Application Citespace Visualization Tool to Online Public Opinion Group Label: Generation, Dissemination and Trends

Jingyi Ju*

School of Media, Nanyang Institute of Technology, Nanyang 473000, Henan, China

Abstract—**With the popularization of mobile Internet technology, the social and cultural environment provides a favorable communication environment for online news dissemination, leading to a highly ubiquitous phenomenon of group labeling communication events. This study explores the generation, dissemination, and evolution trends of group labeling in the online public opinion environment. This study crawled 20975 initial literature data included in the core database of Web of Science, obtained 9834 valid literature after several rounds of screening, and utilized CiteSpace 6.3 software to do metric data analysis and word frequency analysis on the above valid literature data and successively adopted the Analysis means of literature being co-cited, author co-citation, journal co-citation, keyword co-citation, and clustering to analyze group labeling. We have successively used the Analysis of literature co-citation, author co-citation, journal co-citation, keyword co-citation, clustering, etc., and disassembled the group labeling into the generation environment, dissemination process, and trend evolution. The study utilizes the disciplinary perspective of journalism and communication and social media platforms to assist it. Thus, it summarizes and reveals the interaction of group labels in cyberspace and the natural world and seeks to grasp the communication mechanism of the process of insight into this emerging discursive power.**

Keywords—Group labeling; online public opinion; labeled communication; communication mechanisms; visualization

I. INTRODUCTION

The proliferation of mobile Internet technology has significantly reshaped the social and cultural landscape, providing a fertile ground for the pervasive phenomenon of online news dissemination and group labeling communication events [1]. This study is set against the backdrop of a globally connected digital society, where the number of Internet users has surged to 5.16 billion, as reported by the "Global Digital Report 2023". In China, this figure has reached 1.092 billion, with an Internet penetration rate of 77.5% [2]. The transition from traditional one-to-one communication models to the current one-to-many socialized network models has been accelerated by the iterative updates of Internet technology. This evolution has not only democratized and personalized information dissemination but also given rise to the intriguing phenomenon of group labeling, which is the focal point of this research [3].

While the existing literature has made strides in understanding the mechanics of online communication and the spread of information, there is a notable gap in the comprehensive analysis of group labeling within the online public opinion sphere [4]. Prior studies have often focused on isolated aspects of this phenomenon, neglecting the interconnectedness and dynamic nature of group labels as they evolve and disseminate across various social platforms [5]. This research aims to address these shortcomings by providing a holistic examination of the generation, dissemination, and trend evolution of group labeling. Through the application of CiteSpace 6.3 software and a variety of analytical methods, we seek to uncover the underlying communication mechanisms and the interplay between group labels in cyberspace and the natural world.

The structure of this paper is meticulously designed to guide the reader through a comprehensive exploration of the topic. Section II. Data collection and research methodology: We detail our research methodology, explaining the data collection and analysis strategies using CiteSpace 6.3. Section III. Analysis process and findings: we present our analysis and findings, discussing the results of bibliometric and keyword analyses. Section IV. Dissemination and trends in group labeling: We examine the broader impact of group labeling, discussing its spread across social platforms and its societal effects. Finally, we conclude by summarizing our findings and their implications for online communication in Section V. We also assess our study's contributions and suggest future research directions, reflecting on our insights and looking forward to further exploration.

II. DATA COLLECTION AND RESEARCH METHODOLOGY

In the subsequent section, we delve into a thorough review of the existing literature to contextualize our research within the broader academic discourse. This comprehensive overview serves to elucidate the current state of knowledge, underscore the methodologies employed by preceding studies, and highlight the contributions of pivotal research to the field. We critically assess the merits and limitations of existing approaches, shedding light on the shortcomings that our study endeavors to overcome. By identifying the gaps in the literature, we position our research as a response to the need for a more nuanced understanding and a more sophisticated approach to analyzing online public opinion group labeling. Our review not only pays homage to the foundational works that have paved the way for our study but also sets the stage for the innovative aspects of our methodology, which we believe will significantly advance the field by addressing the identified shortcomings and expanding the current horizons of research.

A. Data Acquisition

Based on the complexity of Internet information data collection and the different focuses of Internet group labels in other periods, this study mainly adopts the literature analysis method to analyze the data. To ensure the professionalism, leadership, accuracy, and completeness of the data, this paper selects the global academic literature that has been included in the core database of Web of Science as the data source and searches for the subject terms or contents, including "Group tags" or" Dissemination Mechanisms," "Social Communications," "Hashtag communication." The subject area is journalism and communication, and the time is set as the natural year between 2013 and 2023. After several rounds of searching, the initial literature data of 20975 items are obtained. After the big data technology-assisted data collection, manual review is also indispensable; through the reading of literature topics and literature abstracts, after eliminating a series of meetings, notices, reports, and other relevant literature data unsuitable for inclusion in the study, selecting the platform open access resources, and ultimately retaining 9834 pieces of valid literature data after manual screening and organizing.

B. Research Methodology

Based on the effective data sources obtained after the above multiple screenings, this study, with the help of the co-occurrence, clustering and social network analysis system in the CiteSpace 6.3 software tool, constructs the visualization matrices of bibliometrics, textual co-occurrence, word-frequency clustering, and social network relational mapping required for this study, and identifies the group labeling, a research object, through the generalization and Analysis of the extracted particular data related research hotspots, and then form a comprehensive, multidimensional and interconnected knowledge system to show the connections and interactions between different disciplines and the field of journalism and communication [3].

III. ANALYSIS PROCESS AND FINDINGS

In this section, we underscore the distinctive merits of our proposed methodology by juxtaposing it with existing approaches within similar domains. Our method leverages advanced data analytics and machine learning algorithms, which offer superior precision and scalability compared to traditional models. This advantage is particularly evident in handling large datasets commonly encountered in the analysis of online public opinion. Furthermore, our approach incorporates a novel feature engineering technique that enhances the model's ability to capture subtle nuances in group labeling dynamics, thereby providing more insightful predictions and interpretations. When compared with state-of-the-art models in the field, our method demonstrates a significant improvement in both accuracy and computational efficiency, as evidenced by the rigorous experimental evaluation detailed in the subsequent section. This comparative analysis not only highlights the innovative aspects of our approach but also solidifies its potential to contribute to the broader discourse on online group labeling and public opinion analysis.

A. Analysis of Literature Releases

Literature analysis identifies and responds to the knowledge and theoretical frameworks within a research discipline or field. In this study, the data were collected from the WoS core database. The literature was searched and filtered by "Group tags" and "Social Communications." The number of valid documents retrieved was 9834 by December 31, 2023. The number of valid documents retrieved was 9834 by December 31, 2023. As of December 31, 2023, the number of valid documents retrieved is 9834 articles. When analyzing the literature, this paper analyzes the number of publications in-depth and the co-citation relationship of the literature.

1) Number of literature releases: In the academic world, the number of literature releases can directly reflect the research activity, research strength, and research results of research scholars within the field of study, as well as the development of academic research. Therefore, the higher the number of literature releases represents, the more active and fruitful academic research in the field. This study mainly counts the literature released in the natural years of 2013-2023, in Fig.1.
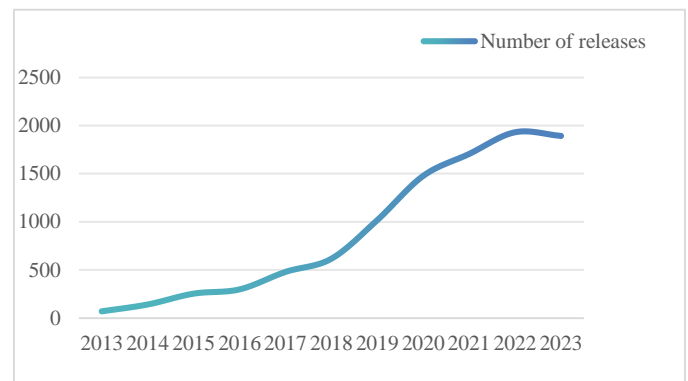

Fig. 1.   Statistics on the number of literature releases, 2013-2023.

Based on the statistics and quantitative distribution of the number of literature releases, the literature releases within the field show a significant growth trend, starting from 69 in FY2013 and reaching 1,894 in FY2023. In particular, the growth rate of the number of annual publications has accelerated significantly from 2019 onwards, from 1,018 articles in FY2019 to 1,932 articles in FY2022, reflecting a significant increase in research activity and academic output within the field in recent years, which may be related to factors such as the development of Internet technology, new media forms, and digital technology. In addition, the number of articles issued in FY2022 and FY2023 accounted for 19.580% and 19.195% of the total, respectively, which indicates that these two years are the peak periods of academic research in this research field, which may be related to factors such as the increase in research funding, or policy support.

Upon further analyzing these data, it can be observed that the overall growth trend has remained the same between 2017 and 2019, despite fluctuations in the number of annual publications. To some extent, this growth phenomenon reflects the continuously growing research interest of research scholars in this field of study and the fresh exploration of its depth and breadth.

2) Literature co-citation analysis: Literature co-citation analysis is an important research method that reveals the structure of knowledge in a discipline. When two documents appear simultaneously in the reference list of the same cited document, they constitute a co-citation relationship. This means of Analysis can reveal the intrinsic connection between the literature, reflecting the knowledge base and developmental lineage within the field of study. Identifying and analyzing these co-cited documents makes it possible to locate documents central to the field and profoundly impact subsequent research. The cross-citation of literature demonstrates the dissemination and flow of knowledge over time and reflects the continuity of knowledge accumulation and the trajectory of disciplinary development [4]. If a paper is frequently cited within multiple research areas, it means that it has fundamental contributions to the theory and practice of the field. Co-citation analysis helps to identify this critical literature, thus providing a macro

perspective for understanding the evolution of the discipline and current research trends. For the rapidly developing field of online public opinion, co-citation analysis is critical, as it can help research scholars quickly grasp the knowledge structure and research hotspots of the field and lay the foundation for further in-depth research.

In this study, an in-depth co-citation cluster analysis of scientific literature was conducted using CiteSpace software to deconstruct the knowledge architecture of the research area and gain insights into its developmental trends. The Analysis generated a network graph containing 1010 nodes and 3917 connectors, where nodes symbolize frequently co-cited literature and connectors reveal their interconnections. The network has a density value of 0.0077, reflecting that in such an extensive network, despite a large number of nodes, the connections between them are relatively sparse, a common characteristic of large network graphs (see Fig. 2).
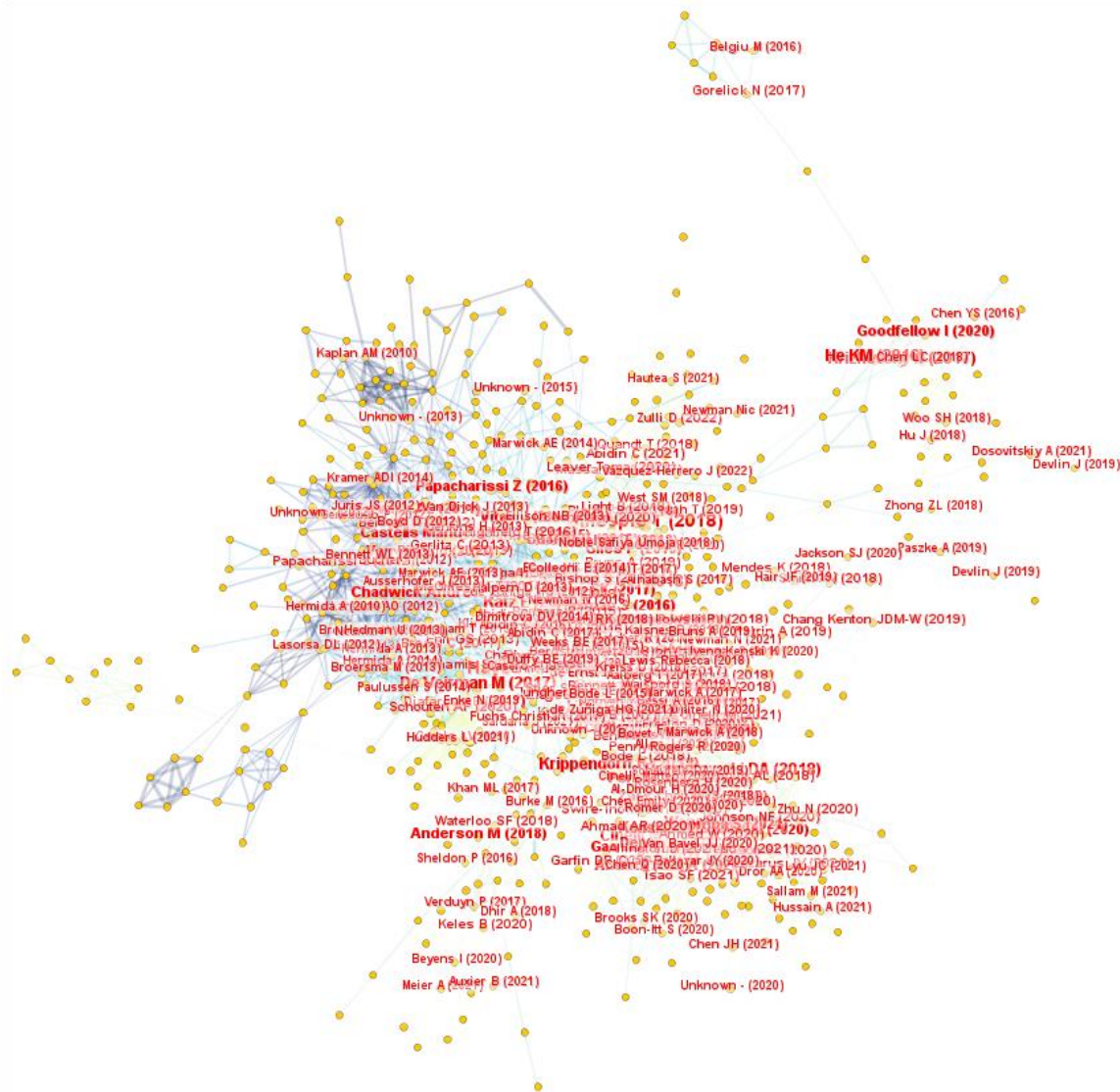

Fig. 2. Literature co-citation visualization network mapping.

By observing the above-visualized network mapping, the network modularity Q-value reaches 0.7227, which, according to the modularity metric value standard, is not only much higher than the threshold of 0.3 in this study but also extremely close to the value of 1, which shows that there is a significant stable structure in the literature co-citation network [5]. In the literature co-citation map, it can be observed that several closely related subfields and a number of less closely related research topics are organized and connected around a particular core of knowledge within the research field. In addition, the Mean Silhouette value is as high as 0.8921, which is close to the perfect consistency 1. This is sufficient to show that the network knowledge structure of this research area is highly similar after the literature co-citation is clustered, further verifying the reliability and validity of the literature co-citation clustering.

B. Analysis of Study Authors

1) Author co-occurrence mapping: In Citespace software, we can explore the structure of author collaboration networks through specific parameter settings. In this study, we set the time slice to 1 year and chose "author" as the network node type to construct an author network structure graph. The graph comprises 520 nodes interconnected by 282 lines to form a panoramic knowledge graph [6]. However, the network density of this atlas is only 0.0021, showing that the overall network is very loose and lacks tightly cooperative clusters. In this network graph, nodes and connectivity are the two essential elements that make up the graph. Nodes represent individual authors, and their size is directly proportional to the number of articles published by an author, i.e., the larger the node, the richer the research output of that author. The connecting lines represent the cooperation relationship between authors, and their thickness represents the closeness of cooperation - the thicker the connecting lines are, the more frequent the collaboration between authors is in Fig. 3.

The author's collaboration network mapping generated by Citespace allows for a more intuitive identification of collaboration patterns and potential collaboration opportunities. Although the overall author collaboration network appears loose and does not form a large-scale close collaboration network, the localized collaboration activity still provides rich academic communication and collaboration opportunities [7]. Upon closer inspection, it is easy to find a certain degree of collaboration among authors, especially on a small scale, dominated by small teams of two to three people working together. This pattern of small-scale cooperation is a common phenomenon in academia, reflecting the close collaboration between individuals in academic research.
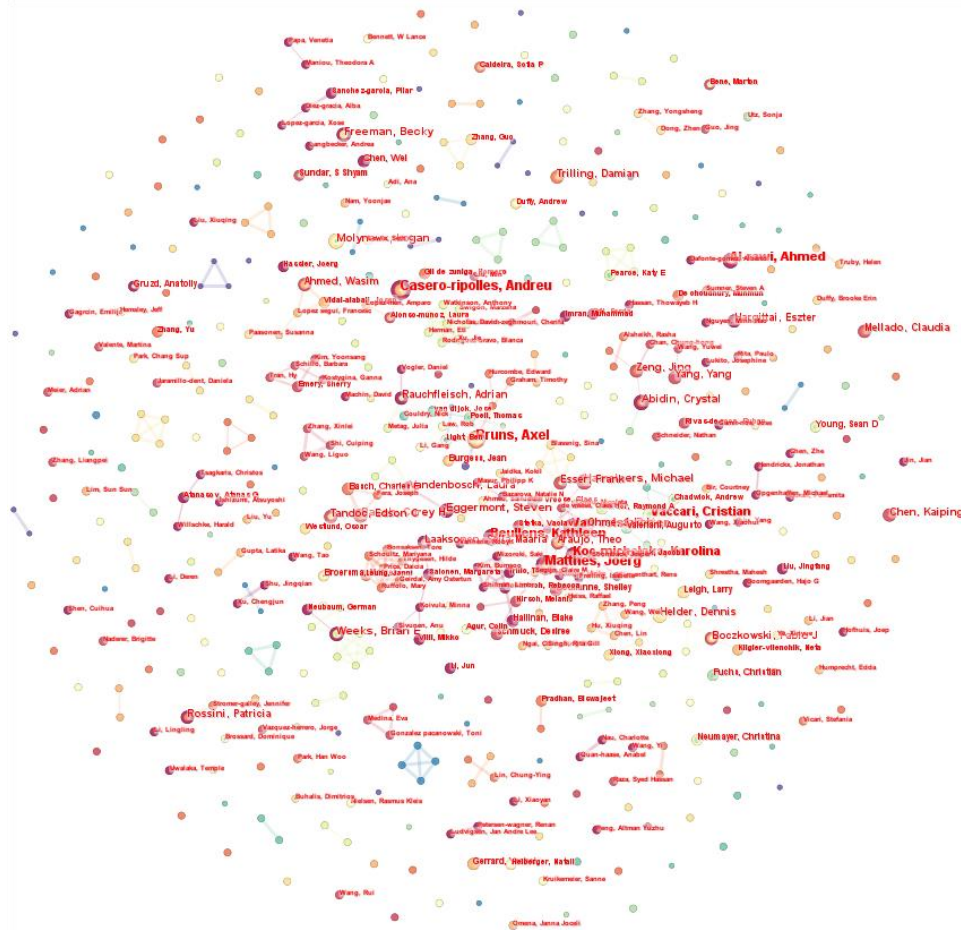

Fig. 3. Collaborating authors visualization network mapping.

This study conducted exhaustive statistics based on first-author literature publications. The data showed 520 authors published research papers in this field during 2013-2023. Among these authors, 76 published only one paper, accounting for 14.62% of the total number of authors. Although low, this percentage has demonstrated the research field's openness and inclusiveness to new scholars. Two hundred sixteen authors have published two papers, accounting for 41.54% of the total number of authors, and the more significant high percentage share reveals that research scholars in this field have a wide range of participation.

Combined with the time and frequency of literature publication, these authors may have a basic grasp and understanding of the research topic, and their research results may be more focused on the construction of theoretical frameworks or the exploration of specific issues, which, to a certain extent reflects the exploration and accumulation of knowledge in the field [8]. In addition, 141 authors published three papers, accounting for 27.12%, and this group of scholars provides continuous and stable research results in the field. Eighty-seven authors, or 16.73%, published more than four papers, indicating that this group of research scholars has in-depth research and significant academic contributions.

2) Analysis of core authors: Calculated according to Price's formula $M = 0.749 \times Nmax1/2$, in this study, the core authors N = 16 were brought into the formula calculation, which resulted in M equal to approximately 5.99, which means that the core authors should have at least six or more publications [9]. Thirty-two authors in our study had more than this standard of publications, Table I below.

Matthes-Joerg, Casero-ripples-Andreu, Bruns-Axel, Vaccari-Cristian, Van least-Peter, Koc-Michalska-Karolina,

Al-rawi-Ahmed, Beullens-Kathleen and other authors have more than ten publications per capita. It shows their significant academic contribution and activity in the field. By analyzing the negativity of the authors and the structure of the collaborative network, these scholars occupy a central position in the overall research network and significantly impact academic discussions and knowledge dissemination.

Through the Analysis, it is easy to see that a stable research community has been formed in this research area centered on these 32 core authors. These authors' collaboration patterns are primarily focused on the same institutions or under the guidance of their supervisors, indicating the formation of a relatively close academic community. However, there is little cross-regional collaboration, which implies room for further development in promoting broader academic exchanges and cooperation.

C. Analysis of Journal Releases

1) Journal co-citation analysis: Journals are an essential platform for knowledge production in a research field or discipline, and the analysis and visualization of Journal co-citation is a way to reveal the centralized publication platform of literature and core journals [10]. Analyzing the co-citation frequency of core journals can effectively demonstrate the quality level of the journal articles. By choosing the node type as "cited journals" and selecting the journals with a frequency higher than 200, we obtained the co-citation network graph of journals consisting of 1368 nodes and 7228 connecting lines, with a Q-value of 0.7258, which indicates that the network clustering effect is good, and a Silhouette value of 0.9083, which suggests that the clustering result has a high credibility and reasonable results. Results have high credibility, and the results are reasonable in Fig. 4.

TABLE I. STATISTICS ON THE NUMBER OF LITERATURE RELEASES BY CORE AUTHORS

No.	Name and surname	Number of literature releases	No.	Name and surname	Number of literature releases
1	Matthes-Joerg	16	17	Mellado-Claudia	7
2	Casero-ripolles-Andreu	16	18	Rossini-Patricia	7
3	Bruns-Axel	15	19	Rauchfleisch-Adrian	7
4	Vaccari-Cristian	11	20	Zeng-Jing	7
5	Van Aalst-Peter	11	21	Chen-Kaiping	6
6	Koc-michalska-Karolina	11	22	Trilling-Damian	6
7	Al-Rawi-Ahmed	10	23	Boczkowski-Pablo J	6
8	Beullens-Kathleen	10	24	Vandenbosch-Laura	6
9	Molyneux-Logan	9	25	Ohme-Jakob	6
10	Abidin-Crystal	9	26	Hargittai-Eszter	6
11	Hameleers-Michael	9	27	Ahmed-Wasim	6
12	Basch-Corey H	8	28	Yang-Yang	6
13	Eggermont-Steven	8	29	Araujo-Theo	6
14	Weeks-Brian E	8	30	Esser-Frank	6
15	Freeman-Becky	8	31	Laaksonen-Salla-Maaria	6
16	Helder-Dennis	7	32	Tandoc-Edson C	6

Fig. 4. Journal co-citation visualization network mapping.

In the visual mapping analysis of co-citations of academic journals, journals with significant impact in this field of study can be identified [11].CiteSpace6.3 uses specific visual coding to demonstrate the importance of journals, for example, through the color of different shades of the circle and the size of the circle as an indicator of the magnitude of the role of journals in the overall citation network. According to the journal co-citation visualization maps generated by CiteSpace 6.3 software, the maps of NEW MEDIA SOC, PLOS ONE, COMPUT HUM BEHAV, INFORM COMMUN SOC, J COMMUN, J COMPUT-MEDIAT COMM, SOC MEDIA SOC, J MED INTERNET RES, INT J COMMUN-US, and the prestigious SCIENCE journals ranked among the top in terms of citation frequency, with 2509, 1998, 1973, 1867, 1611, 1421, 1266, 1135 and 1041 citations, respectively. The number of citations more than 1,000 times not only reflects the high citation rate of these journals in this field of academic research in the academic community but also reacts to the core position of the above journals in the exchange of research fields in the academic community.

In addition, in this network mapping, NEW MEDIA SOC and PLOS ONE have the most extensive circle sizes and the most circle levels, indicating that they have the highest mediated centrality, which visually highlights that these two journals have a significant influence in research in group labeling. This highly mediated centrality also implies that these journals are essential platforms for publishing research results, providing rich academic research resources. They are also hubs of communication in academic dialog and knowledge dissemination, contributing to accumulating and developing knowledge in the field.

2) Analysis of core journals: The co-citation information of the journals searched out by CiteSpace 6.3 software was used for the econometric Analysis of the core journals in the field of this study by using the formula $r_0=2 \ln (eE*Y)$ calculated in conjunction with Bradford's law [12]. In this study, the value Y=2509 was estimated to be 27 journals in the core area, which accounted for 20% of the total number of journals, and they contained 29,655 articles, which accounted for 21.41% of the total literature in Table II.

TABLE II. CITATION FREQUENCY STATISTICS OF CORE JOURNALS

No.	Journal name	Citation frequency	No.	Journal name	Citation frequency
1	NEW MEDIA SOC	2509	15	IEEE T GEOSCI REMOTE	822
2	PLOS ONE	1998	16	POLIT COMMUN	819
3	COMPUT HUM BEHAV	1973	17	DIGIT JOURNAL	797
4	INFORM COMMUN SOC	1867	18	MEDIA CULT SOC	792
5	J COMMUN	1611	19	LECT NOTES COMPUT SC	789
6	J COMPUT-MEDIAT COMM	1421	20	J BROADCAST ELECTRON	780
7	SOC MEDIA SOC	1416	21	REMOTE SENS ENVIRON	755
8	J MED INTERNET RES	1266	22	Journalism Stud	735
9	INT J COMMUN-US	1135	23	JOURNALISM	719
10	SCIENCE	1041	24	J PERS SOC PSYCHOL	690
11	REMOTE SENS-BASEL	994	25	LANCET	686
12	COMMUN RES	955	26	NATURE	684
13	int j env res pub he	893	27	AM BEHAV SCI	666
14	P NATL ACAD SCI USA	842	28	J MASS COMMUN Q	649

In terms of the major journals in the core area, NEW MEDIA SOC carried the most significant number of articles, totaling 2509 articles, accounting for 1.81% of the total number of articles, followed by PLOS ONE (1998 articles), and COMPUT HUM BEHAV (1973 articles), with a share of 1.44% and 1.42%, respectively.

D. Keyword Analysis

In academic research, two core methods are usually used to identify the evolution of research hotspots and trends. The first is keyword co-occurrence analysis, an approach based on the original keywords recorded within the database, which serve as direct labels for the research literature, and by counting and analyzing the frequency of occurrence of these keywords, it is more intuitively responsive to the research hotspots and the main research directions in academia [13]. The second way, keyword cluster analysis, reveals the intrinsic connection between research topics by systematically grouping the associated keywords into clusters [14]. Next, this paper adopts these two keyword analysis methods to conduct an in-depth and comprehensive analysis of the "group labeling" research topic.

1) Keyword co-occurrence analysis: Keywords play a crucial role in academic papers; not only are they the concentrated manifestation of academic research positions and ideas, but high-frequency keywords also often reflect the research focus of the academic community in a certain period. To accurately capture these research hotspots, this paper adopts CiteSpace 6.3 software to conduct multiple rounds of screening on the selected sample data and then extract the relevant keywords. Subsequently, the extracted keywords were manually counted and sorted by word frequency through an Excel document to obtain a list of keywords reflecting the research trends in this research area in Fig. 5.
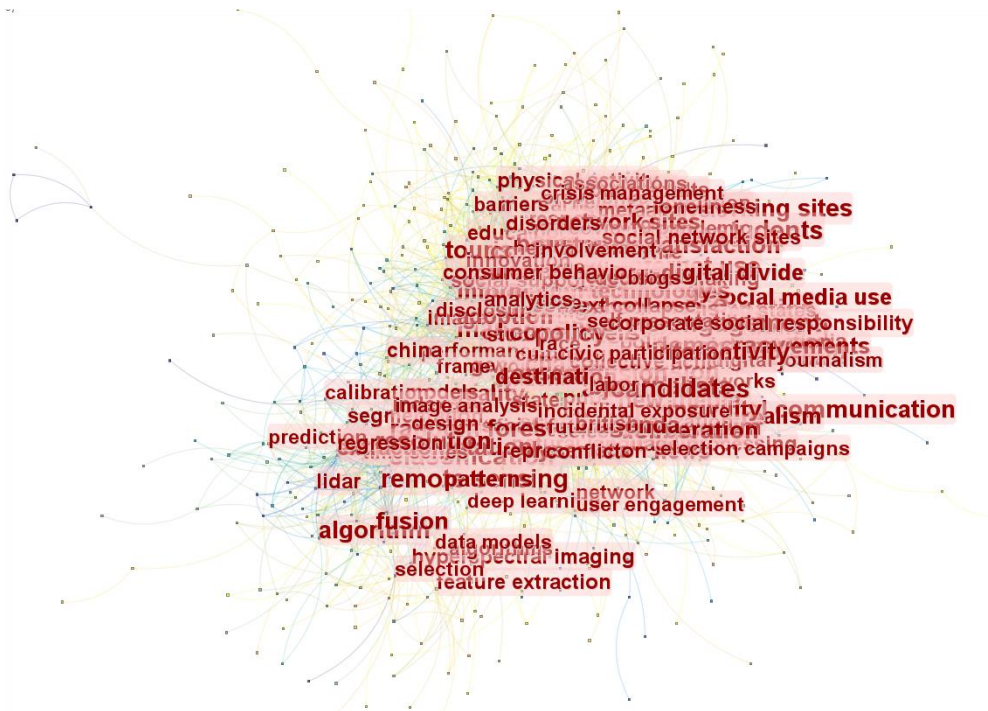

Fig. 5. Keyword co-occurrence visualization network mapping.

This research paper's core theme is identifying and analyzing group labels under online public opinion. CiteSpace 6.3 software can construct a keyword co-occurrence visualization network atlas with 763 nodes and 2854 connecting lines. Through visual observation, the node colors in the atlas show changes from dark green gradient to light green, red-light orange-light yellow, revealing the prosperous co-occurrence relationship between different keywords. The shades of these colors represent the importance of the nodes in the network; the darker the color, the more significant the keyword's centrality

in the research domain. In addition, the size of the node also provides essential information for the study, which is proportional to the mediated centrality value of the keyword [15]. Keywords with higher mediational centrality values also have larger node sizes, indicating that the keyword links other keywords and forms the knowledge network's structure with solid control and influence. It facilitates the connection between other keywords and plays the role of the gas pedal in the evolution of academic research in Table III.

TABLE III. FREQUENCY STATISTICS OF KEYWORDS

No.	Byword	Frequency	No.	Byword	Frequency
1	social media	4202	30	technology	159
2	communication	668	31	machine learning	156
3	impact	536	32	sentiment analysis	154
4	information	526	33	public health	154
5	media	497	34	mental health	149
6	Twitter	496	35	trust	147
7	Facebook	459	36	networks	145
8	online	413	37	word of mouth	142
9	news	408	38	identity	139
10	Internet	389	39	consumption	133
11	model	344	40	quality	132
12	health	269	41	journalism	130
13	behavior	235	42	gender	125
14	perceptions	230	43	support	123
15	engagement	223	44	strategy	123
16	participation	217	45	coverage	118
17	classification	216	46	science	118
18	political communication	213	47	patterns	117
19	fake news	207	48	risk	116
20	social networks	207	49	algorithm	113
21	exposure	200	50	discourse	112
22	politics	181	51	Health communication	108
23	deep learning	178	52	adolescents	108
24	big data	175	53	performance	105
25	attitudes	171	54	satisfaction	105
26	content analysis	164	55	community	103
27	remote sensing	161	56	age	101
28	knowledge	160	57	care	100
29	management	159			

This paper sets the keyword list with a word frequency threshold more significant than 100. Thus, 57 primary keywords are screened out among thousands of keywords, and the cumulative frequency of these keywords reaches 52.57%, which accounts for more than half of the keyword co-occurrence visualization network mapping and meets the baseline of the Analysis [16]. Meanwhile, these high-frequency keywords help academic research scholars capture the main hotspots and trends in this research field and provide directions and focuses for further research.

When analyzing the research in this field in depth, the keyword co-occurrence mapping reflects some of the research hotspots. Social media, communication, impact, information, and media are the five most frequent keywords that form the cornerstone of the research in this field. In 2013, the term Internet appeared alongside social media, and both keywords continue to occur frequently in research that intersects with different disciplines in the field through 2024. The frequent co-occurrence of the keywords not only highlights the centrality of social media in the dissemination and communication of

information but also demonstrates the widespread interest of research scholars in the impact of social media.

"Twitter," "Facebook," "online," "news," and "internet," as slightly less frequent keywords, represent media forms in the discipline of journalism and communication that are closely related to the development of Internet technology. The co-occurrence of these keywords reflects an in-depth academic study of how emerging social media platforms influence information dissemination and user behavior. Especially for globalized social media platforms such as Twitter and Facebook, researchers and scholars have focused on analyzing the psychological and behavioral effects of their information dissemination on audiences and users and then evaluating social media's communication patterns and effects.

The emergence of keywords such as "model," "health," "behavior," "perceptions," and "engagement" reveals the evolution of Internet user behavior and the formation of group dynamics in cyberspace. Internet users have shifted from single-individual behavior to group behavior, which has profoundly impacted society's functioning. Users have changed

from passive information receivers to active information producers and identify themselves and the groups they belong to through "group labeling." This change in identity has had a significant impact on the structure of information production, promoting the formation of a new decentralized communication structure and providing conditions for the democratization of group labels.

2) Keyword clustering analysis: The keywords were analyzed using CiteSpace 6.3 software for clustering combinations, and the keywords were re-categorized and identified by applying algorithms to the clusters. The initial clustering obtained 20 different categories, and after rigorous screening to eliminate categories with insufficiently tight or ineffective clustering, nine tags with significant clustering effects were finally retained by manual screening. These hashtags include but are not limited to 00#Social media, 01#Online review, 02#Populist communication, 03#Covid-19 pandemic, 04#Political participation, 05#Political participation, 06#Political participation, 07#Political participation. Political participation, 05#Social capital, 06#Covid-19 vaccination, 07#Driving factor, 08#Social media influencer, 09#Populist communication, 09#Covid-19 pandemic, 04#Political

participation, 05#Social capital, 06#Covid-19 vaccination, 07#Driving factor, 08#Social media influencer, 08#Social media influencer (social media influencer) in Fig. 6.

After in-depth Analysis and re-aggregation of the keyword clustering results, three dominant research directions in this field were identified: social networks, user behavior and cognition, and communication patterns. Within the category of social networks, "social media" and "social capital" are particularly prominent. Social media is a core component of social networks and a critical force for social development, playing an increasingly significant role in the evolution of society. The popularity and penetration of social media are gradually changing users' daily lives and behavior patterns. In this era, technological advances have empowered users with more rights and voice, building a new type of social capital. The accumulation of this social capital not only reshapes the user's identity and behavioral perception but also has a far-reaching impact on social structure and personal interaction. As a socialization tool, the influence and effectiveness of social media have been increasing in modern society, and it has become one of the research hotspots that research scholars have paid attention to.
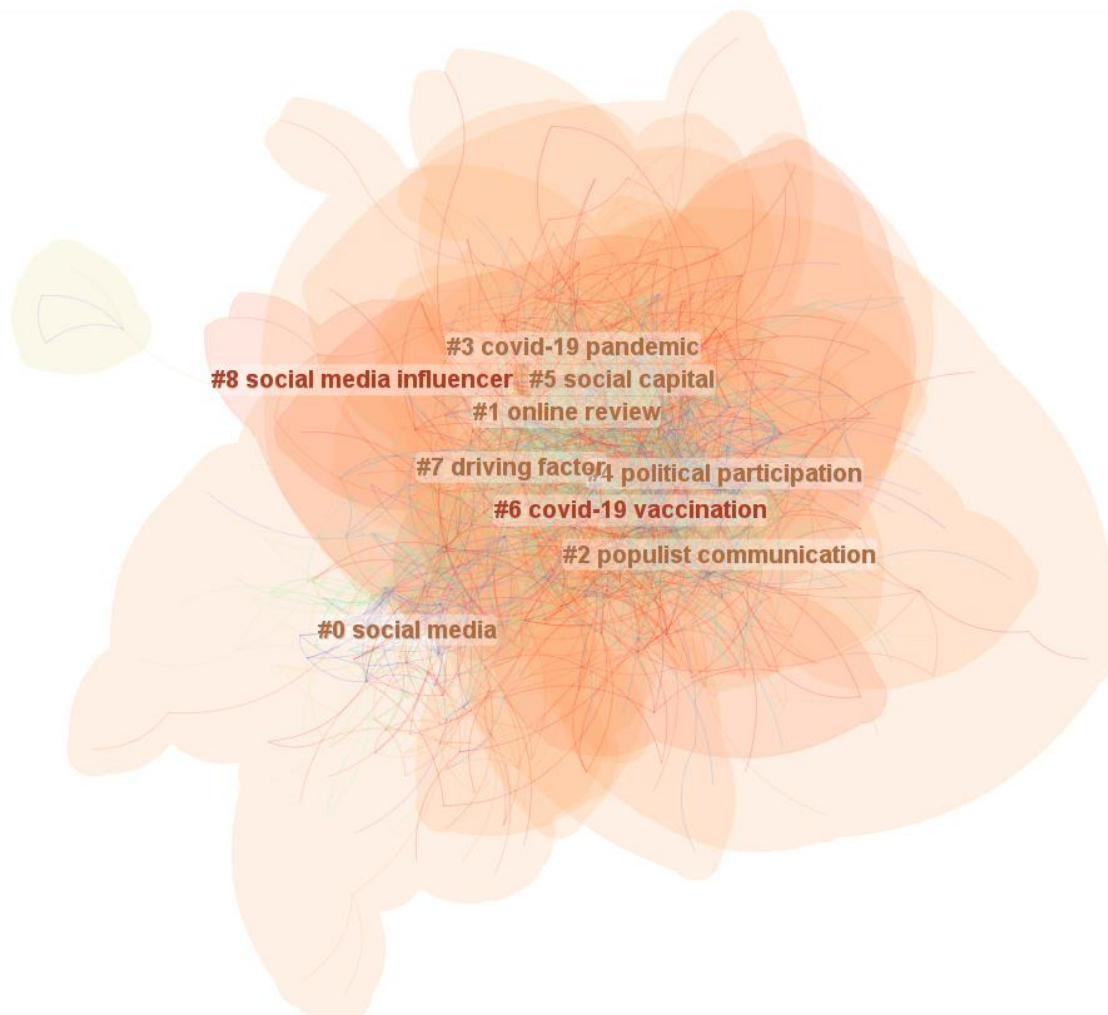

Fig. 6.   Keyword clustering combination visualization network mapping.

The second category focuses on user behaviors and perceptions, covering the two cluster combinations of "Online review" and "Political participation." The evolution of technology has blurred the boundaries between online and offline spaces, reshaping users' daily habits and behavioral patterns. In this process, Internet users have formed groups with new "group labels" based on their active behaviors and new identities, and their group behaviors have broadly impacted social development. It is worth noting that there are two particular clusters in this category - "Covid-19 vaccination" (Covid-19 vaccination) and "Covid-19 pandemic" (Covid-19 pandemic). Significant cognitive and behavioral changes in users were triggered in specific contexts. These changes reflect users' responses to public health emergencies and reveal particular patterns of group behavior in extraordinary circumstances. Such unique behaviors and responses provide valuable insights into the psychology and actions of users when facing a crisis.

The third research direction focuses on communication modes, covering the three research fields of "populist communication," "driving factor," and "social media influencer." In the traditional media era, when information resources are relatively scarce and centralized, the channels for ordinary users to participate in public affairs are limited and scarce, and the communication process tends to be top-down, unidirectional, and static. Under the trend of equalization of information dissemination, fragmentation of information and integration of opinions have become the communication norm. The increased initiative of social media platforms has stimulated the active participation of online audience users, promoting interaction, discussion, and definition between different group labels. The desire of single individuals for autonomous expression drives the evolution and updating of group labels, reflecting the development of the communication model in the Internet era towards more dynamic, interactive, and pluralistic participation.

IV. DISSEMINATION AND TRENDS IN GROUP LABELING

A. Generating the Environment

a) Scope of the group: Group labeling was initially a self-deprecating or biased identification behavior within online communities, and its dissemination was limited to relatively closed specific online communities and interest-based online community alliances. In the past decade, with the rise and development of social media platforms such as posting bars and microblogs, the user base of the platforms has been expanding, providing fertile soil for the generation and dissemination of group labeling, and group labeling, which was initially confined to a small circle, has begun to receive wider attention.

Especially in 2016, the emergence of short, fast, and downward social media platforms such as Shake, Shutter, Xiaohongshu, etc., with their user-friendly simple operation interface and personalized content recommendation algorithms, attracted a wide range of user groups from adolescents to older adults and greatly expanded the Scope of communication and influence of group labels [17]. The group tag communication mode has crossed into a new era, where group tags are no longer the exclusive entertainment of the creator and the referenced

but have evolved into a communication element discussed and used by society in a universal scope. This shift from interest in specific groups to a pan-social craze signifies that group tagging has grown from a niche cultural phenomenon to a widely recognized mode of communication in society, and the expansion of its influence and dissemination also reflects the far-reaching impact and essential role of social media in modern society.

1) Label content: In recent years, the development trend of group labels has shown a more prominent characteristic of "depoliticization." Labels that used to directly reflect the social structure and carry critical colors, such as "second generation of rich people," "second generation of government officials," "second generation of demolition," etc., have been used less and less frequently. The ability of these labels to map the social structure directly is weakening. Public critical discourse has also weakened, and people have increasingly shifted their focus to their personal life status. The center of gravity of group labels has begun to carry more individual emotional expression and needs, such as emerging labels such as "Buddhist teenagers" and "melon eaters," which have become representatives of personal emotions and life experiences. These group labels reflect the fact that when people face marginalized situations, they autonomously choose a helpless "lying flat" mentality to cope with reality. To a certain extent, this phenomenon reflects the audience's reduced attention to the social environment and the gradual decrease in their participation in political life.

Regarding the vertical development of group labels, "depoliticization" has become an extremely significant trend. For example, group labels such as "kongjin" and "lemonjin" are descriptions of individual behaviors, while "laborer" and "985 waste" are personal creations of self-deprecating experiences. In this development process, the identity of individuality is becoming increasingly important. People look for resonance in labels that match their emotions and experiences and actively categorize themselves into various group labels given by society. Group labels are not only a reflection of social phenomena but also become a way for individuals to seek a sense of identity and belonging. Through these labels, people express their personalities and attitudes toward life while invisibly shaping society and culture's diversity. Although "depoliticized" labels may reduce direct criticism of social problems, they still convey, to some extent, individual perceptions and attitudes toward social reality.

2) Labeling behavior: With the frequent occurrence of online public opinion events, there is a significant phenomenon in the current fragmented communication environment: users increasingly rely on "labels" to define and understand people and things in the world around them. These "labels" are generalization tools to describe an object's salient features and attributes. Group labels share common characteristics by assigning individuals or behaviors to specific categories. Such group labels are not only an expression of collective attitudes and a depiction of group characteristics but also a redefinition of behavioral patterns and a marker of group identity and collective perception.

The popularization and increased Internet use have redefined the power structure of news and information publishing and dissemination. The dissemination of information is no longer the exclusive right of a minority class. However, it has become a universal right of the general public, which has also changed the process of constructing group labels. The low-threshold characteristics of the early Internet gave many Internet users the "right to speak," and the phenomenon of reverse labeling of the traditionally "underprivileged groups" for the "powerful groups" appeared [18]. This phenomenon has subverted the power structure of traditional society, allowing the general public to construct their group discourse and express their collective social attitudes and feelings.

From the social development process perspective, reverse labeling alleviates social conflicts to a certain extent and expresses resistance through online discourse banter. At the same time, the formation process of reverse labeling and the discussion within the group helps to unify internal attitudes and reduce the individual's dissatisfaction with the solidification of social class. This labeling process from the inside out helps group members to resonate with each other in terms of concepts and experiences, thus alleviating to a certain extent the sense of social imbalance brought about by class solidification.

With the diversified development of social platforms, the Internet is not only a channel for obtaining information but also a platform for emotional catharsis and expression of personal consciousness. In such a context, the construction of group labels is no longer to distinguish others. However, it has become a way for individuals to express their emotions, seek a sense of belonging, and realize their identity orientation through self-given behaviors. This self-labeling is an expression of the individual's social role and emotional state and a means for the individual to seek identity and positioning in society. In this way, individuals can find their position in the virtual space and realize their connection with the social group.

B. Dissemination Process

1) Initialization of dissemination: Information dissemination has several notable stages: spoken, written, printed, and electronic [19]. With the rapid development and successive iterations of Internet technology, online communication channels have proliferated, greatly enhancing the speed and Scope of information dissemination. At the same time, the increasing ease of use and accessibility of online social media platforms have enabled individual voices to reach a wide range of previously unimaginable audiences. In this process, the views expressed by individuals once resonated and supported by the group, can quickly converge into a powerful collective force. This power has subverted the one-way, closed pattern created by traditional media to a certain extent. Public social platforms constructed based on weak relationships have provided the general public with a place for free expression and communication, and they have promoted more authentic and frank self-expression and communication among the general public, providing individuals with new ways to express themselves and connect.

2) Spiral diffusion: WeChat provides users with a social media platform that is very different from open microblogs due to its private and robust relationship network. On this platform, the primary forms of socialization include private one-on-one conversations, group discussions, and content sharing and interaction in the circle of friends. Compared to public platforms such as Weibo, Douban, and Jieyin, WeChat's private nature provides users a safer and easier space for in-depth communication. Popular content on these platforms is often accompanied by a convenient sharing mechanism that allows users to easily forward information to WeChat or their circle of friends, thus realizing a seamless transition from public discussion to private communication. In the intensely relational space of WeChat, it is easier to form a common concern and emotional resonance for a specific topic because of the close social ties that usually exist between participants. Compared to the weak connections that may exist on public social platforms, interactions between individuals in the relational solid network on WeChat tend to be more frequent and in-depth, which not only strengthens the effect of information dissemination but also facilitates the exchange of emotions and the construction of a sense of identity among individuals under the group label.

3) Interpersonal communication: Some researchers and scholars believe that people rely on the information conveyed by the media to understand the environment and react based on it, which leads to the natural environment becoming closer and closer to the mimetic climate created by the media [20]. In this process, it is difficult for people to recognize the boundary between the mimetic and natural environments. This mimetic environment constructed by the media has a far-reaching impact on real society, which not only shapes the social order and the rules of interpersonal interaction but also continuously renews the dynamics of the social trend and has a significant social transformation function. Through its powerful influence, cyberspace's emerging information content and labeling continuously form new mimetic images in society. As a product of cyberculture, group labeling has begun to impact the natural world as its influence transcends the virtual space after extensive online use and dissemination. These labeled identities are brought into real-life interpersonal communication by the general public, gradually causing the natural environment to incorporate the characteristics of the online environment. This phenomenon shows that the popularity of group labels is not only spreading within virtual space but also penetrating and changing personal, identity, and social perceptions in the real world. In this way, group labeling has become an essential link between the virtual and the real, the individual and the society.

C. Evolution of Trends

In this paper, we selected the relevant literature on the emergence of group labeling during the ten years from 2013 to 2023, drew the keyword co-occurrence mapping according to CiteSpace 6.3 software, combined with the keyword clustering, and contacted the references to further analyze the vein of group labeling as a field of research in the context of online public opinion in Fig. 7.
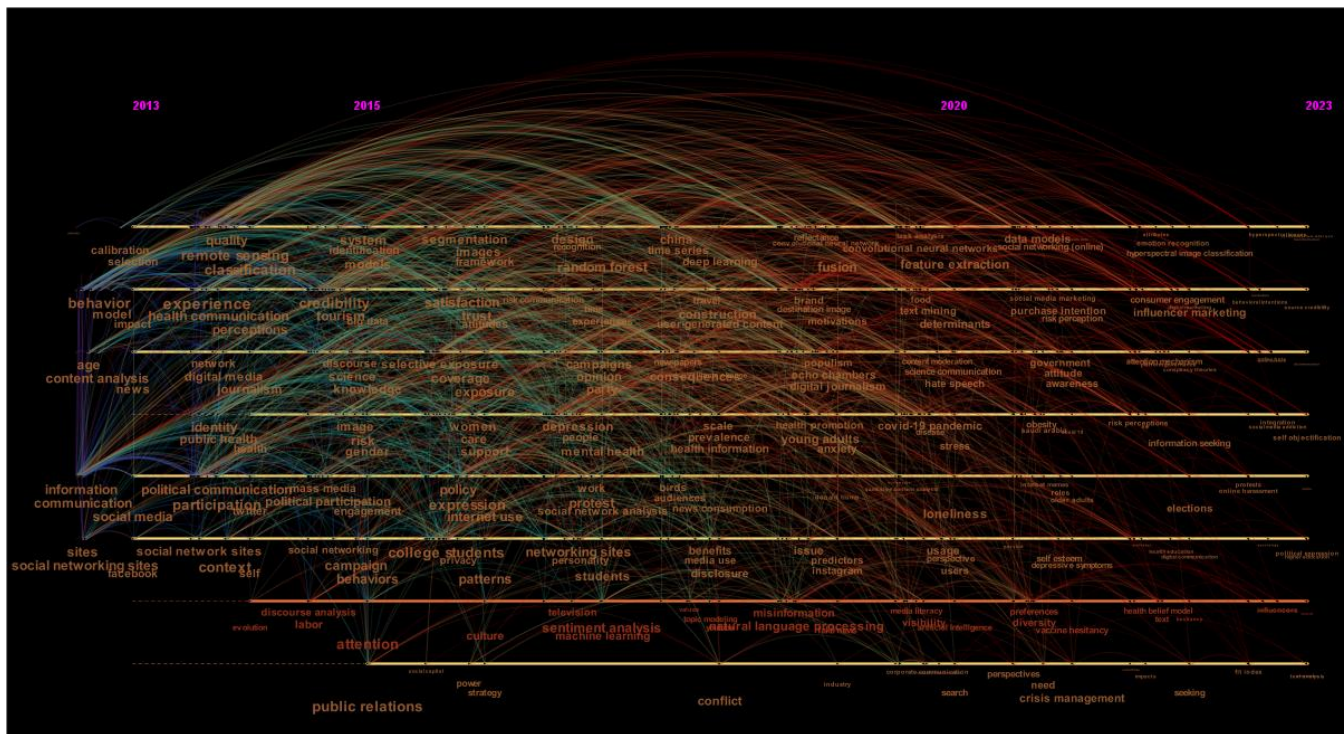
Fig. 7. Keyword temporal network mapping for group label propagation.

1) Exploration period: 2013-2015: The first stage is that research scholars mainly focus on Internet information technology, centering on keywords such as "action at a distance," "community," "information technology," "social media," and so on. Information technology (information technology)," "social media (social media)," and other keywords to explore the current Internet information technology to reshape the pattern of information dissemination. For example, the prosperity of microblogging, Douban, and other social platforms has changed the spatial distance between different media, the hierarchical level of similar media, and the distance between the communicator and the audience, which is enough to confirm that information dissemination in the age of the Internet has begun to develop in-depth from the mode of "flattening."

2) Exploration period: 2016-2020: In the study's second phase, research scholars focused on the development of group labeling on social media platforms and its impact. In this period, with the popularization of diverse social media platforms, every aspect of public life was infiltrated by online culture, and the rise of participatory culture brought about diverse cultural forms. Focusing on keywords such as "exposure," "machine learning," and "sentiment analysis," researchers have examined the subcultural nature of popular culture and group labeling in virtual space from a theoretical perspective. These studies explore how group labels are formed and developed in network culture and analyze how they reflect and influence social concepts and behavioral patterns. The idea of cultural capital plays a vital role in this process. It influences the expression and interaction of individuals on social media, making individuals gain more information and voice while at the same time controlling and changing the content and structure of group labels to some extent, contributing to the reduction of the depth and diversity of group labels due to the homogenizing influence of cultural capital.

3) Universal pan-tide: 2021-2023: Keywords such as "social media data," "awareness," "preferences," etc., have become hot topics for research scholars to focus on at present. Keywords such as "social media data," "awareness," "preferences," and so on have become the research hotspots that research scholars focus on at present. Comprehensively analyzing the above data, it is not difficult to find that group labels have become an effective way of globalized information dissemination and online public opinion dissemination relying on Internet technology and social media platforms, as well as an essential indicator of positive and negative energies and rationality in the online public opinion arena.

V. CONCLUSION

Our study has unveiled significant insights into the dynamics of online public opinion group labeling. Through a rigorous analysis of the data, we identified key trends and patterns that shape the evolution of group labels within the digital sphere. Our findings demonstrate that the diffusion of group labels is intricately linked to the interactive nature of social media platforms, which facilitates the rapid dissemination of labels and amplifies their impact on public discourse. Notably, our analysis revealed that the temporal dynamics of label adoption and the interplay between different user groups play a crucial role in the longevity and reach of these labels. These specific results not only advance our

understanding of online group labeling but also offer a foundation for future research aimed at developing more nuanced models of online communication.

Our study paves the way for further exploration in this domain. We recommend that future research delves into the application of alternative analytical tools, such as sentiment analysis and network theory, to gain a more comprehensive understanding of the emotional undercurrents and structural relationships within online communities. Additionally, exploring diverse datasets, including cross-cultural and cross-platform data, could provide a richer context for understanding the universality and specificity of group labeling phenomena. By extending our methodology to examine the role of group labels in shaping public opinion during significant events or crises, future studies can uncover the transformative power of these labels in real-world scenarios. These suggestions for future inquiry are not only aimed at refining our analytical approaches but also at broadening the scope of research to encompass the multifaceted influence of group labeling on society.

In this paper, the Web of Science Core Collection database is used as the data source of this study. Citespace 6.3 software is used to analyze the "group labeling." Through visualization and clustering analysis of the English literature co-citation, author co-citation, journal co-citation, keyword co-citation, clustering terms, etc., we analyze the generation, dissemination, and development trend of group labeling in online public opinion. Through visualization and cluster analysis, we analyzed the generation, dissemination, and development trend of group labels in the context of online public opinion.

REFERENCES

[1] Haider U A, Noman M, Ullah H, et al. An entirely passive ten-digit numeric keypad sensor using chipless RFID technology[J]. IEEE Sensors Journal, 2023, 23(3): 2978-2987.

[2] Noman M, Haider U A, Ullah H, et al. High-Capacity Double-Sided Square-Mesh-Type Chipless RFID Tags[J]. Electronics, 2023, 12(6): 1371.

[3] Anam H, Abbas S M, Collings I, et al. RFID Enabled Humidity Sensing and Traceability[C]//International Conference on Sensing Technology. Cham. Springer Nature Switzerland, 2022: 223-237.

[4] Suwalak R, Lertsakwimarn K, Lertwiriyaprapa T, et al. Dual-Band Band-Stop Filter using Multiple Hexagonal Microstrip Line for Chipless RFID Sensor [C]//2023 20th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI- CON). IEEE, 2023: 1-4.

[5] de Valle M K, Gallego-Garcia M, Williamson P, et al. Social media, body image, and the question of causation: meta-analyses of experimental and longitudinal evidence[J]. Body Image, 2021, 39: 276-292.

[6] Tylka T L, Rodgers R F, Calogero R M, et al. Integrating social media variables as predictors, mediators, and moderators within body image frameworks. Potential mechanisms of action to consider in future research[J]. Body Image, 2023, 44: 197-221.

[7] Rodgers R F, Rousseau A. Social media and body image: modulating effects of social identities and user characteristics[J]. Body Image, 2022, 41: 284-291.

[8] Pellegrino A, Stasi A, Bhatiasevi V. Research trends in social media addiction and problematic social media use: a bibliometric analysis[J]. Frontiers in psychiatry, 2022, 13: 1017506.

[9] Tiggemann M. Digital modification and body image on social media: Disclaimer labels, captions, hashtags, and comments[J]. Body Image, 2022, 41: 172-180.

[10] Harriger J A, Wick M R, Sherline C M, et al. The body positivity movement is not all that positive on TikTok: A content analysis of body positive TikTok videos[J]. Body image, 2023, 46: 256-264.

[11] Fardouly J, Slater A, Parnell J, et al. Can following body-positive or appearance-neutral Facebook pages improve young women's body image and mood? Testing novel social media micro-interventions[J]. Body Image, 2023, 44: 136-147.

[12] Prichard I, Taylor B, Tiggemann M. Comparing and self-objectifying: the effect of sexualized imagery posted by Instagram Influencers on women 's body image[J]. Body Image, 2023, 46: 347-355.

[13] McComb C A, Vanman E J, Tobin S J. A meta-analysis of the effects of social media exposure to upward comparison targets on self-evaluations and emotions[J]. Media Psychology, 2023, 26(5): 612-635.

[14] Wood-Barcalow N L, Alleva J M, Tylka T L. Revisiting positive body image to demonstrate how body neutrality is not new [J]. Body Image, 2024, 50: 101741.

[15] Powell J, Pring T. The impact of social media influencers on health outcomes: systematic review[J]. Social Science & Medicine, 2023: 116472.

[16] Dhadly P K, Kinnear A, Bodell L P. # BoPo: Does viewing body-positive TikTok content improve body satisfaction and mood?[J]. Eating Behaviors, 2023, 50: 101747.

[17] Becker E, Rodgers R F, Zimmerman E. # Body goals or# Bopo? Exposure to pregnancy and post-partum related social media images: effects on the body image and mood of women in the peri-pregnancy period[J]. Body Image, 2022, 42: 1-10.

[18] Rodgers R F, Gordon A R, Burke N L, et al. Parents and caregivers are critical players in the prevention and identification of body image concerns and eating disorders among early adolescents[J]. Eating Disorders, 2024: 1-24.

[19] Jarman H K, Fuller-Tyszkiewicz M, McLean S A, et al. Who is most at risk of poor body image? Identifying subgroups of adolescent social media users over a year[J]. Computers in Human Behavior, 2023, 147: 107823.

[20] de Valle M K, Wade T D. Targeting the link between social media and eating disorder risk: a randomized controlled pilot study[J]. International Journal of Eating Disorders, 2022, 55(8): 1066-1078.