# A Comparative Study of User Privacy Concerns and Emotions in Online Privacy Dispute Events (Postprint)

**Authors:** Tan Fang, Yang Yang, Zhuo Yiling, Xu Jian, Xiao Zhuo

**Date:** 2023-04-01T16:02:45+00:00

## Abstract

[Purpose/Significance] The rapid development of artificial intelligence, big data, and other fields has intensified the tension between commercial development and privacy protection. Through comparative sentiment and topic analysis of Weibo comments on different types of online privacy dispute incidents, this study explores the similarities and differences in internet users' privacy attitudes across various contexts, as well as the underlying mechanisms. [Method/Process] We collected relevant Weibo comments on online privacy dispute incidents from 2012 to 2019 and preprocessed them as experimental data; calculated sentiment intensity values for each comment based on a sentiment lexicon, categorized privacy dispute incidents into privacy collection, privacy exposure, and privacy policy types, and comparatively analyzed sentiment trends in user comments across different contexts; constructed a bipartite network of user privacy discussion objects and sentiment expressions, built a single-vertex network through bipartite network projection, and conducted bipartite network and projection analysis incorporating node centrality and other metrics. [Results/Conclusion] The results indicate that users' overall privacy concern exhibits an upward trend; the intensity levels of negative sentiment vary across different types of privacy dispute incidents; users' focus areas differ significantly across various privacy dispute contexts, with distinct characteristics in emotional expression. These findings demonstrate that user privacy concerns and emotional expressions exhibit notable differences across different contexts.

## Full Text

# A Comparative Study of User Privacy Concerns and Sentiment Characteristics in Online Privacy Controversial Events

**Tan Fang**[1], **Yang Yang**[1], **Zhuo Yiling**[1], **Xu Jian**[1], **Xiao Zhuo**[2] [1]School of Information Management, Sun Yat-sen University, Guangzhou 510006 [2]Sun Yat-sen University Library, Guangzhou 510275

**Abstract:**
[**Purpose/Significance**] The rapid development of artificial intelligence, big data, and related fields has intensified the conflict between commercial development and privacy protection. Through comparative analysis of sentiment and topics in Weibo comments on different types of online privacy controversial events, this study explores the similarities, differences, and underlying mechanisms of internet users' privacy attitudes across various contexts.
[**Method/Process**] We collected and preprocessed Weibo comments related to online privacy controversial events from 2012 to 2019 as experimental data. Based on a sentiment dictionary, we calculated the sentiment intensity values of each comment and classified privacy controversial events into three categories: privacy collection, privacy exposure, and privacy agreement. We then conducted comparative analysis of user comment sentiment trends across different contexts. Additionally, we constructed a user privacy discussion object-emotional expression bipartite network and projected it into single-vertex networks, analyzing both the bipartite network and its projections using node centrality metrics.
[**Result/Conclusion**] The results show that users' overall privacy concerns exhibit an upward trend; users demonstrate different levels of negative sentiment intensity across different types of privacy controversial events; and users' hotspots of concern vary significantly across different privacy controversial contexts, with distinct emotional expression characteristics. These findings indicate that user privacy concerns and emotional performance differ significantly across various contexts.

**Keywords:** internet privacy, privacy concern, sentiment analysis, bipartite network
**Classification Number:** G203
**DOI:** 10.13266/j.issn.0252-3116.2021.02.009

The rapid development of emerging fields such as big data and artificial intelligence has made it possible to mine associated privacy information based on user data, collectively posing significant risks to privacy security. The "2019 First Half China Internet Network Security Situation" report released by the National Internet Emergency Center states that "phenomena such as mobile apps forcing authorization, excessive permission requests, and collection of personal information beyond scope are widespread, with prominent issues of illegal and non-compliant use of personal information" [1]. In the face of such severe

privacy leakage, some users resort to providing false information or even refusing to provide personal information to protect their privacy [2-3]. However, for enterprises, users' privacy disclosure is crucial for technological development, business operations, and commercial value realization. Therefore, research on user privacy attitudes and behaviors is of great importance. Currently, privacy research in the information field primarily focuses on exploring influencing factors or mechanisms of privacy concerns and behaviors of specific groups in social media, big data, and targeted advertising contexts [4-6]. Studies have shown that privacy cognition and behavior are highly context-dependent [7], meaning that users' privacy cognition and behavior differ across various contexts such as user characteristics, cultural backgrounds, business environments, and platform types [8-10]. This study organizes and classifies online privacy controversial events, conducting comparative analysis of privacy concerns and emotional characteristics based on user comments to explore the similarities and differences in internet users' privacy attitudes across different contexts, providing theoretical references for privacy practice decisions by enterprises and relevant policymakers.

## 1 Literature Review

Privacy was first legally defined as "the right to be let alone, free from external interference" [11]. To date, four perspectives on privacy definition have emerged: rights theory, commodity theory, control theory, and state theory [12-14]. The control theory of privacy originates from information control theory, as A.F. Westin proposed "the right of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others" [15]. The control theory has been widely applied in online privacy definition and privacy concern research [16]. However, the connotation of privacy continuously evolves with information technology development, necessitating a more inclusive definitional approach [13]. Communication Privacy Management theory, proposed by S. Petronio [17], defines privacy from a cybernetic perspective and views privacy boundaries as dynamic, gaining widespread application in online privacy research [18-19]. Based on this theory, this paper defines online privacy as "information presented on the internet that individuals believe belongs to them and can be controlled and managed in terms of its dissemination scope."

Privacy concern refers to "people's subjective cognition that privacy information may be leaked" [20]. In information science, privacy concern research encompasses influencing factors and behavioral intentions. Li He et al., based on construal level theory, found that privacy concern only affects social media users' long-term privacy disclosure intentions [21]. Shen Wang et al. discovered that social media users' privacy concerns negatively impact both their short-term and long-term disclosure willingness [22]. However, both studies are limited to social media contexts, and whether their conclusions apply to other contexts remains worth exploring. Chen Xiaoyan et al. studied the mechanism of user

privacy concern in LBS contexts, finding that the negative effect of privacy concern on disclosure willingness still holds in mobile service environments [23]. Consequently, some scholars note that meaningful research has shifted from studying users' disclosure willingness to examining different user reactions in privacy intrusion contexts [24]. Some researchers have investigated individual differences in privacy concern; for example, E. Vanden Broeck et al. divided adulthood into three stages and found that middle-aged adults are more concerned about their privacy than emerging and young adults [25]. Yang ? et al. studied users' privacy concern status regarding targeted advertising, finding that users of different ages and education levels exhibit varying privacy concern levels, with "users under 19 and those with college education or below showing lower privacy concern levels" [26]. However, these studies often focus on usage scenarios on a particular platform and do not explore and differentiate privacy concerns across different contexts. Gao Shanchuan et al. distinguished platform types and found that different platform types affect users' authorization willingness differently [10]. L. Yu et al. considered the impact of platform type on user privacy disclosure, discovering that users' perceived privacy risk in utilitarian platforms negatively affects disclosure willingness far more than in emotional platforms [7]. Such research confirms that user privacy cognition and behavior differ across contexts, but beyond platform type, other contextual differences—such as privacy intrusion methods and types of violated privacy information—remain unknown regarding their impact on user privacy concern and cognitive performance. These differences can provide valuable references for privacy practices in the internet industry.

Privacy controversial events, as potential incidents that intensify user privacy concerns, represent valuable contexts and data sources for privacy concern research. Users' emotional and cognitive reactions following different types of events are worth investigating [27]. Based on this, this study examines online privacy controversial events and their Weibo comments from 2012 to 2019. By reviewing existing literature and classifying events according to different privacy intrusion methods, this research attempts to answer the following questions: (1) How do users' privacy concerns trend across different online privacy controversial contexts? (2) Do users exhibit different privacy emotional characteristics across different online privacy controversial contexts? (3) Do users' privacy concern hotspots differ across different online privacy controversial contexts, and what are these differences? This paper examines and compares users' emotional characteristics and concern hotspots across different contexts, combining bipartite network and single-vertex network analysis methods to uncover the underlying differential mechanisms. Its innovation lies in analyzing the emotional dimension of users' privacy concerns and attitudes across different contexts, providing a new perspective for privacy analysis based on mass media.

## 2 Research Design

This research design comprises six components: data acquisition and processing, sentiment value calculation, network node extraction, sentiment value comparison, bipartite network construction, and comparative analysis of user privacy concern hotspots and emotional characteristics, as shown in Figure 1 [Figure 1: see original paper].

### 2.1 Data Acquisition and Processing

We used China News [28] as the primary source for news event collection, with Sohu News Platform [29] and Sina Weibo Platform [30] as supplements, searching using keywords such as "app privacy," "social privacy," and "technology privacy." Collection criteria were: (1) focus on event-based news; (2) exclude illegal incidents such as online fraud and hacker attacks. Data collection occurred in October 2019. Using the collected events as units, we searched for these events on the Sina Weibo platform, requiring comments to meet the following conditions: (1) more than 30 valid comments; (2) primarily privacy-related discussion with diverse stances. A total of 31 events from 2012 to 2019 were obtained. Next, we used the Octopus Collector [31] to crawl Weibo comments and performed preprocessing, including deduplication and comment noise reduction. The final dataset comprised 20,471 comments, with evolution trends shown in Figure 2 [Figure 2: see original paper].

The number of events and comments fluctuated across years, but the overall trend was upward. Due to collection timing, 2019 data was incomplete, showing a slight decrease from 2018, which is normal. The number of privacy controversial events and comments varied across years. For instance, in 2015 and 2016, fewer privacy controversial events were identified, while more cybersecurity illegal incidents such as online fraud and information trading were exposed, possibly due to information security policy implementation, enhanced cybersecurity efforts, and the popularization of real-name registration systems. To make the results more representative and considering comment volume fluctuations and balanced time intervals, we grouped 2012-2015 as Period I (9 events, 6,360 comments) and 2016-2019 as Period II (22 events, 14,111 comments).

### 2.2 Sentiment Value Calculation and Comparison

Based on the sentiment lexicon, we calculated sentiment intensity values for comment data. For a comment C containing n positive sentiment words and m negative sentiment words, the sentiment intensity value is calculated using Formula (1):

$$SentiC = \sum_{i=0}^{n}(D_{iT}i + E_i) - \sum_{j=0}^{m}(D_{jT}j + E_j)$$

where SentiC represents the sentiment intensity value of comment C; i and j denote the sequence numbers of positive and negative sentiment words in the

comment, respectively; D corresponds to the degree score of adverbs or negation words preceding the sentiment word (1 if none); T corresponds to the sentiment score of the word; and E corresponds to the exclamation mark score after the sentiment word (0 if none, 2 if present) to enhance weight [36].

We used the sentiment vocabulary ontology from Dalian University of Technology as the basic sentiment dictionary [32], supplemented by the danmaku sentiment dictionary from Baidu Wenku, removing emotion words that express feelings but cannot express users' privacy views [33], and adding privacy-view expression words based on comment data. Additionally, we collected common degree adverbs and negation words to form corresponding dictionaries, ultimately obtaining sentiment words [34], common degree adverbs, and negation words [35] to jointly constitute the sentiment lexicon for this study.

Classifying privacy controversial events based on existing literature and comparing their sentiment intensity value changes can reflect users' emotional characteristics toward privacy across different contexts, aiding privacy practice decisions in various fields, with evolution trends also providing references for predicting user privacy attitude trends.

### 2.3 Network Node Extraction and Bipartite Network Construction

A bipartite network is an important form of complex networks, consisting of two types of nodes and their connections, where only nodes from different groups can connect [37]. Projecting a bipartite network can form two single-vertex networks composed of node categories to reflect relationships among similar nodes, as illustrated in Figure 3 [Figure 3: see original paper] [38]. Bipartite networks can reveal deeper network characteristics and have been applied in various fields such as personalized recommendation [39] and author co-authorship studies [40]. This paper uses bipartite networks to visualize different types of privacy controversial events, constructing discussion object-emotional expression node relationships to understand users' privacy concerns and emotional presentation patterns across event types.

The implementation steps are as follows:

(1) Conduct high-frequency word statistics on user comments, filter out privacy "discussion objects" and "emotional expression" vocabulary to form corresponding dictionaries [41]. Based on these dictionaries, use the Harbin Institute of Technology LTP tool to extract discussion object-emotional expression word pairs, constituting two groups of nodes in the bipartite network. Ultimately, 1,206 relationship groups were extracted to form an association word list as the original dataset.

(2) Based on the dataset, use Gephi complex network analysis software [42] to construct the bipartite network, selecting appropriate parameters and clustering methods to improve visualization effects.

(3) To further explore the core presentation patterns of "discussion objects"

and "emotional expression" in user privacy concerns, import the dataset into the social network analysis tool Pajek [43], transform the bipartite network into single-vertex networks through weighted projection, and use Gephi [42] for visualization.

By comparing and analyzing users' concern subjects and emotional expressions across different types of privacy controversial events, we further explore the specific content of users' privacy concerns and similarities and differences in emotional expression across contexts, analyzing underlying differential mechanisms to provide theoretical references for enterprise decision-makers.

## 3 Comparative Analysis of User Privacy Concerns and Emotional Characteristics

### 3.1 Privacy Event Classification

The 2016 "Cybersecurity Law of the People's Republic of China" comprehensively stipulates the network information security system in Chapter 4 [45]. Based on the definition of online privacy in Section 1 of this paper, Articles 40-45 are directly related to online privacy infringement events. Specifically, Articles 41 and 43 regulate network operators' collection and use of user information; Article 42 addresses network operators providing user information to third parties; and Article 44 covers third-party theft and trafficking of personal information [45]. These provisions primarily target the collection, use, theft, and resale of privacy information. Combined with the collected events, online privacy controversial events are classified into three types, as shown in Table 1 .

**Table 1 Classification of Privacy Controversial Events**

| Type | Definition | Example Events |
|---|---|---|
| Privacy Collection | Events where organizations/products/services/technologies are suspected of collecting user personal data without permission, or using collected information for precision marketing, R&D, or providing to third parties for various commercial purposes, thereby causing privacy controversies. | "Cookies tracking user behavior technology causing privacy controversy""QQ chat window inputting 'Wilderness Action' pops up chicken dinner game ads, QQ suspected of monitoring user chat records" |

| Type | Definition | Example Events |
|---|---|---|
| Privacy Exposure | Events where organizations/products/services/technologies publicly disclose user personal data without permission to achieve various commercial purposes, thereby causing privacy controversies. | "What you ordered for takeout, your friends all know! Takeout app binding with social features accused of privacy infringement""360 camera live streaming online causing privacy controversy—eating and fitness can be seen clearly" |
| Privacy Agreement | Events where privacy controversies are directly caused by non-standard privacy agreements, or forced or default authorization. | "'ZAO' app user privacy agreement non-standard controversy""Alipay default checks Sesame Credit agreement" |

The following analysis adopts the classification standard shown in Table 1.

### 3.2 Comparison of Privacy Event and Comment Quantity Characteristics

We classified the 31 events obtained in Section 2.1 according to the types defined in Table 1. The trends in event numbers and comment quantities for each type are shown in Figure 4 [Figure 4: see original paper].

Figure 4 shows that the quantity of all event types increased over time, while comment volume changes exhibited distinct characteristics. Privacy collection events were the most numerous in both periods, followed by privacy exposure events, with privacy agreement events being the least. However, privacy collection events showed the smallest sentiment intensity value decline (56%), indicating relatively modest changes in privacy concern, attributable to "privacy fatigue" caused by increased effort costs and reduced effectiveness of privacy protection in this context. Privacy exposure events showed the largest sentiment intensity value decline (over 400%), dropping from 0.79 in Period I to -2.57 in Period II. Privacy agreement events emerged only in Period II, with comment volumes surpassing those of privacy exposure events, representing a new hotspot in privacy controversies.

### 3.3 Comparison of Sentiment Values Across Different Controversial Event Types

We compared the emotional evolution of users across different privacy controversial events, with trend changes shown in Figure 5 [Figure 5: see original paper].

The sentiment of early comments for all event types showed a downward trend. Specifically, privacy exposure events' sentiment intensity values dropped from 0.79 in Period I to -2.57 in Period II; privacy collection events decreased from -0.96 to -1.50; privacy agreement events, emerging only as a new type of privacy controversy, dropped from 0 to -1.47, aligning closely with privacy collection events. Enhanced negative emotion often signifies increased user privacy concern [20]. This indicates that user privacy concerns across different event types are rising, correlating with the proliferation of big data applications and frequent exposure of privacy infringement incidents—i.e., changes in the volume of negative media exposure positively impact user privacy concern levels, consistent with Guo Longfei's findings [46]. Notably, despite privacy collection events being the most numerous, they showed the smallest increase in privacy concern, while privacy exposure events showed the largest increase. This suggests that differences in user privacy concern across contexts are not significantly correlated with the static proportion of negative media exposure. A plausible explanation is that privacy collection phenomena have long existed with the highest frequency across stages, increasing users' effort costs for privacy protection while reducing effectiveness, leading to deeper "privacy fatigue," which can promote privacy-negative behaviors [47], such as adopting an "indifferent" attitude toward personal information misuse, as reflected in comments like "Whatever, we're already transparent anyway."

### 3.4 Comparison of Privacy Event Concern Hotspots and Emotion Across Different Types

This section constructs bipartite networks and node projections to further understand users' privacy concern hotspots and emotional presentation patterns across different privacy controversial event types.

**3.4.1 Bipartite Network Analysis** Following the steps described in Section 2.3, we constructed discussion object-emotional expression bipartite networks for the three controversial event types, shown in Figure 6 [Figure 6: see original paper]. In Figure 6, node size is sorted by "degree"—larger nodes indicate more frequent mentions; edge size and color depth represent "edge weight"—thicker edges indicate more frequent co-mentions of the two nodes.

In privacy collection events, larger nodes include "Apple" and "phone," while in emotional expression, larger nodes include "logout" and "steal." Reviewing original comments reveals significant user concern about app logout issues, such as "Logging out doesn't mean your information is deleted; it still exists in others' databases," indicating concern for the right to be forgotten. In privacy exposure events, larger nodes include "function" and "camera," while emotional expression nodes primarily include "garbage" and "rogue," reflecting user disgust, as seen in comments like "Privacy is indeed violated, and being live-streamed when going out, ** is garbage." In privacy agreement events, larger nodes include "Alipay" and "face recognition," while emotional expression nodes include

"leak" and "security." In this context, users particularly concern themselves with biometric information privacy security, as reflected in comments like "I don't dare to use Alipay's face recognition."

**3.4.2 Discussion Object-Emotional Expression Single-Vertex Network Analysis** Bipartite networks cannot intuitively reflect direct connections between privacy-related discussion objects and emotional expressions. Therefore, by extracting "discussion object" and "emotional expression" nodes and obtaining single-vertex networks through weighted projection, we can further reflect users' attention distribution to these two node groups. Figures 7 [Figure 7: see original paper] to 9 [Figure 9: see original paper] show the discussion object and emotional expression single-vertex networks for the three privacy event types. Inter-node weights represent the number of shared connection nodes, with different grayscale colors representing different clustering categories. To optimize visualization, nodes were appropriately merged, selecting only those with degree values of 1 or above.

**(1) "Privacy Collection" Events.** As shown in Figure 7(a), "phone" has the highest betweenness centrality (535.66), serving as the core word connecting various communities. The network density shows close connections among communities. Through appropriate merging, three community types emerge: (1) Core nodes including "phone" and "Apple," accounting for 58% of nodes, involving disputes over privacy collection in communication devices and tools, as reflected in comments like "Only Apple can steal from iPhones; all apps can steal from Android phones, so I still choose Apple"; (2) Core nodes including "Baidu" and "WeChat," accounting for 21% of nodes, involving privacy collection discussions about search engines and social software, such as "WeChat also reads your chat records for precise ad targeting"; (3) Core nodes including "data" and "APP," accounting for 18% of nodes, with discussion objects involving e-commerce and lifestyle service platforms, focusing on privacy information types such as "phone number" and "bank card," as in comments like "Once you log in with your phone number, all information leaks."

In Figure 7(b) "Emotional Expression" single-vertex network, "dare not" has the highest betweenness centrality (1,122.77), serving as the core connecting communities. This network also divides into three types: (1) Community centered on "dare not" and "cautious," accounting for 47.18% of nodes, primarily expressing fear; (2) Community centered on "unspeakable" and "despicable," accounting for 27.46% of nodes, expressing disgust; (3) Community centered on "good" and "nothing unusual," accounting for 25.35% of nodes, expressing positive or neutral emotions.

In summary, users in privacy collection events primarily focus on privacy collection issues in communication devices, with discussion objects including devices themselves and various applications on them, dominated by fear-based emotions mixed with contradictory disgust and positive emotions. From the perspective of social exchange theory [50], users can exchange browsing, interaction, and

shopping information for more personalized and convenient services. However, as this marketing strategy becomes widely adopted by tech companies, users' perception of privacy risks strengthens. Since users cannot directly supervise corporate information usage, privacy risks often carry significant uncertainty. Consequently, increasing numbers of users question the effectiveness of their privacy protection behaviors—i.e., reduced privacy protection self-efficacy—leading to rising proportions of fear-based emotions while other emotion types remain relatively stable.

**(2) "Privacy Exposure" Events.** As shown in Figure 8(a), "live streaming" has the highest betweenness centrality (272.74), serving as the core connecting communities in the "discussion object" network. Through appropriate merging, two major categories emerge: (1) Core nodes including "live streaming," "camera," and "personal information," accounting for 55.32% of nodes. Reviewing original comments reveals user concerns about public surveillance being used for live streaming and its connection to personal privacy leakage, such as "Cameras installed in public places, doesn't matter if they're live streaming or not" and "I don't want to be live streamed while eating"; (2) Core nodes including "enterprise," "Zhou Hongyi," and "Dianping," accounting for 21.27% of nodes, primarily referring to privacy exposure behaviors in various lifestyle service applications, such as "Meituan can also see what WeChat friends ordered, scaring me from ordering."

In Figure 8(b) "Emotional Expression" single-vertex network, "redundant" serves as the core connecting communities with betweenness centrality of 449.13. Communities can be merged into: (1) Core nodes including "redundant," "brain-dead," and "refuse to use," accounting for 37.66% of nodes. Combined with comment data, these represent evaluations of discussed objects' functions, indicating low functional value not worth the privacy cost; (2) Core nodes including "ulterior motives" and "security," accounting for 22.08% of nodes, reflecting users' cautious attitudes toward privacy security; (3) Core nodes including "not bad" and "apology," accounting for 12.99% of nodes, representing users' forgiving attitudes after platform or corporate apologies, such as "The attitude of admitting mistakes is quite good."

In summary, the main concern hotspots in privacy exposure events are social entertainment applications like "live streaming," with users expressing relatively strong negative emotions. Unlike privacy collection events, users in these events have clear perceptions of privacy leakage and usage, such as various surveillance videos "leaked" in monitoring live streaming incidents, leading to strong dissatisfaction and questioning. Additionally, based on privacy calculus theory, when facing privacy issues, users evaluate expected benefits against perceived privacy risks to make privacy decisions [51]. In this situation, users subconsciously weigh product value (primarily functional value). When perceived benefits are lower than the cost of selling privacy but privacy risks are forced upon them, strong disgust and resistance emotions emerge.

**(3) "Privacy Agreement" Events.** As shown in Figure 9(a), "Alipay" serves

as the core connecting communities in the "discussion object" network with betweenness centrality of 334.92. Through appropriate merging, two categories emerge: (1) Core nodes including "Alipay," "user," and "risk," accounting for 75% of nodes. Combined with original comments, this reflects user concerns about privacy risks caused by non-standard user agreements, such as "Alipay's agreement clearly poses serious personal privacy security risks"; (2) Core nodes including "personal information" and "face recognition," accounting for 17.31% of nodes, involving concerns about biometric information privacy, such as "A facial photo can leak many biometric features, and can unlock some Android phones on the market, so use with caution."

In Figure 9(b) "Emotional Expression" single-vertex network, "tricky" serves as the core connecting communities with betweenness centrality of 822.59. Through appropriate merging, communities divide into: (1) Core nodes including "tricky," "infringement," and "refuse to use," accounting for 53.42% of nodes, expressing disgust and resistance; (2) Core nodes including "full of holes," "wrong," and "awesome," accounting for 27.40% of nodes. Reviewing original comments reveals condemnation of corporate privacy infringement and recognition of product value, reflecting controversial attitudes among different users.

In summary, in privacy agreement events, users are more concerned about privacy issues in lifestyle service applications, with negative emotions dominating. L. Yu categorized platforms into emotional and utilitarian types, with the former primarily including social and entertainment platforms with emotional interactions, while the latter mainly includes transaction-oriented platforms [7]. In this study, discussion subjects in privacy collection and exposure events primarily involve emotional applications, while privacy agreement events mainly involve utilitarian platforms. Compared to emotional platforms, users in utilitarian platforms have specific purposes for privacy disclosure, thus behaving more rationally and conducting more careful assessments of platform privacy policies and cost-benefit evaluations [7]. User agreements serve as basic written guarantees for user privacy. When non-standard, they trigger disgust and resistance among most users, also verifying E.S. Wang's conclusion from another perspective that reduced perceived effectiveness of privacy policies significantly undermines user confidence in usage behavior and directly impacts their actions [52].

## 4 Conclusion and Limitations

The rapid development of the internet has made online privacy issues increasingly severe. Existing research primarily focuses on the causes of privacy concerns and behavioral intentions, with limited studies on user privacy concerns and emotions across different contexts. This paper focuses on privacy controversial events in legal regulatory gray areas, examining online privacy controversial events and Weibo comments from 2012 to 2019. We conducted differential analysis of user privacy concerns, emotions, and concern hotspots across contexts,

including comment data acquisition and processing, sentiment value calculation, network node extraction, sentiment value comparison, bipartite network construction, and comparative analysis of user privacy concern hotspots and emotional characteristics. This study supplements research perspectives in the privacy concern field, with conclusions providing practical references for decision-makers in internet and related industries. The main conclusions are:

(1) User privacy concerns in different types of privacy controversial events show an upward trend. Differences in user privacy concern performance across event types are not significantly correlated with the static proportion of negative media exposure. Based on existing literature, privacy controversial events are classified into "privacy collection," "privacy exposure," and "privacy agreement" categories. Among them, privacy collection events have the largest quantity (22 events) and comment volume (12,223 comments) across periods, yet the smallest sentiment intensity value decline (56%). The relatively modest change in privacy concern can be attributed to "privacy fatigue" caused by increased effort costs and reduced effectiveness of privacy protection in this context. Privacy exposure events showed the largest sentiment intensity value decline (over 400%), dropping from 0.79 in Period I to -2.57 in Period II. Privacy agreement events emerged only in Period II, with comment volumes surpassing those of privacy exposure events, representing a new hotspot in privacy controversies.

(2) Different types of privacy controversial events exhibit different levels of negative sentiment intensity. Based on construal level theory, when psychological distance is closer—i.e., when perceived risks of privacy infringement are clearer—users exhibit stronger negative emotions and privacy concerns in privacy controversial events. In privacy exposure events, users have clear perceptions of privacy leakage, while in privacy collection and agreement events, privacy risk perceptions are often hypothetical. Consequently, the former shows stronger negative emotions than the latter two.

(3) Users' privacy concern hotspots differ significantly across contexts, with distinct emotional expression characteristics. In privacy collection events, users primarily focus on information collection issues in communication devices. However, since users cannot directly supervise corporate information usage, perceived privacy risks carry significant uncertainty, resulting in fear-based emotions. In privacy exposure events, concerns focus on social entertainment applications, with users expressing relatively strong negative emotions. In privacy agreement events, which involve utilitarian platforms, users behave more rationally and cautiously. Reduced perceived effectiveness of privacy policies significantly impacts user emotions, also dominated by disgust-based emotions, verifying E.S. Wang's conclusion from another perspective [52].

This paper has several limitations: First, Weibo comments only represent pri-

vacy attitudes of a subset of users, without considering users on other platforms and non-internet user groups, thus the conclusions have certain limitations. Second, this paper treats all internet users as a single group; future research could further investigate differences in online privacy emotions and concern hotspots across different user groups.

## References

[1] National Computer Network Emergency Response Technical Team/Coordination Center of China. 2019 First Half China Internet Network Security Situation [R/OL]. [2020-03-24]. http://www.cac.gov.cn/2019-08/13/c_{1124871484}.htm.

[2] Wang Le, Wang Luyao, Sun Zao. The Influence Mechanism of Privacy Invasion Experience on Internet Users' Self-Disclosure [J]. Systems Engineering—Theory & Practice, 2020, 40(1): 79-92.

[3] MARTIN K D, MURPHY P E. The Role of Data Privacy in Marketing [J]. Journal of the Academy of Marketing Science, 2017, 2(45): 135-155.

[4] Yu Tingting, Yang Yunhan. Research on Privacy Concerns and Their Influencing Factors in Precision Advertising [J]. Journalism Research, 2019(9): 101-116.

[5] Liu Qian. Research on Motivations for Privacy Management Among Young Users in WeChat Moments [J]. Contemporary Communication, 2019(4): 84-89.

[6] ALASHOOR T, HAN S, JOSEPH R C. Familiarity with Big Data, Privacy Concerns, and Self-Disclosure Accuracy in Social Networking Websites: An APCO Model [J]. Communications of the Association for Information Systems, 2017, 41(4): 62-96.

[7] YU L, LI H, HE W, et al. A Meta-Analysis to Explore Privacy Cognition and Information Disclosure of Internet Users [J]. International Journal of Information Management, 2020, 51: 1-10.

[8] JEONG Y, COYLE J R. What Are You Worrying About on Facebook and Twitter? An Empirical Investigation of Young Social Network Site Users' Privacy Perceptions and Behaviors [J]. Journal of Interactive Advertising, 2014, 2(14): 51-59.

[9] JEONG Y, KIM Y. Privacy Concerns on Social Networking Sites: Interplay Among Posting Types, Content, and Audiences [J]. Computers in Human Behavior, 2017, 69: 302-310.

[10] Gao Shanchuan, Wang Xinyi. The Influence of Network Platform and Benefit Types on Information Privacy Decision-Making [J]. Chinese Journal of Applied Psychology, 2019, 25(4): 364-371.

[11] WARREN S D, BRANDEIS L D. The Right to Privacy [J]. Harvard Law Review, 1890, 5(4): 193-220.

[12] Duan Weiwen, Ji Changlin. Privacy Rights in the Internet and Big Data Era [J]. Science and Society, 2014, 4(2): 90-100.

[13] Liu Jian, Chen Zhuo. Three Discourse Analyses of "Privacy" Research [J]. Journal of Shaanxi University of Technology (Social Sciences Edition), 2017, 35(2): 47-51.

[14] Lv Yaohuai, Xiong Jiechun. The BIT Model for Defining Information Privacy [J]. Library Theory and Practice, 2011(6): 35-39.

[15] WESTIN A F. Privacy and Freedom [M]. New York: Athenaeum, 1967.

[16] Li Rui. Measurement and Empirical Research on Privacy Leakage Tolerance in Mobile Internet Environment [D]. Dalian: Dalian University of Technology, 2014.

[17] PETRONIO S. Boundaries of Privacy: Dialectics of Disclosure [M]. Albany, New York: State University of New York Press, 2002.

[18] Zhong Ying, Liu Lifang. Privacy Infringement and Protection in Information Dissemination [J]. News and Writing, 2018(2): 23-26.

[19] EISENSPANNER J. Privacy Online: Perspectives on Privacy and Self-Disclosure in the Social Web [J]. New Media & Society, 2013, 8(15): 1401-1402.

[20] Chen Hao, Li Wenli, Ke Yulong. Social Media Continuance Usage Research: Mediating Role of Emotional Response [J]. Journal of Information Systems, 2016, 30(3): 83-95.

[21] Li He, Yu Lu, Xu Yiming, et al. Social Network Privacy Paradox Research from Construal Level Theory Perspective [J]. Information Science, 2020, 38(8): 120-127.

[22] Shen Wang, Gao Xueqian, Dai Wang, et al. Social Network Privacy Paradox Research Based on Construal Level Theory and Regulatory Focus Theory [J]. Information Science, 2018, 37(1): 1-13.

[23] Chen Xiaoyan, CLIQUET G. User Privacy Concern Research from Privacy Trade-Off Theory Perspective [J]. Technical Economics and Management Research, 2020(5): 9-13.

[24] MARTIN K D, MURPHY P E. The Role of Data Privacy in Marketing [J]. Journal of the Academy of Marketing Science, 2017, 45(2): 135-155.

[25] VANDEN BROECK E, POELS K, WALRAVE M. Older and Wiser? Facebook Use, Privacy Concern, and Privacy Protection in the Life Stages of Emerging, Young, and Middle Adulthood [J]. Social Media + Society, 2015, 1(2): 671657067.

[26] Yang ?, Wen Xiuyan. The Mediating Effect of Privacy Protection Willingness: Privacy Concern, Privacy Protection Self-Efficacy, and Precision Advertising Avoidance [J]. Journalism and Mass Communication, 2020(7): 41-52.

[27] Xie Yi, Gao Chongyan, Tong Zelin. Consumer Privacy Concern Research Review and Prospects [J]. Foreign Economics and Management, 2020, 42(6): 111-125.

[28] China News [EB/OL]. [2020-03-24]. https://www.chinanews.com/.

[29] Sohu News [EB/OL]. [2020-03-24]. https://news.sogou.com/.

[30] Sina Weibo [EB/OL]. [2020-03-24]. https://weibo.com/.

[31] Octopus Collector [EB/OL]. [2020-03-24]. https://www.bazhuayu.com/.

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

[33] Deng Shuqing, Li Wanwei, Xu Jian. Emotional Word Recognition Based on Syntactic Dependency Rules and Part-of-Speech Features [J]. Information Studies: Theory & Application, 2018, 41(5): 137-142.

[34] Sentiment Dictionary [EB/OL]. [2020-03-24]. https://figshare.com/articles/12240224.

[35] Common Degree Adverbs and Negation Words Dictionary [EB/OL]. [2020-03-24]. https://figshare.com/articles/12240233.

[36] Xu Jian, Wu Siyang. Research on Sentiment Divergence Quantification Algorithm for Online User Reviews [J]. Journal of the China Society for Scientific and Technical Information, 2020, 39(4): 427-435.

[37] Lu Weicong, Xu Jian. Sentiment Analysis of Online User Reviews Based on Bipartite Networks [J]. Information Studies: Theory & Application, 2018, 41(2): 121-126.

[38] LATAPY M, MAGNIEN C, VECCHIO N D. Basic Notions for the Analysis of Large Two-Mode Networks [J]. Social Networks, 2008, 30(1): 31-48.

[39] Li Shuqing, Xu Xia, Xu Minjia. A Method for Measuring Book Recommendability and Personalized Book Recommendation Service Based on Reader-Book Bipartite Networks [J]. Journal of Library Science in China, 2015, 41(4): 1401-1402.

[40] Zhang Jinzhu, Han Tao, Wang Xiaomei. Research on Co-Authorship Relationship Prediction in Author-Keyword Bipartite Networks [J]. Library and Information Service, 2016, 60(21): 74-80.

[41] Privacy Discussion Object-Emotional Expression Dictionary [EB/OL]. [2020-03-24]. https://figshare.com/articles/{___}/12401426.

[42] BASTIAN M, HEYMANN S, JACOMY M. Gephi: An Open Source Software for Exploring and Manipulating Networks [C]//The Association for the Advancement of Artificial Intelligence. Proceedings of the Third International Conference on Weblogs and Social Media. Menlo Park, California: The AAAI Press, 2009: 361-362.

[43] BATAGELJ V, MRVAR A. Pajek—Program for Large Network Analysis [J]. Connections, 1998, 2(21): 47-57.

[44] BURGOON J K, PARROTT R, LEBPOIRE B A, et al. Maintaining and Restoring Privacy Through Communication in Different Types of Relationships [J]. Journal of Social and Personal Relationships, 1989, 2(6): 131-158.

[45] Standing Committee of the National People's Congress. Cybersecurity Law of the People's Republic of China [EB/OL]. [2020-03-24]. http://www.moj.gov.cn/Department/content/2016-11/23/592_{2013322}.html.

[46] Guo Longfei. Research on Dynamic Influencing Factors and Behavioral Patterns of Social Network Users' Privacy Concerns [D]. Beijing: Beijing University of Posts and Telecommunications, 2013.

[47] CHOI H, PARK J, JUNG Y. The Role of Privacy Fatigue in Online Privacy Behavior [J]. Computers in Human Behavior, 2018, 81: 42-51.

[48] TROPE Y, LIBERMAN N, WAKSLAK C. Construal Levels and Psychological Distance: Effects on Representation, Prediction, Evaluation, and Behavior [J]. Journal of Consumer Psychology, 2007, 2(17): 83-95.

[49] BRANDES U. A Faster Algorithm for Betweenness Centrality [J]. Journal of Mathematical Sociology, 2001, 2(25): 163-177.

[50] Li Bingquan. Research on Privacy Infringement, Boundary Delimitation, and Protection in Social Media Context [D]. Changchun: Jilin University, 2020.

[51] Liu Yinghua, Zhai Jiajing. Review of Information System Users' Privacy Calculus Research [J]. Journal of Academic Libraries, 2020, 38(3): 113-119.

[52] WANG E S. Effects of Brand Awareness and Social Norms on User-Perceived Cyber Privacy Risk [J]. International Journal of Electronic Commerce, 2019, 23(2): 272-293.

## Author Contributions

**Tan Fang:** Designed research framework, collected data, conducted experiments, wrote manuscript
**Yang Yang:** Data preprocessing
**Zhuo Yiling:** Collected data, improved sentiment dictionary
**Xu Jian:** Designed research framework, revised final manuscript
**Xiao Zhuo:** Proposed research ideas, revised final manuscript

## Comparison of Privacy Concern and Sentimental Characteristics of Users in Internet Privacy Controversial Events

**Tan Fang[1], Yang Yang[1], Zhuo Yiling[1], Xu Jian[1], Xiao Zhuo[2]**
[1]School of Information Management, Sun Yat-sen University, Guangzhou

510006
[2]Sun Yat-sen University Library, Guangzhou 510275

**Abstract:**
[**Purpose/Significance**] The rapid development of AI, big data, etc., makes an increasingly fierce confrontation between business development and privacy protection. This paper finds the differences in privacy concern and sentimental characteristics of users in different situations through the comparative analysis of sentiment and topics of Weibo comments on different types of online privacy controversial events. [**Method/Process**] Firstly, Weibo comments related to privacy controversial events from 2012 to 2019 were collected and preprocessed as experimental data. Secondly, the sentiment value of each comment was calculated based on the sentiment dictionary. Then, the privacy controversial events were classified into privacy collection categories, privacy exposure categories, and privacy agreement categories. The sentiment evolution of user comments under different situations was compared and analyzed. Finally, a bipartite network of "discussed object-emotional expression" was constructed, and a single-vertex network was constructed by the bipartite network projection. The analysis of both were carried out by combining the indexes such as node centrality. [**Result/Conclusion**] The result shows that users' overall privacy concern shows an upward trend; users of different types of privacy controversial events have different levels of negative sentiment intensity; the hotspots of users' concern vary widely across privacy controversial events, and emotional expression have different characteristics. These findings indicate that there are significant differences in user privacy concern and sentimental performance in different situations.

**Keywords:** internet privacy, privacy concern, sentiment analysis, bipartite network

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