Research Progress on Online Public Opinion and Path Analysis of Its Thematic Relationships: Postprint

Authors: Sun Zhuo, Zhao Hong, Wang Zongshui

Date: 2023-04-01T00:00:00+00:00

Abstract

[Purpose/Significance] Reviewing the research progress of online public opinion helps clarify the intrinsic relationships and thematic evolution paths among the knowledge base, transmission patterns, early warning mechanisms, and governance strategies in online public opinion research. [Method/Process] First, we analyze the theoretical knowledge base of online public opinion, then combine existing research to categorize online public opinion studies into four themes—influencing factors, transmission paths, early warning mechanisms, and guidance and governance strategies—according to the principle of content progression, and employ content analysis and social network analysis to conduct thematic correlation analysis and evolution path exploration for online public opinion. [Results/Conclusion] The results indicate that Life Cycle Theory, Cognitive Set Theory, Spiral of Silence, Group Polarization Theory, Butterfly Effect Theory, and Governance Theory often serve as the theoretical knowledge base for online public opinion research; six factors—online media environment, social structural pressure, netizens' psychology, triggering events, effective mobilization, and social control forces—are regarded as important influencing elements of online public opinion evolution; and six themes—online public opinion, public opinion events, social media, stakeholders, big data, and information dissemination—are closely related to other research content and play an important bridging role in thematic evolution paths.

Full Text

Preamble

Network Public Opinion Research Progress and Thematic Association Path Analysis

Sun Zhuo[1], Zhao Hong[1,2], Wang Zongshui[3]

[1]School of Economics and Management, University of Chinese Academy of Sciences, Beijing 100190
[2]Sino-Danish College, University of Chinese Academy of Sciences, Beijing 100190
[3]School of Economics and Management, Beijing Information Science & Technology University, Beijing 100192

Abstract: [Purpose/Significance] Reviewing research progress on network public opinion helps clarify the internal relationships and thematic evolution paths among its knowledge base, transmission laws, early warning mechanisms, and governance strategies. [Method/Process] This paper first analyzes the theoretical foundations of network public opinion research, then divides the field into four themes—influencing factors, transmission paths, early warning mechanisms, and guidance/governance strategies—according to progressive content logic. Content analysis and social network analysis are employed to explore thematic associations and evolution paths. [Result/Conclusion] The results show that lifecycle theory, cognitive set theory, spiral of silence, group polarization theory, butterfly effect theory, and governance theory commonly serve as theoretical foundations. Six factors—network media environment, social structural pressure, netizen psychology, triggering events, effective mobilization, and social control—are identified as key influences. Additionally, six themes (network public opinion, opinion events, social media, stakeholders, big data, and information dissemination) maintain close relationships with other research content and play important bridging roles in thematic evolution.

Keywords: network public opinion; content analysis; association relationship; path analysis; dynamic evolution
Classification Number: G250
DOI: 10.13266/j.issn.0252-3116.2021.07.014

1. Introduction

With the development of mobile internet technology and the proliferation of social media platforms such as Weibo and WeChat, the internet has become a second arena for people to express beliefs, attitudes, opinions, and emotions, attracting increasing attention from all sectors of society. Compared with traditional media, information dissemination through online social platforms features rapid transmission speed, wide influence range, and strong interactivity. These characteristics not only break spatiotemporal boundaries but fundamentally alter the path relationships between information disseminators and recipients, transforming network information from a one-way, deep, and dispersed pattern to a multi-centered, comprehensive coverage model that easily creates agglomeration effects. Against this backdrop, every netizen serves as both information receiver and disseminator, expressing viewpoints in real time and interacting with others, thereby creating a dynamic network environment where communi-

cators intertwine, themes multiply, and effects become unpredictable.

Due to its openness, immediacy, and convenience, the internet allows sudden social events or hot topics to be published and spread rapidly through online media platforms, easily forming network public opinion under the joint promotion of social organizations, media platforms, and netizens. Network public opinion represents a form of social sentiment—a collection of strongly influential and tendentious statements and viewpoints that netizens hold regarding certain popular real-life events in the internet environment. It plays a dual role in social order stability and national governance. On the positive side, netizens can extensively participate in discussions of public events and express their attitudes, views, and opinions in real time, which enhances public engagement in social affairs while providing intellectual support for governments, enterprises, and other organizations to understand public sentiment and formulate valuable references. On the negative side, network anonymity inevitably allows harmful information to infiltrate alongside genuine information, creating situations where truth becomes indistinguishable and "bad money drives out good," potentially triggering large-scale, multi-centered opinion crises that pose significant challenges to social stability and governance.

Faced with endless network public opinion incidents, academic and government circles urgently need to grasp evolution patterns, identify netizen viewpoints accurately, and construct precise early warning models to provide effective guidance for national governance. After decades of development, scholars have conducted extensive research on network public opinion from multidisciplinary backgrounds, multiple theoretical perspectives, and various methodological approaches. However, existing studies remain fragmented, limited by time constraints, or lack dynamic perspectives, making it difficult to present a complete picture of network public opinion research. Traditional text clustering and LDA models often create redundant content associations. Therefore, this paper provides a panoramic review by dividing network public opinion research into four themes—influencing factors, transmission paths, early warning mechanisms, and guidance/governance strategies—according to progressive content logic. It innovatively employs keyword co-occurrence to construct sub-networks and association matrices for these themes, using social network analysis to calculate closeness centrality between sub-networks and reveal dynamic evolution paths.

---

2. Theoretical Foundations of Network Public Opinion

Public opinion is an ancient concept distinct from "public voice" and "popular will," with its connotation continuously enriched over time. Early scholars defined it as a socio-political attitude that the public holds toward governors when stimulated by social intermediary events, where the public constitutes the subject, governors the object, and intermediary events the core. Later re-

search expanded this definition to include the sum of emotions, attitudes, and values formed by the public during the evolution of social events, representing a concentrated expression of "social conditions and public sentiment."

In the internet era, the rapid development of online and social media has broken spatiotemporal limitations in information dissemination and expression, expanding the connotation, forms, and transmission modes of public opinion while posing new challenges for scientific communication and rational policy-making. Some scholars view network public opinion as a conceptual extension of public opinion in cyberspace—a mapping of real public sentiment onto virtual space—suggesting traditional governance models could address new challenges. However, Wang Guohua and Fang Fujian argue that while this extension is reasonable, viewing it as a simple mapping is questionable, as online public opinion exhibits characteristics of directness, bias, randomness, and concealment that differ from real-world public opinion in both form and transmission channels.

Building on traditional concepts, refined definitions describe network public opinion as the collection of consciousness, emotions, attitudes, and behavioral tendencies that netizens express through text, images, video, or audio regarding hot issues, events, or phenomena they care about. Multiple theories have been applied to network public opinion research, including:

Lifecycle Theory: Network public opinion from emergence to demise follows regular, periodic, and continuous patterns. Scholars have proposed three-stage, four-stage, five-stage, and six-stage models, plus various improved versions, to accurately depict stage characteristics and evolution paths.

Cognitive Set Theory: Netizens form cognitive sets based on existing emotional tendencies and social experiences. When public opinion involves objects matching these sets, they actively publish and disseminate information, creating resonant consciousness that diffuses throughout society.

Spiral of Silence Theory: When netizens perceive positive attitudes toward their views, they participate more actively; when they anticipate isolation, they remain silent. Over time, positively perceived views spread rapidly while negative ones become marginalized.

Group Polarization Theory: In group decision-making scenarios, final consensus typically carries higher risk than individual decisions, creating extreme public opinion that severely impacts social stability and economic development.

Butterfly Effect Theory: Netizen behaviors or decisions in virtual space that align with personal desires but conflict with social needs may trigger group harm events, where conflicts escalate due to the butterfly effect.

Governance Theory: Addressing risks in public opinion evolution, this theory provides theoretical support and pathways for government and social intervention through three modes: society-centered governance, network governance, and meta-governance.

3. Analysis of Network Public Opinion Research Content

3.1 Influencing Factors of Network Public Opinion

In the new technological wave, omnimedia and visualization have made network public opinion more dynamic and variable, complicating the exploration of system factors and their interactions. Current research identifies six main influencing factors: network media environment, social structural pressure, netizen psychology, triggering events, effective mobilization, and social control force.

Network Media Environment and Social Structural Pressure constitute the technical hard environment and social soft environment. Technologically, internet and social media development has transformed spatial distance perception and enabled "many-to-many" decentralized diffusion networks, broadening dissemination scope. Mobile internet fully utilizes netizens' fragmented time, providing strong momentum for information aggregation. Socially, relative deprivation theory and value accumulation theory explain how wealth gaps, unfairness, and deprivation feelings drive netizens to express dissatisfaction online. When people experience long-term structural tension, exploitation, and oppression, they vent resentment through extreme behaviors or intense language to gain emotional resonance and behavioral support.

Netizen Psychology serves as the medium converting concrete events into abstract emotions. According to uses and gratifications theory, netizens participate with agency to satisfy needs and expectations. Two psychological mechanisms operate: (1) Empathy theory—netizens show more appreciation for favored objects and sympathy for those with similar experiences, achieving emotional generalization and contagion; (2) Cognitive dissonance theory—when attitudes and behaviors conflict, netizens release tension by publishing information matching their true attitudes. Network virtuality, autonomy, and openness provide convenient environments for polarized psychology and public opinion formation.

Triggering Events, Effective Mobilization, and Social Control are crucial external social factors. Only topics with conflict (attention-grabbing), incitement (emotional resonance), abnormality (curiosity), and drama (heated discussion) become effective triggers. In highly developed network media, triggering events can quickly form general beliefs among netizens, creating representative polarized viewpoints under conformity and spiral of silence effects. Effective mobilization by authoritative opinion leaders during formation stages can guide netizen emotions positively. However, due to event complexity and network anonymity, strong social control (keyword blocking, timely rumor refutation, official announcements, information screening) remains necessary to prevent harmful information spread and maintain social order.

3.2 Transmission Paths of Network Public Opinion

Network public opinion transmission constitutes a complex system with multiple influencing factors and participants, exhibiting strong extensibility and variability. Research shows the evolution chain involves interactions among opinion environment, information, and elements, forming various stage division models.

Wang Laihua proposed a "three-stage model" (occurrence-change-end) based on public opinion's stage characteristics. Xu Jinghong et al. adapted lifecycle theory to develop a "generation-dissemination-aggregation" three-stage model. Liu Yi criticized these models as overly abstract and insufficiently considering threshold settings, proposing a "fluctuation-order change-conflict-attenuation" linear continuous model based on system dynamics. Xie Yungeng and Rong Ting incorporated internet characteristics and netizen heterogeneity to create a four-stage model (formation, outbreak, alleviation, and 平复). As network environments grew more complex, five-stage and six-stage models emerged with stronger generalizability and inclusiveness, though their multi-stage divisions and threshold settings require further empirical validation.

3.3 Early Warning Mechanisms for Network Public Opinion

Network public opinion early warning involves collecting, analyzing, processing, and classifying massive amounts of online information to make value and trend judgments. Four key technologies support this process:

(1) Information Collection and Extraction: Using AI, natural language processing, and neural networks to acquire knowledge and build early warning databases, followed by named entity recognition and keyword/syntactic analysis for information extraction.

(2) Topic Detection and Tracking: The TDT (Topic Detection and Tracking) system uses data mining to segment massive real-time information, conduct similarity analysis, and identify new topics or track existing ones.

(3) Sentiment Analysis: Two main approaches exist—semantic knowledge base methods (using sentiment dictionaries like Tsinghua's Chinese sentiment lexicon, HowNet, or NTUSD) and machine learning methods (treating sentiment analysis as a classification problem). The former is intuitive but dictionary-dependent; the latter offers higher accuracy but struggles with sentiment word divergence.

(4) Evolution Analysis: Using big data and AI to judge development trends based on topic duration, discussion frequency, and sentiment tendencies. Popular methods include pattern recognition, neural networks, topic detection/tracking, and integrated approaches.

Early warning indicator systems have also developed substantially. Dai Yuan and Yao Fei built a security system with three primary indicators: circulation volume, opinion elements, and trend patterns. Zeng Runxi divided indicators

into three categories—warning sources (12 items), warning signs (9 items), and warning situations (9 items). Wang Qing et al. conducted E-R model analysis to categorize indicators into theme popularity, content intensity, growth patterns, and audience tendencies. Specialized indicators have been developed for terrorism incidents, earthquake disasters, and multimedia public opinion decline phases.

3.4 Network Public Opinion Guidance and Governance Strategies

Network anonymity and traffic-driven economics flood the internet with false and genuine information, disrupting quality judgment and creating "bad money drives out good" phenomena. Under relative deprivation and herd effects, false information spreads rapidly, causing social chaos and panic. Governance strategies operate at three levels:

Individual Netizens: As both disseminators and recipients, netizens play crucial roles. Lacking discrimination ability, they may become promoters of false opinion due to empathy. Opinion leaders can integrate scattered views into representative opinion viewpoints. Netizens must improve discrimination ability and media literacy to "not create, believe, or spread rumors," while opinion leaders should actively spread positive energy.

Social Organizations: Enterprises, non-profits, communities, and volunteers should participate effectively to fill the space between government macro-guidance and individual micro-practice. Different social subjects should leverage their unique advantages for tight integration with government and individuals, following principles of collaborative and classified governance.

Government: As the ultimate bearer of residual governance risks, government must adhere to people-centered principles, release authoritative information promptly to stabilize social emotions, and demonstrate determination through actions. It should improve top-level design, form multi-agent emergency governance collaboration mechanisms, monitor external opinion environments, and uphold information disclosure and equal development principles.

---

4. Thematic Association and Evolution Path of Network Public Opinion

Keywords represent research content and themes. This study conducted keyword co-occurrence analysis through the following process: (1) Content mining of four themes to identify search terms; (2) CNKI database searches using term combinations like "network public opinion + factors/elements/motivation/mechanism," "network public opinion + evolution/transmission/path," etc., filtering CSSCI sources to obtain 172, 563, 332, and 832 sample documents; (3) Using Bicomb2 for keyword extraction

and synonym merging to form co-occurrence matrices (frequency > 5); (4) Constructing keyword co-occurrence networks using Pajek [FIGURE:1-4].

Key findings: Beyond the six main influencing factors, opinion leaders, negative public opinion, and stakeholders significantly impact network public opinion. Grounded theory, structural equation modeling, spatiotemporal feature models, and association models are widely used. In the transmission path network [Figure 2: see original paper], opinion events, social media, information dissemination, dynamic evolution, and stakeholders show large nodes and dense connections, indicating that public opinion evolution represents dynamic gaming among stakeholders in media environments. Social network analysis, sentiment analysis, SIR evolutionary game models, epidemic models, LDA-ARMA integrated models, and content analysis dominate research methods. Weibo public opinion receives extensive discussion as a key component.

In the early warning mechanism network [Figure 3: see original paper], semantic knowledge base-based sentiment analysis and web crawler-based machine learning provide technical support for dynamic monitoring, with big data and AI becoming important foundations. In the guidance/governance network [Figure 4: see original paper], network public opinion governance, rumors, Weibo opinion, and network governance appear frequently, reflecting the complexity of the omnimedia era. Concepts like internet thinking, big data, collaborative governance, top-level design, and social visibility offer new solutions.

To explore inter-thematic associations, a 4$×$4 cross-coordinate matrix was constructed [Figure 5: see original paper]. Analysis reveals six common themes across all four research areas: network public opinion [0.94], opinion events [0.73], social media [0.71], stakeholders [0.70], big data [0.63], and information dissemination [0.63], indicating their bridging roles. Two-stage sub-network intersections show varying connections: spatiotemporal feature models [0.62] connect influencing factors and transmission paths; opinion space [0.68] and coordination [0.61] link transmission paths with early warning and governance; Weibo public opinion & sentiment analysis [0.64] connect transmission paths with early warning; public supervision [0.67] links transmission paths with governance; and social visibility [0.71] serves as a bridge between early warning and governance.

---

5. Conclusions and Future Research

5.1 Conclusions

Network public opinion has become a true reflection of value transmission, interest expression, and social supervision in virtual space. Academic research has extensively explored influencing factors, evolution mechanisms, and governance strategies from multidisciplinary perspectives. Thematic association analysis reveals varying attention levels across topics, with governance discussions being

most extensive, followed by evolution paths and early warning mechanisms. Although different stages lack direct connections, they link through network public opinion, opinion events, social media, stakeholders, big data, and information dissemination, achieving healthy panoramic ecosystem operation.

However, limitations remain: (1) Insufficient interdisciplinary integration with computer science, information management, and intelligence studies; (2) Research focuses on sudden incidents and false information, lacking standardized indicator systems and predictive validation; (3) Studies concentrate on post-event analysis rather than real-time prediction and model validation; (4) Limited attention to individual netizen psychological mechanisms during transmission.

5.2 Future Research Directions

(1) Deepen Interdisciplinary Integration and Strengthen Quantitative Modeling: Future research should enhance connections with computer science, intelligence studies, and information management, incorporating mathematical modeling, system simulation, grounded experiments, and field surveys to evaluate existing theoretical models and indicator systems.

(2) Construct Standardized Models with Enhanced Validation: Comprehensive analysis should clarify distinctions and connections among various information types to build logically rigorous systems. Methodologically, existing indicator models should be matched with different opinion event characteristics, evolution mechanisms, and applicable scenarios to achieve standardized model construction for specific event types. Dynamic assessment and validation should enable precise predictions.

(3) Focus on Netizen Psychology and Explore Individual Cognitive Mechanisms: Current research emphasizes opinion leaders and group viewpoints while neglecting individual psychological trajectories. Future studies should integrate cognitive psychology, social psychology, and quantitative psychology with existing models (epidemic models, SIR evolutionary game models, LDA-ARMA, system dynamics) to investigate individual psychological traits, participation motivations, and decision-making processes.

References

[References 1-59 are preserved exactly as in the original text]

Author Contributions

Sun Zhuo: Data collection, paper conceptualization, and manuscript writing
Zhao Hong: Research question formulation and manuscript revision

Wang Zongshui: Research framework design, data analysis, and manuscript revision

---

Abstract: [Purpose/significance] This paper aims to sort out the research progress of network public opinion, which helps to clarify the internal relationship and thematic evolution path among the knowledge base, transmission law, early warning mechanism, and governance strategy of network public opinion research. [Method/process] Firstly, this paper analyzed the theoretical knowledge base of network public opinion. Then, according to the law of content progression, the network public opinion research was divided into four themes: influencing factors, transmission path, early warning mechanism, and guidance governance strategy. Content analysis and social network analysis were used to analyze the thematic correlation and explore the evolution path of network public opinion. [Result/conclusion] The results show that lifecycle theory, cognitive set theory, spiral of silence, group polarization theory, butterfly effect theory, and governance theory are often used as the theoretical knowledge base of network public opinion research. In addition, the network media environment, social structure pressure, netizens' psychology, trigger events, effective mobilization, and social control force are regarded as the important factors influencing the evolution of network public opinion. Moreover, the six themes of online public opinion, public opinion events, social media, stakeholders, big data, and information dissemination are closely related to other research contents, and they play an important bridging role in the evolution path of themes.

Keywords: network public opinion; content analysis; association analysis; path analysis; dynamic evolution

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