional fluctuation trends at various time nodes and statistical analysis of positive, neutral, and negative discourse. By constructing time-sentiment trend visualizations, we can grasp the emotional attributes during public opinion propagation.

4 Empirical Research: A Vaccine Incident Case Study

Sina Weibo, a microblogging service launched by Sina.com, ranks highly among global websites [21] with 234 million active users, making it an important platform for netizens to express opinions. We therefore selected Weibo as our data source.

4.1 Data Normalization

We used the Gooseeker web crawler to collect Weibo data on the "Changsheng vaccine incident," including search keywords, user IDs, usernames, post content, URLs, publication time, terminal type, forwarding/comment counts, and like counts. On July 15, 2018, the National Medical Products Administration announced that a biotechnology company had falsified production records for freeze-dried human rabies vaccine. Using "XX vaccine" as the search keyword, we collected 34,384 relevant posts from July 1 to September 30, 2018. After manually cleaning the raw data to remove irrelevant posts and eliminating images, links, and videos, we used ROSTCM6.0 (ROST Content Mining System Version) [22] for Chinese word segmentation and stop-word filtering. We constructed a stop-word list to remove meaningless terms such as "among," "collapse," "display," and "full text" (see Table 1). We then conducted word frequency statistics on the segmented and filtered data, storing it as structured normalized data (see Table 2) to build a keyword corpus for the vaccine incident.

4.2 Crowd Attribute Analysis: Opinion Leaders

Opinion leaders exert widespread influence and even guiding effects on ordinary users regarding emotional tendencies and topic direction in public opinion propagation. Identifying and quantifying their influence in the vaccine incident is significant for guiding public opinion direction. Referencing the Weibo opinion leader evaluation system [17] and incorporating four indicator variables—comment count (X_1), forwarding count (X_2), like count (X_3), and follower count (X_4)—with corresponding weights of 0.25, 0.25, 0.125, and 0.375, we defined the influence value S as:

$$S = w_1 \cdot X_1 + w_2 \cdot X_2 + w_3 \cdot X_3 + w_4 \cdot x_4 \quad \text{(Formula 1)}$$

Quantitative calculation identified opinion leaders in this topic. Based on the maximum weight for follower count and the comprehensive influence indicator S, we constructed a radar chart of opinion leader influence scope (see Figure 1 [Figure 1: see original paper]). The three most influential bloggers were People's Daily, CCTV News, and Toutiao News. Analyzing the 11 most influential opinion leaders in the vaccine incident through bar charts of comments, forwards, and likes (see Figure 2 [Figure 2: see original paper]) revealed that Toutiao News, People's Daily, and China Business News had relatively high influence. Notably, 10 of the 11 opinion leaders were central, local, or online authoritative

news media, demonstrating that government and news media play a guiding role in public opinion events concerning people's livelihood, and that the public places greater trust in officially released authoritative data on major safety issues, confirming the core position of authoritative institutions in guiding public opinion on major livelihood issues.

4.3 Content Attribute Analysis: Event Evolution

4.3.1 Evolutionary Analysis of Public Opinion Content Drawing on existing research [23-24], we divided the entire lifecycle into incubation, outbreak, spread, and decline phases. Using the vaccine incident keyword corpus, we created a time-series line chart of daily post volumes, dividing the lifecycle into: incubation (July 3-14, 118 posts), outbreak (July 15-24, 12,244 posts), spread (July 25-August 23, 18,028 posts), and decline (August 24-September 30, 3,873 posts), as shown in Figure 3 [Figure 3: see original paper].

This division enables effective segmentation of different temporal phases in overall propagation trends. To dynamically visualize public opinion information, we displayed keywords and frequencies for each phase (see Table 3). "Changchun," "Changsheng," and "vaccine" remained the highest-frequency keywords across all four phases, while other core keywords varied by phase: the incubation phase featured "cervical cancer" and "biology," indicating information about Changsheng Biotechnology's upcoming cervical cancer vaccine; the outbreak phase focused on "Changsheng," "rabies," and "falsification," reflecting public concern about the cause—the rabies vaccine falsification; the spread phase emphasized "state," "launch," and "investigation," showing a shift in focus to government handling measures; and the decline phase saw attention shift to "vaccination," "revaccination," and "medicine," reflecting public concern about solutions after understanding the government's stance.

To further analyze the relationships among high-frequency keywords throughout the incident, we constructed a social network knowledge graph based on co-occurrence frequency weights for nodes and edges using VOSviewer (see Figure 4 [Figure 4: see original paper]). Larger nodes indicate higher keyword weight and more central positions, while thicker edges represent stronger connections. "Changsheng" and "vaccine" occupied core positions, with "Changchun" and "biology" as secondary cores. Other keywords radiated outward from these four core terms, with "rabies" adjacent to the core and peripheral terms like "supervision," "administration," "unqualified," "batch," "violation," and "vaccination" forming distinct clusters. This visualization reveals that public attention evolved sequentially from understanding the event subject, to the cause, to government attitudes and measures, and finally to feedback and solutions—demonstrating the dynamic content attribute evolution across the lifecycle.

4.3.2 Thematic Analysis of Public Opinion Content While understanding evolutionary characteristics, we must also grasp overall thematic distribution. The social network knowledge graph of high-frequency keywords reveals

four main discussion themes: (1) the perpetrator—Changchun Changsheng Biotechnology's vaccine falsification; (2) the official body—the National Medical Products Administration; (3) the cause—unqualified batch vaccines administered in violation; and (4) historical context—the company's previous illegal activities regarding DPT vaccines.

4.4 Sentiment Attribute Analysis: Emotional Trends

User emotional fluctuations significantly affect propagation speed and scope throughout public opinion evolution. Analyzing overall sentiment trends toward the vaccine incident reflects public attitudes toward livelihood issues. The ROSTEA sentiment analysis tool developed by Wuhan University assesses sentiment orientation for each post, assigning positive scores (>0), negative scores (<0), and neutral scores (=0), with both positive and negative sentiments further divided into general (0-10), moderate (10-20), and high (>20) intensity levels.

Analyzing the entire keyword corpus revealed predominantly negative attitudes. Using data from July 24 (the peak posting day) as an example (see Table 4), most users expressed negative sentiment. The time-sentiment trend chart (see Figure 5 [Figure 5: see original paper]) shows that during the outbreak phase, most users held negative attitudes toward the vaccine incident, with occasional strong positive sentiments, indicating serious public concern about livelihood safety issues that drove the incident to attract widespread attention and government response.

5 Results and Discussion

5.1 Theoretical Model Innovation

Our model innovatively integrates the three social attributes (people, event, emotion) with three external manifestations (opinion leaders, events, sentiments), connecting online society with offline reality to comprehensively interpret profound sociological issues behind online public opinion. While existing studies explore online public opinion, most examine single factors, failing to elucidate interactions among multiple influencing factors. Our model provides an operational theoretical framework that considers both the anonymous, episodic nature of online public opinion and its real-world social context, offering new perspectives for interpreting public opinion's essential characteristics.

5.2 Summary of Empirical Findings

Under our model's guidance, analyzing the vaccine incident across three attributes yielded these conclusions: (1) **Content attribute analysis** revealed that public attention evolved through the lifecycle following the sequence: event subject → cause → government attitude → follow-up measures. Thematic analysis using social network methods identified four discussion themes (perpetra-

tor, official body, cause, historical context), demonstrating the effectiveness of examining both evolutionary process and thematic distribution. (2) **Crowd attribute analysis** showed that for livelihood issues, the public trusts government departments and official media more, confirming that authoritative institutions should play a core guiding role, with their directives actively interpreted and supported to alleviate negative emotions. (3) **Sentiment attribute analysis** revealed predominantly negative public sentiment toward vaccine safety incidents, with accumulating negative emotions accelerating incident fermentation and development, highlighting the need for government guidance and psychological 疏导.

Online public opinion, as a manifestation of social issues in networked society, possesses relevant sociological attributes. Our model centers on public opinion with three internal social attributes (crowd, content, sentiment) and three external manifestations (opinion leaders, events, emotions), explaining their one-to-one correspondence. This integration of sociological perspective and information science methodology provides a feasible model that considers both online characteristics and offline social context, better assisting governments and regulators in understanding essential issues and implementing effective governance and emotional guidance. The vaccine incident case study validates the model's effectiveness while providing comprehensive interpretation across the three dimensions, demonstrating that this innovative sociological perspective combined with quantitative and visualization methods from information science enables both macro-level grasp and in-depth interpretation of public opinion events, offering new analytical perspectives for future research.

References

[1] Hu Yuning, Qi Binbin, Zhu Xuefang. From "empty nest" mentality to "cluster" behavior: Perspective and network mapping of the "empty nest youth" phenomenon [J]. China Youth Study, 2017(8): 36-43, 84.
[2] Huang Siqi. Research on social network public opinion analysis methods based on the Weibo platform [D]. Nanjing: Nanjing University of Posts and Telecommunications, 2015.
[3] Ding Xuejun, Fan Rong, Yang Jinyi. Research status and review of online public opinion on public health emergencies [J]. E-Government, 2017(6): 55-64.
[4] Li Fei, Ke Ping, Gao Haitao, et al. Research on mobile public opinion propagation models and regulatory strategies based on social network analysis [J]. Modern Intelligence, 2017, 37(9): 97-102.
[5] LEE M J, CHUN J W. Reading others' comments and public opinion poll results on social media: Social judgment and spiral of empowerment [J]. Computers in human behavior, 2016, 65(C): 479-487.
[6] BLINDER S. Imagined immigration: The impact of different meanings of 'immigrants' in public opinion and policy debates in Britain [J]. Political studies, 2015, 63(1): 80-100.
[7] Jia Hongyu, Zhao Xueyan, Qi Chenzi. Analysis of Weibo public opinion

maps based on complex networks [J]. Information Science, 2015, 33(7): 104-109.
[8] Huang Wei, Zhang Xiaojun. Visual analysis of research progress on China's online public opinion based on knowledge graphs [J]. Information Science, 2016, 34(6): 87-92.
[9] Zhang Baosheng, Qi Xiaoting. Visual analysis of China's online public opinion research based on scientific knowledge graphs [J]. Information Exploration, 2017(5): 111-120.
[10] Su Nan, Zhang Xuan, Yang Honggang, et al. Visual analysis of domestic online public opinion research based on knowledge graphs [J]. Information Magazine, 2012, 31(10): 42-47, 58.
[11] Xu Xing, Xi Pengfu, Qin Tian. Social network-based public opinion information analysis and visualization—Taking Sina Weibo as an example [J]. Computer CD Software and Applications, 2013, 16(12): 94-95, 98.
[12] Chang Rui. Research on network public opinion and its governance models and mechanisms in mass incidents [D]. Changchun: Jilin University, 2012.
[13] Liu Quan. Research on online public opinion evolution models based on individual social attributes [D]. Dalian: Dalian University of Technology, 2016.
[14] Luo Chuying. Analysis of irrational online public opinion dissemination based on social psychology [D]. Nanchang: Nanchang University, 2017.
[15] Song Xiangli, Cao Shunxian. Multi-dimensional interpretation of the social influence of online public opinion [J]. Henan Social Sciences, 2014, 22(7): 30-34.
[16] Xie Lulu. Sociological analysis of online public opinion in mass incidents [D]. Nanjing: Nanjing Normal University, 2015.
[17] Li Yuzhen, Hu Yong, Xiong Xi, et al. Evaluation model for Weibo opinion leaders [J]. Information Security and Communications Confidentiality, 2013(2): 79-81.
[18] Zhou Xiuyun, Lou Cequn. Concepts, processes, and characteristics of information ecological community succession [J]. Information Theory and Practice, 2011, 34(6): 12-14.
[19] Yu Hong, Ma Xueyi, Li Ruifang. Analysis of the dynamic mechanism of sensitive hotspot public opinion evolution from the perspective of emotional sociology [J]. News Forum, 2018(1): 78-81.
[20] Zhao Rongying, Wang Jing. Visual analysis of social network analysis (SNA) research hotspots and frontiers [J]. Library and Information Knowledge, 2011(1): 88-94.
[21] Wang Xiwei, Zhao Dan, Zhang Changliang, et al. Research on new media network public opinion information dissemination based on social networks—Taking anti-corruption topics as an example [J]. Information Magazine, 2016, 35(3): 103-110, 102.
[22] Fang Qin. Research on the impact of ROST content mining system on content analysis method [J]. Curriculum Education Research, 2014(1): 234-235.
[23] Ji Shunquan, Li Zhuozhuo. Analysis of Weibo public opinion patterns based on information lifecycle—Taking food safety incidents as an example [J]. E-Government, 2015(5): 58-65.
[24] Li Kaiqing. The role of Weibo in guiding public opinion in public health

emergencies—Taking the "Wei Zexi" incident as an example [J]. Radio & TV Journal, 2016(9): 123-124.

Author Contributions:
Mou Dongmei: Guided topic selection, conceptualization, writing, and revision;
Shao Qi: Responsible for research design, theoretical model construction, and initial draft;
Han Nannan: Knowledge graph construction and paper revision;
Wang Ping: Multi-level model construction and adjustment, paper revision;
Jin Shan: Literature review and paper revision;
Jin Chunyan: Literature review and paper revision.

Note: Figure translations are in progress. See original paper for figures.