Post-print: Construction of a Technical Framework for Targeted Guidance of Deviant Audience Behavior in Online Public Opinion

Authors: Huang Wei, Liu Yi, Sulin Guo

Date: 2023-10-08T00:00:00+00:00

Abstract

[Purpose / Significance] The anomalous behavior of online public opinion audiences is an important factor in drastic changes in public sentiment. This paper focuses on constructing a technical framework for targeted guidance of anomalous behaviors, providing technical means for implementing guidance on such behaviors. [Method / Process] Based on the analysis of targeted guidance for anomalous behaviors in online public opinion audiences, a framework model is established, and each module of the framework model is described in detail; the core technologies required for functional implementation of each module and ideas for further research are systematically outlined, and the framework is validated using Weibo public opinion as an example. [Result / Conclusion] Targeted guidance is a method combining theory and application. The technical framework model for targeted guidance of anomalous behaviors in online public opinion audiences can, from a technical perspective, address the current defect of guidance strategies being overly macroscopic and lacking in specificity, thereby enhancing the operability of targeted guidance.

Full Text

Technology Framework Model for Targeted Guidance of Anomie Behavior in Network Public Opinion Audiences

Huang Wei[1], Liu Yi[1], Guo Sulin[1],[2]

[1]School of Management, Jilin University, Changchun 130022
[2]School of Architecture and Urban Planning, Jilin Jianzhu University, Changchun 130119

Abstract

[**Purpose/significance**] The anomie behavior of network public opinion audiences constitutes a crucial factor in drastic fluctuations of public opinion. This paper focuses on constructing a technical framework for the targeted guidance of anomie behavior, providing technical means for implementing such guidance.

[**Method/process**] Based on the analysis of targeted guidance for anomie behavior among network public opinion audiences, this study establishes a framework model and provides detailed descriptions of each module. It systematically outlines the core technologies required for functional implementation of each module and proposes directions for further research, using Weibo public opinion as a case study to validate the framework.

[**Result/conclusion**] Targeted guidance represents a method combining theory and application. The technical framework model for targeted guidance of anomie behavior in network public opinion audiences can address the current limitation of overly macroscopic and insufficiently specific guidance strategies from a technical perspective, thereby enhancing the operability of targeted guidance.

Keywords: network public opinion; anomie behavior; targeted guidance

1 Introduction

Currently, the barrier to entry for self-media is low, and activity levels have reached unprecedented heights. Due to the lack of "gatekeepers" to control information content and behavior, some network public opinion audiences exhibit non-normative information behaviors during participation in public opinion events, causing adverse effects on the network ecological environment. Consequently, the discovery and targeted guidance of anomie behavior among network public opinion audiences has become a prominent research topic.

Existing research on targeted guidance technology for anomie behavior among network public opinion audiences includes multiple approaches such as university guidance and government guidance. Regarding university guidance, young students have active minds but lack extensive social practice experience, exhibiting characteristics such as impulsive behavior, susceptibility to influence, and simplistic logical thinking [1]. As an important group among network public opinion audiences, university students are also prone to forming online public opinion [2]. Therefore, research on improving universities' public opinion guidance capabilities and consolidating ideological positions in universities [3] represents a key focus of current public opinion guidance studies. Regarding government guidance, government public opinion guidance strategies are more complex, requiring long-term, coordinated, and interactive mechanisms for public opinion guidance regulation. This necessitates simultaneous action from government, media, and the public, as well as the formulation of corresponding laws and regulations [4]. In government public opinion guidance research, Luo

Xiaoxiao et al. argue that when handling public emergencies, the government must promptly release authoritative information and establish smooth information communication channels to correctly control and guide public opinion [5]. Wang Xiwei et al. believe that during the process of guiding public opinion, the government should pay attention to differences between mobile platform users and non-mobile platform users [6]. Zhao Xiaohang et al. believe that government responses to public opinion must voice timely opinions in the early stages of a crisis [7].

Existing research on network public opinion guidance primarily focuses on macro-level strategies, mostly enumerating response, governance, and recovery strategies for network public opinion one by one. Guidance strategies appear somewhat weak in terms of practical implementation and application. In this study, taking network public opinion events as the background and focusing on researching anomie behaviors of network public opinion audiences in different events, we propose a technical framework for targeted guidance for different anomie behaviors. This framework attempts to use information technology as a foundation to enhance more effective and precise technical guidance functions, thereby compensating for the current weakness in operational capability of network public opinion guidance and providing new ideas for network public opinion guidance.

2.1 Conceptual Definition of Anomie Behavior

Yu Zhenhai believes that the anomie behavior of network media is behavioral deviation that violates social norms and inappropriately uses the internet, thereby deviating from mainstream social requirements [11]. Xie Hui believes that anomie behavior includes opinion violence, opinion kidnapping, opinion extremism, and opinion distortion [12]. Yang Qi believes that anomie behavior deviates from socialist core values and will cause adverse effects on society [13]. Synthesizing the above definitions of anomie behavior, this paper defines anomie behavior of network public opinion audiences as all behaviors of online public opinion audiences that produce negative effects, lead public opinion in wrong directions, and cause public opinion crises during network public opinion events.

2.2 Conceptual Definition of Targeted Guidance

Targeted guidance for anomie behavior of network public opinion audiences is guided by three dimensions: network public opinion audience, network public opinion event, and anomie behavior of network public opinion audience. Targeted guidance is implemented on the basis of progressive layering and focusing across these three dimensions. Its analysis process involves analyzing the background of network public opinion events to obtain independent event features, determining whether audiences with independent features have the potential to participate in network public opinion events, and then fitting and predicting their participation behavior based on audience independent features, thereby

proposing targeted guidance strategies according to the fitting and prediction results. Targeted guidance is a technical method based on machine learning algorithms that automatically proposes corresponding guidance strategies according to features in each dimension. Its notable characteristic is achieving accurate and automated public opinion disposition.

3 Goals and Principles of Technical Framework Model Construction for Targeted Guidance of Anomie Behavior of Network Public Opinion Audience

3.1 Goals of Technical Framework Model Construction

The construction of the technical framework model for targeted guidance of anomie behavior among network public opinion audiences must follow four basic steps and target processes: real-time data acquisition, parallel data processing, targeted matching, and fully automatic output. First, real-time data acquisition means that under the entire framework model, data in any module is continuously updated and enriched, requiring real-time data acquisition to achieve timeliness in targeted guidance. Second, parallel data processing means that under the framework model, the individual typing module, event matching module, behavior matching module, and trigger point discovery module follow distributed parallel processing requirements. Both audience profiling and public opinion event perception operate according to multi-threaded distributed computing without temporal sequence. Third, targeted matching means the framework must match more reasonable guidance strategies and paths according to network public opinion audiences, event contexts, and anomie behaviors of audiences in event contexts to achieve targeted guidance. Fourth, automatic output means the entire system operation within the framework does not require manual intervention; from data acquisition and processing to audience profiling, public opinion event perception, anomie behavior fitting, and targeted guidance is an automatic process.

3.2 Principles of Technical Framework Model Construction

The construction of the technical framework model for targeted guidance must follow three principles: intelligent data storage, tight model integration, and streamlined operational processes. Intelligent data storage refers to the optimized storage of data after crawling throughout the framework structure. After data acquisition, most data belongs to raw data, which is transformed into usable data through data cleaning and semantic recognition processes, and then forms a unified tag system through network public opinion audience profiling and network public opinion event perception. This tag data needs to be stored in databases in table form. Meanwhile, targeted guidance strategies for anomie behavior among network public opinion audiences also need to be stored in databases in tag form. The database includes three databases: audience profiling database, event perception database, and targeted guidance strategy

database, which have interconnectivity and data interoperability functions.

Tight model integration refers to the seamless connection of multiple mathematical models throughout the framework, including anomie behavior fitting and targeted guidance of network public opinion audiences, such as data acquisition models, semantic recognition models, behavior prediction models, event matching guidance models, audience typing guidance models, and anomie behavior matching guidance models. These models need to be integrated through the technical framework model for targeted guidance of network public opinion to form unified data interfaces, achieving seamless connection from previous model outputs to subsequent model inputs.

Streamlined operational processes refer to the rationalization of module operation sequences throughout the framework. Module operation exists in both sequential and parallel processing situations. The framework must determine which modules can operate in parallel and which must follow sequential order, as well as which steps data must undergo from input to output for anomie behavior fitting and targeted guidance of network public opinion audiences. Rationalization of overall processes not only clarifies the operational flow of the framework model but also improves framework operational efficiency through parallel computing and processing of data and models.

4 Technical Framework for Targeted Guidance of Anomie Behavior of Network Public Opinion Audience

The technical framework model for targeted guidance of anomie behavior among network public opinion audiences is constructed according to a three-dimensional system of network public opinion audience, network public opinion event, and anomie behavior of network public opinion audience. It includes components such as data acquisition module, targeted guidance strategy database, anomie behavior fitting module for network public opinion audiences, audience profiling database, event perception database, individual typing guidance module, behavior matching guidance module, event matching guidance module, and trigger point discovery module. The framework process runs from data acquisition to behavior fitting to targeted guidance strategy output, with each module interconnected. The framework structure for targeted guidance of anomie behavior among network public opinion audiences is shown in Figure 1 [Figure 1: see original paper].

4.1 Module Composition Analysis of Technical Framework for Targeted Guidance

The module composition of the technical framework for targeted guidance of anomie behavior among network public opinion audiences includes seven parts: data acquisition module, anomie behavior fitting module, database module, individual typing guidance module, behavior matching guidance module, event

matching guidance module, and trigger point discovery module. Specific module positions are shown in Figure 1.

First, the data acquisition module includes audience data acquisition and event data acquisition. The audience discovery module extracts, collects, and processes information such as audience homepages, posts, and comments. The public opinion monitoring module extracts, collects, and processes blog posts and comment information from network public opinion events in which network public opinion audiences participate. The data acquisition module serves as the information input unit for the entire framework for anomie behavior fitting and targeted guidance of network public opinion audiences, implemented through web crawler tools as a continuous processing procedure.

Second, the anomie behavior fitting module includes network public opinion audience profiling, network public opinion event context perception, and anomie behavior fitting for network public opinion audiences, corresponding to audience profiling models, event context perception models, and audience anomie behavior prediction models. Profiling and context perception function to conduct type matching for secondary processed data of network public opinion audiences and events and store processed data in databases. Audience anomie behavior prediction combines improved S-O-R (Stimulus-Organism-Response) models, treating multi-type audiences in the network public opinion audience profiling database as organism groups, allowing each organism type to be stimulated by event contexts separately to predict behaviors and discover anomie behaviors. The anomie behavior fitting module is a data processing unit that prepares for targeted guidance of anomie behavior among network public opinion audiences.

Third, the database module includes the audience profiling database, event perception database, and targeted guidance strategy database. The audience profiling database and event perception database are built upon network public opinion audience profiling and network public opinion event context perception, forming user classification sets and event classification sets under corresponding tag systems. The targeted guidance strategy database is a collection of targeted guidance strategies, including guidance strategy sets and guidance path sets, both annotated according to the three major elements of audience, behavior, and event. Different specific guidance strategies and paths exist for different audiences, behaviors, and events. The database module serves as the data storage unit for the technical framework model for targeted guidance of network public opinion, forming the data foundation for implementing targeted guidance of anomie behavior among network public opinion audiences.

Fourth, the individual typing guidance module is an interface connecting network public opinion audience profiling with targeted guidance strategies for network public opinion audiences. It automatically matches guidance strategies and paths from the targeted guidance strategy database according to unique features of individual network public opinion audiences.

Fifth, the behavior matching guidance module is an interface connecting anomie

behavior fitting for network public opinion audiences with targeted guidance strategies for network public opinion audiences. It automatically matches guidance strategies from the targeted guidance strategy database according to different anomie behaviors of network public opinion audiences.

Sixth, the event matching guidance module is an interface connecting network public opinion event context perception with targeted guidance strategies for network public opinion audiences. It automatically matches guidance strategies from the targeted guidance strategy database according to different network public opinion events.

Seventh, the trigger point discovery module is the activator for implementing targeted guidance. Different network public opinion events and different network public opinion audiences have varying conditions for activating targeted guidance. Based on results from anomie behavior prediction and fitting for network public opinion audiences, trigger conditions are automatically calculated. Since behavior occurrence times and patterns differ across audiences, and whether and when to trigger guidance also varies for different events, determining the optimal targeted guidance trigger time for specific audiences is extremely important.

4.2 Operational Flow Analysis of Technical Framework for Targeted Guidance

The basic steps for anomie behavior fitting and targeted guidance of network public opinion audiences are data update, behavior fitting, audience positioning, strategy matching, guidance triggering, and targeted guidance.

First, data update involves the continuous collection and updating of audience data and event data. This process is based on crawling models, proceeding from capturing hot public opinion event content to capturing audiences of hot public opinion events. Data updating is a continuous process along the timeline. Due to Weibo's anti-crawling mechanism that results in account bans, data crawling frequency is slightly reduced, with the next event crawl occurring after each time interval. After data updating, network public opinion event data enters the context perception database through the context perception model. If the public opinion event already exists in the context perception database, the event data is updated; if not, the event is added and initialized. Network public opinion audience data is stored in the audience profiling database through audience profiling. If the audience already exists in the audience profiling database, the data is updated; if not, the audience is added and initialized.

Second, behavior fitting involves continuous execution of the data acquisition module, where new network public opinion audiences are continuously stored in the audience profiling database through audience profiling, and new public opinion events are continuously stored in the event perception database through context perception. Unlike network public opinion audiences entering the audience profiling database, new public opinion events are screened, and events

with high crisis levels serve as the stimulus component in S-O-R models, affecting corresponding network public opinion audiences in the profiling database. Appropriate anomie behavior prediction models for network public opinion audiences are invoked to predict potential audience behaviors, forming anomie behavior fitting results. Behavior fitting results output event information and behavior prediction information.

Third, audience positioning identifies individual network public opinion audiences who may produce extreme behaviors or cause adverse impacts based on anomie behavior fitting results, positioning audiences within the audience profiling database and outputting audience information. On the basis of audience positioning, audiences are matched with strategies in the strategy database to conduct targeted strategy guidance.

Fourth, strategy matching utilizes the targeted guidance model for anomie behavior among network public opinion audiences to match strategies in the targeted guidance strategy database according to previously output audience information, network public opinion event information, and behavior prediction information. Matching results are integrated and prepared for output. The purpose of strategy matching is targeted guidance, a targeted mechanism generated through further mining of network public opinion audience features, network public opinion event features, and anomie behavior features of network public opinion audiences. It enables guidance and control of events, audiences, and consequently audience anomie behaviors with the most accurate strategies at the optimal timing.

Fifth, guidance triggering utilizes the trigger point discovery model to calculate reasonable guidance trigger timing during strategy matching and strategy preparation output. When timing criteria are met, guidance strategy output and guidance path output are triggered. Strategy output is conducted under the premise of global control of events, network public opinion audiences, and anomie behaviors of network public opinion audiences. Therefore, targeted guidance needs to consider and accurately judge trigger conditions and timing for different situations, enabling precise guidance implementation and avoiding problems such as blind guidance, untimely guidance, excessive guidance, or inaccurate guidance.

Sixth, targeted guidance outputs targeted guidance strategies and paths to network public opinion control professionals to provide guidance. Through targeted guidance, the goal of timely, appropriate, moderate, and precise guidance of anomie behavior among network public opinion audiences is achieved. Timely means taking measures to guide audience voice before anomie behavior occurs; appropriate means adopting accurate guidance measures without blind guidance; moderate means guidance measures must be proper, neither excessive nor ineffective; precise means guidance must accurately lock onto individual audiences and implement targeted guidance according to specific audience features. During framework model operation, audience profiling and context perception in behavior fitting, as well as strategy matching and guidance triggering, belong to parallel operation modules, each controlled by different control units for par-

allel computing. Other steps follow temporal sequence. The targeted guidance flow for anomie behavior among network public opinion audiences is shown in Figure 2 [Figure 2: see original paper].

5 Core Technologies for Targeted Guidance of Network Public Opinion

The core technologies for targeted guidance of network public opinion include data processing technology, behavior fitting and matching technology, and database technology. Data processing technology includes data acquisition, data cleaning, word segmentation, and word vectors. Behavior fitting and matching technology includes semantic recognition, user profiling and context awareness, and behavior prediction.

5.1 Data Processing Technology

5.1.1 Data Acquisition First, based on all Weibo posts related to a certain event, typical Weibo posts at event turning points are selected according to event status, party status, and Weibo comment volume. The Weibo post with the most original comments and richest emotions during the event outbreak stage is selected as the research object. Second, all comments under this Weibo post are traversed and crawled, during which each public opinion audience's homepage link is extracted and audience homepage information is collected. Third, all Weibo posts published by the public opinion audience are traversed and crawled. Finally, during the traversal process, if participation in other network public opinion events is discovered, the process returns to the first step to select all Weibo posts related to that event as background, forming a circular public opinion information crawling flow.

Data crawling should center on S-O-R theory, capturing only necessary data for research. The public opinion event and key Weibo link belong to event context, representing the stimulus component in S-O-R theory. Audience information and audience Weibo posts belong to the audience organism, representing the organism component in S-O-R theory. Audience Weibo posts and Weibo comments belong to organism response, representing the response component in S-O-R theory. The entire Weibo data crawling process uses the public opinion information crawling flow loop as its framework and S-O-R theory as its core, forming the data acquisition module shown in Figure 3 [Figure 3: see original paper].

Overall, the network public opinion information data acquisition module can be divided into sub-modules such as public opinion events, key Weibo posts, Weibo comments, audience information, and audience Weibo posts, collectively manifested as three components: stimulus, organism, and response. As the foundation for data modeling and data analysis, effective data cleaning, data screening, and data preprocessing operations are essential.

5.1.2 Data Cleaning During data crawling, issues such as data source problems, data format problems, and network fluctuations during crawling often result in chaotic data formats, non-uniform formats, data duplication, and missing values [14]. Therefore, data cleaning is necessary to ensure data integrity and uniform format.

During data cleaning, different crawling formats for public opinion events, key Weibo posts, Weibo comments, audience information, and audience Weibo posts are numbered with humans, events, and texts as centers respectively, forming one-to-one corresponding datasets of human-text-event. For different datasets, text format portions are unified into UTF-8 format; portions containing numbers are uniformly defined as Double format; portions containing time are uniformly organized according to yyyy/mm/dd format; all data unit cells apply the Strip function to remove extra spaces.

Due to network delays and data source issues during crawling, numerous missing values appear. Common missing value handling methods include global constant filling, central measure filling, same-group mean filling, and most likely value filling. For missing value problems, this study processes them as follows: since missing values contain both text and numeric data, any data containing missing values is uniformly deleted to facilitate subsequent processing.

5.1.3 Word Segmentation This study adopts Python programming technology, applying the Chinese word segmentation tool jieba to achieve Chinese word segmentation for text content. Considering special words such as pronouns, prepositions, and modal particles that are irrelevant to tweet semantics, this study adopts the stop word tool in jieba for directional segmentation. In this experiment, only 20 word categories are adopted from segmentation results, including general nouns, personal names, transliterated personal names, place names, transliterated place names, institutional group names, other proper names, noun idioms, noun morphemes, new words, location words, general verbs, auxiliary verbs, gerunds, verb morphemes, adjectives, adverbial adjectives, adverbs, and distinctive words. Additionally, stop words not filtered out such as “have,” “don’t have,” “still,” “is,” “also” are added [15].

5.1.4 Word Vectors On the basis of word segmentation, transforming human-understandable text information into computer-understandable information has become a key technology in natural language processing. The initial word vector model represented the presence or absence of different words in a matrix using (1,0) values, but word vector matrices formed in this manner were overly sparse and lacked similarity relationship descriptions between words. Therefore, in 2013, T. Mikolov et al. proposed the Word2vec model [16], which can quickly and efficiently train word vectors and reflect association relationships between words. The Word2vec model includes two neural network-based training models: CBOW (Continuous Bag of Words) model and Skip-gram model. The CBOW model predicts the current word through context, while

the Skip-gram model predicts context through the current word. These two opposite training methods correspond to different data requirements: CBOW performs well in small corpora, while Skip-gram performs better in large corpora. Since both algorithms have unprecedentedly large parameter training scales when dealing with massive data, consuming enormous time, T. Mikolov introduced two optimization algorithms: Hierarchical Softmax and Negative Sampling. Combining the two training algorithms with the two optimization algorithms yields four frameworks.

This study adopts the 200-dimensional word vector model trained by Tencent AI Laboratory [17]. This model is an optimized version of the Skip-gram model called Directional Skip-gram. The Directional Skip-gram algorithm considers the relative positions of word pairs in addition to their co-occurrence relationships within text windows, thereby improving the accuracy of semantic representation of word embeddings. It features rich training samples, fresh corpora, and high accuracy. However, applying large-scale data training to improve accuracy cannot meet the efficiency requirements of the word vector process, making semantic recognition operations using word vector models slow. Applying neural network models requires only one training session to obtain training parameters, after which semantic recognition processes simply need to call these training parameters without affecting recognition efficiency.

5.2 Behavior Fitting and Matching Technology

5.2.1 Multi-Scale Convolutional Neural Network Semantic Recognition Algorithm The network public opinion audience semantic recognition model macroscopically utilizes convolutional neural network models to classify Chinese corpora published by network public opinion audiences. Microscopically, it encapsulates comments and original Weibo posts published by network public opinion audiences into computer-readable digital information through word vector processing, then uses the semantic recognition and classification functions of convolutional neural networks to categorize them according to emotional intensity into five levels: very low, low, medium, high, and very high; according to emotional polarity into three categories: negative, neutral, and positive; and according to topic category into: social livelihood, political economy, culture and entertainment, education and technology, natural disasters, ideology, public safety, judicial police affairs, and others. Since natural language processed by word vector models exhibits distance relationships between word vectors, analyzing sentence semantics requires considering not only individual word meanings but also connecting multiple words for comprehensive consideration. Convolutional kernels in convolutional neural networks have step-scanning functions that can achieve comprehensive analysis of multiple words by controlling convolutional kernel sizes. Based on convolutional neural network models proposed by X. Yang and K. Shrivastava [18-19], this study constructs a multi-scale convolutional neural network model using different convolutional kernel sizes, as shown in Figure 4 [Figure 4: see original paper].

5.2.2 User Profiling and Context Awareness Modeling User profiling and context awareness are substantially similar processes. User profiling involves structured tag descriptions of individual network public opinion audience characteristics, analyzing basic features, behavioral characteristics, and preferences of the audience. Context awareness involves analyzing the environment in which network public opinion audiences are situated, essentially providing tagged descriptions of public opinion events faced by audiences. Both audiences and events on networks consist of heterogeneous multimedia data such as text, images, and video [20], requiring specific semantic recognition algorithms to interpret these data. This study focuses on text data processing, applying the semantic recognition model mentioned in the previous section. User profiling and context awareness modeling are shown in Figure 5 [Figure 5: see original paper].

5.2.3 Bayesian Network-Based Network Public Opinion Audience Information Behavior Prediction Bayesian networks are a class of probability models based on Bayes' theorem that can solve classification problems. They can reflect interrelationships between data graphically and provide probabilistic descriptions of events. In research on network public opinion audience information behavior prediction, Bayesian networks can combine network public opinion audiences with network public opinion events for analysis, utilizing the user profiling model and context awareness model mentioned in the previous section to derive specific emotional intensity, emotional polarity, behavioral preferences, and behavioral performance types of specific network public opinion audiences in specific network public opinion events, thereby calculating the probability of specific audiences producing specific anomie behaviors. The specific Bayesian network model is shown in Figure 6 [Figure 6: see original paper].

5.3 Database Technology

The technical framework model for targeted guidance of anomie behavior among network public opinion audiences includes three categories of databases: audience profiling database, event perception database, and targeted guidance strategy database. The audience profiling database and event perception database both provide tagged descriptions of specific network public opinion audiences and events. The targeted guidance strategy database is a collection of guidance strategies proposed according to historical records of network public opinion audiences, network public opinion events, and anomie behaviors of network public opinion audiences.

Database design can be divided into three stages from abstract to concrete: conceptual design, logical design, and physical structure design [21]. The following sections briefly describe conceptual design, logical design, and physical design using the network public opinion audience profiling database as an example.

5.3.1 Database Conceptual Design Database conceptual design addresses conceptual design patterns for database requirements. In database conceptual design, database conceptual models must be constructed based on database requirements analysis, namely E-R diagrams. In the technical framework model for targeted guidance of anomie behavior among network public opinion audiences, three categories of databases need to be designed. On the Weibo platform, taking the network public opinion audience profiling database as an example, the conceptual design must integrate historical records of network public opinion audiences, including audience homepages, audience comments, audience Weibo posts, and network public opinion audiences themselves. Network public opinion audiences and audience homepages have a one-to-one entity correspondence, as audiences can independently modify basic information such as nickname, location, and gender through editing their homepage, all obtained through homepage crawling. Network public opinion audiences have a one-to-many entity correspondence with Weibo posts through publishing behaviors (posting or forwarding), with one audience posting or forwarding multiple Weibo posts containing data such as Weibo ID, content, and repost count obtained from data crawling models. Network public opinion audiences have a many-to-one entity correspondence with comments through commenting behaviors, as one audience publishes multiple comment messages collected during Weibo comment information collection. Weibo posts also have a one-to-many entity correspondence with comments, as one Weibo post often contains multiple comments. Through the above conceptual analysis, the E-R diagram for network public opinion audience profiling database conceptual design can be derived as shown in Figure 7 [Figure 7: see original paper].

5.3.2 Database Logical Design Database logical design transforms conceptual relationships into logical relationships. During database logical design, three normal forms must be satisfied to eliminate data redundancy and improve database efficiency.

The first normal form requires database tables to be two-dimensional tables where fields contain only basic data types with single attributes. All fields in database tables cannot contain multiple meanings or multiple attributes, and each field can be represented through basic attributes such as CHAR, VARCHAR, and INT.

The second normal form requires each table to have only one primary key. No primary key of any database table can be duplicated or contain multiple attributes.

The third normal form eliminates transitive dependency on primary keys. For example, when column C has a dependency relationship with column B, and column B has a dependency relationship with column A, the relationship needs to be split into multiple tables for representation and cannot exist in a single table.

Taking the correspondence between network public opinion audiences and Weibo posts in the network public opinion audience profiling database as an example, Tables 1 and 2 briefly illustrate database logical design.

5.3.3 Database Physical Design In database physical design, DDL (Data Definition Language) is applied in MySQL databases to create database schemas, create table frameworks, modify tables, and delete tables. Subsequently, DML (Data Manipulation Language) is used to add data, modify data, and query data.

5.4 Core Technology Evaluation for Targeted Guidance

The core technologies for targeted guidance of network public opinion include data processing technology, behavior fitting and matching technology, and database technology. The most critical aspect of data processing technology is data acquisition technology, whose core is the network information radar. Network information radar has capabilities for deep mining, continuous collection, real-time updates, remote updates, and optimized task scheduling, enabling vertical retrieval and directional, three-dimensional, comprehensive, and rapid collection of data.

The most critical aspect of behavior fitting and matching technology is semantic recognition technology, whose core is the convolutional neural network model. Convolutional neural network models simulate the human process of reading text information, achieving semantic recognition and understanding of text information. Convolutional neural networks feature self-organization, self-adaptation, and self-learning, enabling directional and precise semantic processing.

Database technology is an indispensable component of any information system. Database application ensures effective storage and retrieval of audiences, events, and strategies. During database design, normal forms are applied to minimize data redundancy, control database scale, and improve database operational efficiency.

6 Case Study and Framework Validation

This study uses the "Li Xincao Drowning" event as background to validate the framework by following framework processes without providing specific descriptions of data and modules.

6.1 Data Update

During data updating, the data acquisition module obtains Weibo event data and network public opinion audience data for the "Li Xincao Drowning" event. Context awareness results are shown in Figure 8 [Figure 8: see original paper], and audience profiling results are shown in Figure 9 [Figure 9: see original

paper]. According to data samples, they are stored in the context perception database and audience profiling database. From Figure 8, we can conclude that the "Li Xincao Drowning" event belongs to the social livelihood category of public opinion events, with negative emotions occupying most of the content and strong emotional intensity. In word frequency statistics, terms such as "Li Xincao," "drowning," "police," and "college student" are prominent. From Figure 9, we can conclude that the specific network public opinion audience "mini baby together" has behavioral preferences for social livelihood public opinion events, belongs to the ordinary network public opinion audience type, posts more original content than forwarded content, has more positive statements, relatively mild emotional intensity, and prominent words such as "life," "child," and "news" in word frequency statistics.

6.2 Behavior Fitting and Audience Positioning

Taking the "Li Xincao Drowning" event as the analysis object and fitting network public opinion audiences in the audience profiling database, we can derive behavior fitting results as shown in Table 3 . According to the participation behavior fitting results described in the table, we can determine the matching degree between specific network public opinion audiences and the event, emotional polarity of participation behavior, emotional intensity, and the likelihood of publishing, forwarding, or @ actions regarding the event. Based on matching degree, audiences most likely to participate in this Weibo public opinion event can be positioned.

Taking the audience "Runner-You Xipeng" as an example, we can conclude that this audience has the highest matching degree with the "Li Xincao Drowning" event, with a 70% probability of posting negative messages, emotional intensity of 0.28, and a 72% probability of adopting publishing actions.

6.3 Strategy Matching, Guidance Triggering, and Targeted Guidance

First, event strategy matching: According to event strategy matching module calculation results, the "Li Xincao Drowning" event is a medium-risk, social livelihood category Weibo public opinion event. The system conclusion is: attention should be paid to medium-risk Weibo public opinion events with warning strategies proposed. Warning strategies include three aspects: (1) Tracking warning strategy, requiring Weibo public opinion managers to continuously monitor and consider notifying public opinion response agencies according to public opinion development; (2) Collaborative guidance strategy, cooperating with leader-type and traditional media-type Weibo public opinion audiences during public opinion handling to leverage their strong dissemination power, wide influence, and high credibility to voice support for positive public opinion development; (3) Continuous attention strategy, as medium-risk Weibo public opinion events have low crisis levels but may evolve into high-risk events, they cannot be taken lightly and require constant attention.

Second, audience strategy matching: The Weibo public opinion audience "Runner-You Xipeng" is an ordinary Weibo public opinion audience. The system conclusion is: attention should be paid to ordinary Weibo public opinion audiences with audience communication strategies adopted. Audience communication strategies include three aspects: (1) Strengthening education strategy, enhancing education for ordinary Weibo public opinion audiences through media publicity and science popularization to improve their response capabilities when facing Weibo public opinion events and reduce anomie behavior occurrence; (2) Targeted information push, regularly pushing positive Weibo content to the audience to create a harmonious Weibo environment; (3) Cutting off access to harmful information, minimizing connections and information acquisition channels between the Weibo public opinion audience and harmful audiences to create a harmonious Weibo environment and prevent audience anomie behavior.

Third, behavior strategy matching: The participation behavior that "Runner-You Xipeng" will adopt regarding the "Li Xincao Drowning" event belongs to negative publishing behavior with low emotional intensity, classified as "very low-risk anomie behavior." The system conclusion is: observation strategies can currently be adopted without the need for direct or indirect intervention.

7 Summary and Outlook

This study constructed a technical framework for targeted guidance of anomie behavior among network public opinion audiences and summarized three core technologies to make the proposed framework technically feasible. Using the "Li Xincao Drowning" event as an example, the study 梳理 ed framework operation processes, calculated framework operation results, and proposed specific targeted guidance strategies based on operation results, validating framework effectiveness. However, the proposed technology for targeted guidance of anomie behavior among network public opinion audiences remains in early research stages, with much content not yet integrated to form a more complete system. In future research, the authors will construct and optimize mature mathematical models from the above seven modules for deeper investigation.

References

[1] Zeng Runxi, Wang Guohua, Xu Xiaolin. Control and guidance of university network public opinion[J]. Information Theory and Practice, 2009, 32(11): 79-82.

[2] Wu Hong. How university libraries can use positive elements of popular culture to guide college students in practicing socialist core values[J]. Library and Information Service, 2015, 59(S1): 143-146.

[3] XIA W. Occupy media highland, lead the consciousness development of college students[J]. Canadian Social Science, 2015, 11(2): 86-93.

[4] Zheng Lei, Ren Yali. Research on the current status of Chinese government agency Weibo[J]. Library and Information Service, 2012, 56(3): 13-17.

[5] Luo Xiaoxiao, He Yue, Xiong Tao. Research on the impact of authoritative information on Weibo content in public emergencies: a case study of the Liuzhou cadmium pollution incident[J]. Library and Information Service, 2012, 56(11): 123-127.

[6] Wang Xiwei, Xing Yunfei, Wang Nan'axue, et al. Research on information dissemination and empirical study of network public opinion in emergencies under the new media environment: a case study of the "South China Sea Arbitration" topic on Sina Weibo[J]. Information Theory and Practice, 2017, 40(9):

[7] Zhao Xiaohang. Research on government information disclosure in the "post-Weibo" era based on sentiment analysis and topic analysis: a case study of the "Tianjin Explosion" topic on Sina Weibo[J]. Library and Information Service, 2016, 60(20): 104-

[8] Liu Yong, Xu Yejing. Research on theoretical issues of context-oriented intelligence services[J]. Information Theory and Practice, 2013, 36(11): 1-4, 19.

[9] Liu Yingjie, Huang Wei, Yan Lu. Model construction of emotional dimension space for network public opinion information based on A-V-P[J]. Information and Documentation Services, 2017(6):

[10] Zeng Shengquan. Skills for responding to network public opinion[M]. Guangzhou: Guangdong People's Publishing House, 2019:100-241.

[11] Yu Zhenhai. Analysis and research on anomie behavior of network media based on the new media environment[J]. Modern Intelligence, 2013, 33(8): 172-174, 177.

[12] Xie Hui. Legal regulation of anomie behavior in China's network society: from the perspective of network public opinion[J]. Legal System Expo, 2019(31): 37-39.

[13] Yang Qi. Research on behavior anomie and countermeasures of Weibo opinion leaders[D]. Changsha: Hunan University, 2016.

[14] Huang Wei, Xu Yejing, Liu Yi. Research on data-driven mechanism for concurrent acquisition of multimedia network public opinion in big data environment[J]. Information Theory and Practice, 2019, 42(6): 42-48, 16.

[15] Huang Wei, Liu Yi, Xu Yejing, et al. Construction of a heat measurement model for network public opinion tweets[J]. Library and Information Service, 2019, 63(20): 17-25.

[16] MIKOLOV T, CHEN K, CORRADO G, et al. Efficient estimation of word representations in vector space[EB/OL]. [2018-06-17]. https://arxiv.org/pdf/1301.3781v3.pdf.

[17] YAN S, SHUMING S, JING L, et al. Directional skip-gram: explicitly distinguishing left and right context for word embeddings[C]//The 2018 Conference of

the North American Chapter of the Association for Computational Linguistics. New Orleans: Association for Computational Linguistics, 2018: 175-180.

[18] YANG X, XU S, WU H, et al. Sentiment analysis of Weibo comment texts based on extended vocabulary and convolutional neural network[C]//2018 international conference on identification, information and knowledge in the internet of things. Beijing: Procedia Computer Science, 2018: 361-368.

[19] SHRIVASTAVA K, KUMAR S, JAIN D K. An effective approach for emotion detection in multimedia text data using sequence based convolutional neural network[J]. Multimedia Tools and Applications, 2019, 78: 29607–

[20] Huang Wei, Liu Yi, Sun Yue. Analysis of key technologies for semantic recognition of multimedia network public opinion[J]. Information Theory and Practice, 2019, 42(1): 134-140.

[21] Huang Wei, Zhao Jiangyuan, Xu Yejing. Research on interoperability elements and mechanisms of multimedia network public opinion knowledge base groups[J]. Library and Information Service, 2019, 63(20):

Author Contributions

Huang Wei: Responsible for framework design and content guidance;
Liu Yi: Responsible for paper writing and model construction;
Guo Sulin: Responsible for paper revision and proofreading.

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