Postprint: Measuring Sentiment Polarization Degree of Opinion Clusters in Online Public Opinion Sphere

Authors: Gao Junfeng, Huang Wei

Date: 2023-07-26T00:00:00+00:00

Abstract

[Objective/Significance] This paper proposes a method for measuring the sentiment polarization degree of viewpoint clusters within the online public opinion arena, providing a basis for quantifying the sentiment posture of public opinion audiences and identifying polarized groups. [Method/Process] First, the conditions for sentiment polarization among public opinion audiences are clarified. Viewpoint clusters satisfying these conditions are then filtered by setting appropriate thresholds. Building upon this, three polarization indicators (audience attraction rate, extreme audience growth rate, and polarized sentiment growth rate) are introduced to characterize the degree of sentiment fermentation of viewpoint clusters within the measurement time window. Finally, smoothing weights are employed to conduct a weighted synthesis of the sentiment performance of viewpoint clusters across different measurement stages, yielding their overall sentiment polarization degree value. [Results/Conclusion] The sentiment performance of viewpoint clusters in each time window can provide a diagnostic basis for stage-specific interventions in audience sentiment polarization, while the comprehensive sentiment polarization degree facilitates accurate identification of sensitive topics and groups within the arena, enabling precise management and control of online public opinion.

Full Text

Preamble

The drivers of netizen emotional polarization constitute a key factor in altering the topological structure of topic networks [**?**]. Scholars have summarized the formation mechanisms of online collective behavioral intentions, which to some extent reflect the emotional polarization performance of a particular stance camp and can identify entities within public opinion issues that are prone to

triggering polarized emotions. This provides a quantitative foundation for monitoring sensitive topics and precisely intervening in irrational public opinion groups. Furthermore, researchers have elaborated on the differences among physiological measurement methods, self-report methods, and scale-based measurement approaches for intention prediction, and have proposed the formation cycles and influencing factors of online collective behavioral intentions [**?**]. In terms of quantitative research, some studies have employed quantile regression models to construct quantitative indicator systems for online group polarization early warning, using real data from online forums to investigate the emergency response feasibility of the models and indicators [**?**]. Other research has built game-theoretic models among governments, ordinary netizens, and opinion leaders during the process of online group polarization, using the expected benefits of game participants at different stages as the basis for maintaining stability in online group polarization [**?**]. From a robustness perspective, some studies have examined the capacity of online public opinion viewpoint communities to maintain three aspects—viewpoint dissemination, emotional mobilization, and internal cohesion—when subjected to attacks and core node destruction, using these capabilities as criteria for determining whether a viewpoint community exhibits polarization tendencies and whether intervention measures are necessary [**?**, **?**]. Additional literature has simulated the process of public opinion polarization and predicted the occurrence of group polarization by defining rules for individual viewpoint interaction and viewpoint evolution among audiences within the public opinion field [**?**].

Extensive research demonstrates that monitoring and responding to netizen emotional polarization is central to online public opinion governance. However, existing research findings have not reached consistent conclusions regarding the judgment and measurement of emotional polarization among public opinion audiences. The emergence of emotional polarization among public opinion audiences necessarily relies on their affiliated viewpoint camps and clusters. Therefore, selecting effective and readily obtainable indicators of emotional polarization degree to calculate opinion clusters represents a reasonable approach for quantifying the emotional posture of different public opinion groups within the field. Nevertheless, not all viewpoint clusters in the public opinion field necessarily exhibit emotional polarization. This is because the nature of entities involved in different public opinion matters varies, and both the behavioral categories and severity levels of these entities can differentially affect the degree of emotional evolution among audiences. Consequently, this study directly treats entities involved in public opinion matters as the basis for dividing audience viewpoint clusters. Theoretically, viewpoint clusters formed around a particular entity can present three types: "positive," "negative," and "neutral." Apart from "neutral" clusters, the other two types can be identified as carriers for the emergence of emotional polarization among public opinion audiences. However, as previously stated, not all clusters will necessarily exhibit emotional polarization phenomena. Therefore, this study first establishes several necessary conditions for emotional polarization based on its definition, screens

viewpoint clusters that satisfy these conditions and identifies them as measurement objects, then describes and calculates the states of elements that influence emotional polarization (such as individual audience emotional intensity, total extreme emotion quantity, and cluster scale) to derive the emotional polarization degree of the measured viewpoint cluster during a specific time period. As time evolves, changes in the states of these elements can also be reflected in the measurement results for subsequent time periods. Finally, by introducing the exponential smoothing method to comprehensively calculate the emotional polarization degree of viewpoint clusters across the entire measurement time series, the obtained results can reflect the emotional polarization performance of a stance camp to a certain extent and identify entities within public opinion matters that are prone to triggering polarized emotions.

2. Definition of Related Concepts

2.1 Emotional Polarization of Audiences in Online Public Opinion Fields

In early proposed mechanisms for the formation of online irrational discourse, group emotional polarization is defined as group members holding a certain positional tendency even before viewpoint interaction occurs, which is then strengthened through group deliberation and discussion, causing a particular viewpoint or attitude to far exceed the original group average level and ultimately become extreme and gain a dominant position [**?**]. The above classical definition provides the general nature of polarization: stance consistency and emotional extremity. The third characteristic concerns whether the viewpoints held by an already polarized group must necessarily overwhelm and dominate all other groups. In other words, can more than one group viewpoint exhibit emotional polarization within a defined domain? This study argues that the simultaneous existence of multiple group emotional polarization phenomena is possible. In the online environment, "disembodiment" and "decentralization" coexist simultaneously. On one hand, online citizens need not forcibly submit to elite viewpoints; on the other hand, their speech is not constrained by real-world identity background restrictions. This leads to competitive and antagonistic viewpoints being encouraged rather than constrained as long as they gain certain recognition, potentially creating a situation where multiple stances are "evenly matched" as audiences defend their respective viewpoint camps, providing possibilities for multi-trajectory, multi-directional emotional polarization. Additionally, as mentioned earlier, multiple entities involved in public opinion matters may exist within the online public opinion field. Audience viewpoints clustering around a particular entity may macroscopically converge with viewpoint clusters of other entities, but microscopically, the topic salience and social sensitivity of different entities or different aspects of the same entity are not equivalent, which may also lead to non-unique emotionally polarized groups in the online public opinion field.

2.2 Conditions for Emotional Polarization of Audiences in Online Public Opinion Fields

Based on the above definition, the conditions for emotional polarization can be summarized as follows:

(1) **Condition 1**: The viewpoint cluster formed around an entity must possess a certain scale. Emotional polarization phenomena exhibit obvious group participation characteristics. Otherwise, even if opinion leaders play a role in emotional mobilization and guidance during this process, without triggering responses from public opinion audiences, the overall emotional intensity required for emotional polarization cannot be brewed. Therefore, whether emotional polarization occurs is first constrained by the number of audiences absorbed by the cluster to be measured. Additionally, the scale of a polarized viewpoint cluster is a necessary condition for maintaining competitive advantages over other polarized clusters within the field. Since emotional polarization is the product of online citizen participation and mobilization, regardless of how extreme the emotional attitudes held by individual audiences may be, they cannot represent the group. Isolated extreme attitudes and statements cannot be regarded as group emotional polarization.

(2) **Condition 2**: The emotional intensity of individual audiences within the viewpoint cluster must exhibit an extreme trend. According to the definition, the occurrence of emotional polarization necessarily accompanies a significant extremist tendency in held viewpoints. This significant extremism cannot be simply equated with the simple superposition of emotions held by large numbers of audiences; rather, individuals holding extreme emotions must occupy a considerable proportion within the group. Because in a large-scale group, even if individual audiences generally hold moderate non-neutral positions, simple accumulation of emotional intensity may overwhelm other relatively smaller groups within the field (even if individual emotional attitudes are generally extreme). This situation violates the definition of emotional polarization. Therefore, while confirming the scale of polarized viewpoint clusters, the proportion of polarized individuals within them must still be considered.

(3) **Condition 3**: The total amount of extreme emotions within the viewpoint cluster must reach corresponding intensity. The total intensity of extreme emotions generated around an entity is the focal point for competing with other polarized groups within the field. The trajectory and direction of emotional polarization among audiences in the online public opinion field are also completed through comparisons of extreme emotional intensity among different polarized groups. Therefore, the total amount of extreme emotions within a group is also an important measurement indicator.

The definition and adoption of polarization conditions serve to quantitatively distinguish whether viewpoint clusters appearing in the field have experienced

emotional polarization during the initial measurement time window, thereby enabling the selection of viewpoint clusters that meet the polarization conditions.

3. Measurement Mechanism

The overall measurement approach is "first identify the clusters to be measured, then extract their polarization indicators, then calculate the emotional polarization degree of the clusters in stages, and finally derive a comprehensive emotional polarization value based on the stage results." In previous research, indicators such as group scale, topic discussion volume, and group interaction frequency [**?**] have been used to measure group polarization. The common characteristic of these indicators is that they are easy to obtain and can all reflect group emotional activity states to a certain extent. Following this 思路, this study not only closely generates polarization indicators around the aforementioned conditions but also strives to select readily obtainable indicators that can reflect cluster emotional states to satisfy the measurement work.

The step of judging clusters to be measured can directly set a series of thresholds to screen out viewpoint clusters that meet the conditions. Although cluster polarization indicators can also directly use values corresponding to the conditions (cluster scale, number of polarized audiences in cluster, total polarized emotion in cluster), for viewpoint clusters that already meet polarization conditions in the initial measurement stage, their total polarized emotion, number of polarized audiences, or number of extreme audiences may show different changes over time. Directly using these values for calculation would inevitably weaken the descriptive power of the overall evolution trend of cluster emotions. Therefore, this study selects three ratio indicators—"extreme audience growth rate," "polarized emotion growth rate," and "audience attraction rate"—as the polarization degree indicators of viewpoint clusters. The first two indicators can describe the evolution trend of the cluster's own emotional state, while the last indicator can serve as a base to describe the growth and decline comparison of various viewpoint clusters across the entire field.

These three indicators form a measurement formula that collaboratively completes the stage-by-stage calculation of the emotional polarization degree of each viewpoint cluster. Then, the stage results are processed to comprehensively derive a value for the emotional polarization degree of a viewpoint cluster. The specific measurement steps are shown in Figure 1 [Figure 1: see original paper].

(1) **Determine the observation time series**: Divide the entire measurement period into n equal parts, making each time window consistent in width, then analyze indicator changes of viewpoint clusters from each window.

(2) **Select entities involved in public opinion**: First, use the NLPIR Chinese word segmentation system to segment opinion texts appearing in the public opinion field, decompose content words to extract nouns, verbs, and modifying emotional words, and use ontologies such as "HowNet"

to normalize and merge word semantics, eliminating ambiguity. Then, adopting the TF-IDF word frequency statistical approach, select nouns that meet the frequency threshold "Ω," and manually supplement the public opinion description attributes of nouns according to context to highlight the topic quality of entities involved in public opinion (the same entity may have multiple topics, and audience attitudes toward topics can be further subdivided to form viewpoint clusters). As public opinion events ferment and netizens develop divergent cognition, new entities and viewpoint clusters may emerge. Therefore, this step needs to be repeated for each time window, and the identification of polarized viewpoint clusters in the next step must also be handled accordingly.

(3) **Identify polarized viewpoint clusters**: In the word frequency statistical step, through semantic analysis of viewpoint content, the clustering of public opinion viewpoints within the field can be completed. Simultaneously, viewpoint clusters that do not meet conditions must be excluded. Based on the semantic analysis steps above, by accumulating the polarity and intensity of emotional words used by audience individuals to modify entities, "entity + topic" can be categorized into one of three stances: Positive, Objective, or Negative (viewpoint clusters formed around the same entity topic correspond accordingly, see Figure 2 [Figure 2: see original paper]). Neutral and rational viewpoints cannot form polarized emotions and therefore are not treated as measurement objects. Thus, multiple viewpoint clusters with non-unique emotional polarities may emerge around an entity, requiring further observation of the initial scale of actual polarized groups. Set the current group size threshold Ant (a viewpoint cluster must have a total audience number not less than this value during the measurement period to be considered as meeting Condition 1), the polarized individual emotion amount threshold $Paet$ (if the intensity of emotional words used in individual viewpoints is not less than this value, the cluster is considered to meet Condition 2), the total polarized emotion amount threshold Pet (if the total value of emotional word intensity accumulated from all viewpoints in the cluster is not less than this value, the cluster is considered to meet Condition 3) as polarization judgment thresholds, and set the initial polarized emotion comparison threshold Δ between different clusters of the same entity (the ratio of total emotion amount of the cluster to be measured to other clusters cannot be less than this value) to reasonably select viewpoint clusters.

(4) **Describe attributes of polarized clusters**: For viewpoint clusters to be measured within the field, their polarization characteristics need to be described from three aspects:

Audience Attraction Rate $(Aar) = \frac{\text{Cluster audience increment in } T_i}{\text{Total audience increment in field in } T_i} \times 100\%$

Extreme Audience Growth Rate $(Grea) = \frac{\text{Total polarized audience in cluster in } T_i - \text{Total polarized audience in cluster i}}{\text{Total polarized audience in cluster in } T_{(i-1)}}$
100%

Polarized Emotion Growth Rate $(Grpe) = \frac{\text{Total polarized emotion in cluster in } T_i - \text{Total polarized emotion in cluster in}}{\text{Total polarized emotion in cluster in } T_{(i-1)}}$ 100%

In the measurement of extreme audience growth rate and polarized emotion growth rate, situations may occur where the total polarized emotion or total polarized audience in cluster in $T_{(i-1)}$ is zero, caused by the introduction of new topics. In such cases, this study defaults the cluster's polarized emotion or audience total in $T_{(i-1)}$ to 0. The increment of audiences in the cluster to be measured is based on whether new audiences publish text-based viewpoints within the field. The same audience's viewpoint posting behavior in different observation windows is not considered audience increment, but if the emotional intensity in their viewpoint evolves from mild to extreme (judged according to the aforementioned polarized individual emotion amount threshold), they are treated as audiences triggered with emotional polarization and included in the calculation of extreme audience growth rate and polarized emotion growth rate.

(5) **Calculate stage emotional polarization degree of each viewpoint cluster**: The three indicators obtained in the previous stage correspond to the three conditions of emotional polarization among public opinion audiences. Since viewpoint clusters that failed to meet conditions have already been eliminated in the polarized cluster identification step, these three indicators can directly outline the stage emotional polarization degree of each viewpoint cluster.

First, the audience attraction rate (Aar) describes the scale growth rate of each cluster to be measured within the field across the time series. Because once an audience's explicit stance is captured, even if a stance reversal occurs in the next period, the audience's polarized state at the previous moment still has statistical significance, the value range of indicator Aar is $[0, 1]$. The audience attraction rate can also describe the types of group polarization (single/dual/multi-directional) in different temporal states of the public opinion field. As previously mentioned, emotional polarization is a product of group participation; isolated extreme emotions that cannot incite and attract sufficient audiences cannot be called polarization. Therefore, the stronger a viewpoint cluster's ability to assimilate and attract audiences, the larger the base of audiences triggered with polarized emotions within that cluster, and the more likely the entity forming that viewpoint cluster is to cause netizen emotional polarization. In this sense, the audience attraction rate indicator serves a horizontal comparison function for emotional polarization degrees among various viewpoint clusters in the online public opinion field.

Additionally, the extreme audience growth rate ($Grea$) and polarized emotion growth rate ($Grpe$) describe changes in polarization conditions of viewpoint clusters to be measured from a temporal dimension. Theoretically, if the number of extreme audiences and the total amount of polarized emotions within a cluster can maintain a steady upward trend, the cluster's emotional polarization degree is considered higher. Because these two indicators can directly demonstrate the

polarization trend of viewpoint clusters, they exhibit a positive correlation with emotional polarization degree. Unlike the audience attraction rate, the value ranges of extreme audience growth rate and polarized emotion growth rate are $[0, +\infty)$. Both describe the internal emotional change state of viewpoint clusters. Although measurement values can accurately reflect cluster polarization states, their global descriptive capability is relatively weak. While the audience attraction rate indicator can consider scale changes of various viewpoint clusters from the perspective of the entire online public opinion field, its precision is slightly insufficient if used alone to characterize emotional polarization degree. Therefore, comprehensive consideration of these three indicators is an inevitable approach to calculating the emotional polarization degree of viewpoint clusters per unit time. This study constructs the unit time emotional polarization degree measurement formula:

$$Epd(T_x) = \begin{cases} \delta Grea \cdot Aar + \varepsilon Grpe \cdot Aar & \text{if } Grea + Grpe \neq 0 \\ 0 & \text{if } Grea + Grpe = 0 \end{cases}$$

In the formula, if either the extreme audience growth rate ($Grea$) or the polarized emotion growth rate ($Grpe$) is 0, the other growth indicator must also be 0. In this case, the formula is not calculated, and the emotional polarization degree of the viewpoint cluster in the current measurement window is directly defaulted to 0. Then, the audience attraction rate is used to multiply with the two growth rates respectively to compensate for the limitations of individual indicators in characterizing polarization degree, enabling the statistical significance of all indicators to function. Furthermore, "δ" and "ε" are two indicator weights set by this study to describe the importance of extreme audience growth rate and polarized emotion growth rate in demonstrating emotional polarization degree. Based on the polarization conditions derived from semantic analysis, this study considers that both growth rates have equally important significance in describing the polarization capability of viewpoint clusters, so the two weight values are defaulted to 0.5.

From Formula (1), the emotional polarization degree of viewpoint clusters in each measurement period (T_x) can be obtained. However, the study ultimately needs to derive a specific numerical value. Considering that each stage's emotional polarization degree in the original time series has a different impact on the final result, different weights need to be assigned to each stage's polarization degree. Therefore, this study selects the exponential smoothing method to comprehensively calculate the stage emotional polarization performance of each viewpoint cluster, ultimately obtaining the overall emotional polarization degree of opinion clusters in the online public opinion field.

Let β represent the weight, "OC" be the viewpoint cluster to be measured, and n be the number of measurement time windows. According to the weight selection mechanism, $\beta = \frac{2}{n+1}$, and the weight set forms a geometric sequence with a common ratio of $\gamma = 1 - \beta$. Each time window corresponds to a unique

weight, with the most recent period's polarization degree that has the greatest impact on the final result assigned the maximum weight value [**?**]. The audience emotional polarization degree calculation formula for viewpoint cluster "OC" is:

$$Epd(OC) = Epd(T_n) + \frac{(T_{n-1}) + ... + 2^{n-1}}{(n+1)} Epd(T_1)$$

Due to measurement mechanism limitations, the indicator value in the T_1 stage is 0.

The entire measurement process for emotional polarization degree of viewpoint clusters in online public opinion fields can be formally expressed as follows:

Input: Word frequency statistical threshold Ω for public opinion entity terms, viewpoint cluster size threshold (Ant), polarized individual emotion amount threshold ($Paet$), total polarized emotion amount threshold (Pet), measurement time series ($T_{begin}, T_{end}/\Delta T$), and initial polarized emotion degree comparison threshold Δ between different viewpoint clusters of the same entity.

Output: Emotional polarization degree $EPD(T_X)$ of viewpoint clusters in each observation time window, and overall emotional polarization degree $EPD(OC)$ of each viewpoint cluster.

1. Call the NLPIR Chinese word segmentation system [**?**] to perform semantic processing on viewpoint texts in the online public opinion field, using word frequency statistical threshold Ω to obtain the public opinion entity set Set_1.
2. Call the emotion ontology [**?**] to obtain audience stance orientations in the public opinion field, then obtain the viewpoint cluster set Set_2, and statistically derive the audience number A_n for each viewpoint cluster.
3. Call the emotion ontology again to calculate the total polarized emotion amount (P_e) and polarized audience emotion amount (P_{ae}) for each viewpoint cluster, and use the polarized individual emotion amount threshold "$Paet$" to identify extreme audiences and their numbers (E_{an}) in each viewpoint cluster in the current time window.
4. Call the viewpoint cluster size threshold (Ant), polarized individual emotion amount threshold ($Paet$) initial polarized emotion degree comparison threshold Δ, to reasonably select viewpoint clusters from set Set_2 attached to the same entity, obtaining the viewpoint cluster set Set_3 to be measured.
5. Represent each viewpoint cluster in Set_3 as a quadruple (P_e, P_{ae}, E_{an}, A_n), introduce measurement time elements, calculate the audience attraction rate (A_{ar}), extreme audience growth rate (G_{rea}), and polarized emotion growth rate (G_{rpe}) for each cluster in the current time window, and obtain the emotional polarization degree $Epd(T?/OC?)$ of each viewpoint cluster in the current period.

6. Update the time window and repeat steps 1-5 until reaching the preset measurement end time node, calculate the smoothing weight β, observe the time window sequence and corresponding weight values for each cluster, and obtain the overall emotional polarization degree $Epd(OC?)$ of each viewpoint cluster.

4. Empirical Research

This study takes the recent "high-speed rail seat hogging" incident as the empirical research object, using the thematic discussion area initiated on "Sina Weibo" as the online public opinion field for this research. The data collection period was from 15:00 on August 23, 2018, to 15:00 on August 31, 2018, with $T = 8$ and $\Delta T = 24$h. After data cleaning, the field contained a total of 16,456 audience members who published viewpoints during the measurement period. Since the purpose of the empirical study is to verify the measurement method, excessive requirements for threshold settings in the paper are unnecessary. During the research process, the author set each initial judgment threshold as follows: $\Omega = 0.1$, $Ant = 400$, $Paet = 20$, $Pet = 4500$, and the initial polarized emotion degree comparison threshold between different stance groups of the same entity $\Delta < 0.39$. Considering the high frequency of semantically equivalent terms for "seat hogging man" in the public opinion field, whenever "seat hogging man" or semantically equivalent vocabulary appeared in statistics, the author directly counted its accompanying public opinion attributes, i.e., topic vocabulary. Through semantic disambiguation and manual semantic merging of high-frequency topic vocabulary, the results meeting threshold conditions were obtained as follows: $\{OC_1$, seat hogging man + seat occupation(3493)$\}$, $\{OC_2$, seat hogging man + apology(2942)$\}$, $\{OC_3$, seat hogging man + plagiarism(2349)$\}$, $\{OC_4$, seat hogging man + high education(1952)$\}$, $\{OC_5$, seat hogging man + punishment(2905)$\}$, $\{OC_6$, seat hogging man + wheelchair provocation(2815)$\}$.

Entities such as "Jinan Railway Bureau + investigation results not illegal," "seat hogging man + keen on matchmaking," "online mob + human flesh search," "seat hogging man + academy staff," and "Chinese Academy of Sciences + innocent bystander" all failed to reach the given threshold frequency. Therefore, this study ultimately obtained six viewpoint clusters for one entity.

Call the emotion ontology to identify the emotional polarity in audience individual text viewpoints. The selected emotion ontology in this study categorizes emotional vocabulary into five intensity values ("1, 3, 5, 7, 9") with positive/negative connotations outlined through intensity value signs [**?**]. In this study, the "positive/negative" sign of the total intensity value of emotional modification vocabulary in audience viewpoints corresponds to their "Positive/Negative" stance. The emotional polarity of the entire viewpoint cluster is obtained through superposition of individual emotions. Additionally, the appearance of degree adverbs modifying emotional words often causes fluctuating effects on the current emotional word intensity values. To accurately

identify the audience's stance toward entities, this study adopts the degree adverb weighting scheme proposed in literature [**?**], dividing degree adverbs into seven categories: "extremely, most (1.5)," "very (1.3)," "slightly (1.1)," "moderately (1)," "insufficiently (0.8)," "excessively (-0.5)," and "negation (-1)," then performs multiplication operations with the values of modified emotional words, and finally accumulates emotional intensity/polarity. If the same audience viewpoint contains multiple topic objects (belongs to multiple viewpoint clusters), emotional values for different topic objects are calculated separately and then categorized according to the rules proposed earlier.

Based on the emotional polarity and intensity calculated for each audience, neutral audiences are eliminated, and the polarization performance of viewpoint clusters formed around each entity + topic in the initial time window is further analyzed. The polarized emotion polarity of each viewpoint cluster is predominantly "Negative." Except for OC_6 whose initial window is T_8, the initial windows for all other viewpoint clusters are T_1, meeting the Ant and Pet thresholds. Additionally, the comparison results of emotion amounts from opposing stance (positive) clusters are all less than 0.2, satisfying the comparison threshold condition. The scale of each viewpoint cluster is shown in Figure 3 [Figure 3: see original paper].

Through the emotional polarity intensity calculation in the aforementioned steps, the polarized audience volume and total polarized emotion for each cluster under the measurement time series were obtained, as shown in Figure 4 [Figure 4: see original paper] and Figure 5 [Figure 5: see original paper].

By calling the polarized cluster attribute description formula, the audience attraction rate, extreme audience growth rate, and polarized emotion growth rate for each cluster in different measurement windows were obtained, as shown in Table 1 .

By calling the measurement formulas "$Epd(T_X)$" and "$Epd(OC)$," the emotional polarization performance of each viewpoint cluster in the measurement time series was obtained (see Table 2), as well as the overall emotional polarization degree of each viewpoint cluster (see Figure 6 [Figure 6: see original paper]).

Through observation of the empirical results, the author found that the emotional polarity of viewpoint clusters around the "seat hogging man" in the entire online public opinion field was highly unified macroscopically. Although the two topics of "seat occupation" and "apology" triggered widespread discussion in the early stages of network public opinion field formation, the aggregated viewpoint clusters did not exhibit high emotional polarization degrees. Meanwhile, the two topics of "plagiarism" and "punishment," which the author considered sensitive and likely to trigger emotional polarization, also performed mediocrely throughout the measurement time series. Conversely, the two topics of "high education" and "wheelchair provocation" showed strong emotional polarization.

The above phenomena mainly arise from the following reasons: (1) In the high-

speed rail seat hogging incident, although the entity involved was relatively unified, numerous public opinion sub-topics still evolved, causing audience attention in the online public opinion field to shift jumpily, and emotional mobilization in each viewpoint cluster generally lacked sustainability. Additionally, since $OC_1/OC_2/OC_3/OC_5$ all showed cases where polarization indicators were 0 in certain time windows, their good polarization indicator performance in early stages was offset. (2) Although the viewpoint cluster formed around "high education" showed average polarization performance throughout the measurement cycle, compared with other viewpoint clusters, topics about the entity's identity could more easily connect with newly emerging public opinion topics, thus receiving more sustained attention and achieving higher emotional polarization degrees. (3) The entity provocation event appeared at the end of the measurement period, resulting in the highest emotional polarization degree. Moreover, due to the emergence of this topic, the online public opinion field that had already entered a dormant period was reactivated, and topics that had faded over time regained attention. The topic with good compatibility, "high education," regained polarization vitality in the final window with the emergence of the "wheelchair provocation" topic.

The empirical results demonstrate that predetermined topic characteristics play a significant role in emotional polarization. The two viewpoint clusters of "seat occupation" and "apology" lacked sufficient depth in public opinion fermentation, while "plagiarism" and "punishment" lacked breadth in public opinion fermentation. The viewpoint cluster formed around the "high education" topic possessed both depth and breadth in public opinion fermentation, thus receiving sustained attention.

This also indicates that emotional polarization at a single moment may not necessarily cause shockwaves. Therefore, in the online public opinion field focused on in this empirical study, the viewpoint cluster generated by "wheelchair provocation" may produce further emotional polarization, but its final evolution trend still requires extended measurement time series for judgment. However, topics that receive sustained attention are highly likely to cause unforeseen consequences throughout the entire public opinion event process. Sufficient attention must be paid to such topics in both emotional polarization measurement and online public opinion control practice, and the entity "seat hogging man" should also receive corresponding attention because it has stimulated highly consistent emotional tendencies among public opinion audiences in the field. If no measures are taken, it may interact with other homogeneous events, thereby causing collective emotional shockwaves among netizens again.

The method proposed in this study for measuring the emotional polarization degree of viewpoint clusters in online public opinion fields essentially comprehensively considers the emotional coherence performance of various polarized viewpoint clusters in the field from three perspectives. By calculating polarization indicators within the time series, an overall emotional polarization value for a viewpoint cluster is ultimately obtained, providing a quantitative basis

for intervention governance of emotional polarization among online public opinion audiences. However, due to time and capability limitations, the research method proposed in this study still has certain limitations: measurement must rely on a continuous time series, and the overall judgment of viewpoint cluster polarization degree cannot be completed through polarization indicators at a single time node. It is believed that with further in-depth research, the above problems will be properly resolved.

References

[1] Gao Junfeng, Song Shaocheng. Research on the Measurement of Audience Cognitive Distance of Viewpoints in Online Public Opinion Information [J]. Library and Information Service, 2016, 60(20): 77-85.

[2] Le Guoan, Xue Ting, Chen Hao. A Preliminary Exploration of the Definition and Classification Framework of Online Collective Behavior [J]. Journal of Chinese People's Public Security University (Social Science Edition), 2010(6): 99-106.

[3] McGarty C, Bliuc A M, Thomas E F, et al. Collective action as the material expression of opinion-based group membership [J]. Journal of Social Issues, 2010, 65(4): 839-857.

[4] Yzerbyt V Y, Dumont M, Wigboldus D, et al. I feel for us: The impact of categorization and identification on emotions and action tendencies [J]. British Journal of Social Psychology, 2011, 42(4): 533-549.

[5] Bkowicz P, Sobkowicz A. Dynamics of hate-based networks [J]. The European Physical Journal B, 2010, 83(4): 633-643.

[6] Shi Mi, Liu Jianzhun. Online Collective Behavior Intention: Concept, Measurement and Formation Elements [J]. Journal of Intelligence, 2017, 36(5): 101-105.

[7] Du Yangqin, Huo Youguang, Suo Zhihai. Measurement of Online Group Polarization Based on Quantile Regression [J]. Library and Information Service, 2011, 55(24): 38-43.

[8] Zhang Runlian, Lan Yuexin, Wang Caihua, et al. Evolutionary Game Analysis and Countermeasure Research on Online Group Events [J]. Library and Information, 2016(4): 24-30.

[9] Huang Wei, Song Xianzhi, Gao Junfeng. Measurement and Empirical Research on Connection Robustness of Audience Viewpoint Communities in Online Public Opinion Field [J]. Journal of the China Society for Scientific and Technical Information, 2017, 36(5): 503-511.

[10] Gao Junfeng. Research on Cohesion Robustness Measurement of Audience Viewpoint Communities in Online Public Opinion Field [J]. Journal of Intelligence, 2018, 37(4): 106-113.

[11] Wu Shixian, Zhang Bilan. Public Opinion Polarization Prediction Model Based on Viewpoint Potential Field [J]. Library and Information Service, 2015, 59(19): 108-112.

[12] Sunstein C R. The Republic of Networks: Democratic Issues in Network Society [M]. Translated by Huang Weiming. Shanghai: Shanghai People's Publishing House, 2008.

[13] Chen H M. Group polarization in virtual communities: The case of stock message boards [J]. I-schools, 2013, 4(2): 185-195.

[14] Zha Xianjin. Information Analysis [M]. Wuhan: Wuhan University Press, 2014.

[15] Zhang H P, Yu H K, Xiong D Y, et al. HHMM-based Chinese lexical analyzer ICTCLAS [C]//Second SIGHAN Workshop Affiliated with 41st ACL. Stroudsburg: Association for Computational Linguistics, 2003: 184-187.

[16] Xu Linhong, Lin Hongfei, Pan Yu. Construction of Emotional Vocabulary Ontology [J]. Journal of the China Society for Scientific and Technical Information, 2008, 27(2): 180-185.

[17] Gao Ge, Luo Jingmei, Wang Yu. Text Sentiment Orientation Analysis Based on HNC Theory [J]. Data Analysis and Knowledge Discovery, 2017, 1(8): 85-91.

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