

# Lack of Multiple Comments on Social Media Platforms

Yan Miao<sup>1\*</sup>

<sup>1</sup> New York University, New York, US

\*Corresponding author. Email: xs2119@nyu.edu

## ABSTRACT

In recent years, along with the increase of Internet users and the abundance of information access channels, the online interactions become easier and more frequent. Meanwhile, the psychology of user use under different information platforms has also undergone great changes. Along with these changes, the relationship between the opinion field and the change of users' emotions has also had been affected, especially for the popular platforms represented by social media platforms and search engines. This paper analyzes and studies the diversity of social media users' and search engine users' comments from both quantitative and qualitative perspectives. The author crawled Weibo and Baidu trends lists from September 20 to September 26, and collected user comments for sentiment analysis after selecting the specific topics. Finally, Python is used to determine the diversity of user comments from the perspective of sentiment analysis. The author conducted further interviews to investigate the reasons for data analysis results. Reviewing other social media and search engine comparison studies, this paper is the first to compare user sentiment issues, creating a new direction of qualitative analysis for the comparison of social media and search engines.

**Keywords:** Social media platform, Search engine, Trends comparison, Emotion analysis

## 1. INTRODUCTION

With the enrichment functions and increased use of mobile apps, users are engaging in more and more ways to interact with the web and other users. Likes, favorites, retweets and comments are able to reflect users' attitudes to some extent, while user comments have become an important source of information sharing. Since comments have to take some time and effort to express one's opinion or emotion, the author believes that paying attention to user comments can reflect whether users have different attitudes when using different platforms. In order to compare the diversity of user comments in search engines and social media platforms, the author first chose two representative applications to complete the study: Baidu, which has a huge number of users in China, and Weibo, which has a very high daily activate units.

Carr and Hayes mentioned that social media is an Internet-based channel that provides a platform for users to produce information and interact with specific user groups or a wide range of user groups [1]. At the same time, social media itself realizes value and profits from the perspective of user-produced content. In short, the characteristics of social media include: the Internet as the technical background, the production of information by users as the main body, and the creation of interaction and

connection by users as the goal. As the latest manifestation of social media, Weibo has become the main information exchange platform for the public and the gathering place for social thoughts and opinions. In order to narrow down the scope of the study, the author tried to focus on Weibo, mainly because of its significant social media characteristics and the fact that it can be studied and investigated in a representative manner. On the other hand, Weibo released its first quarter 2021 financial results on May 10. According to the earnings data, as of March this year, Weibo had 530 million monthly active users and 530 million daily active users. The number of active users was 230 million. The large number of users can bring enough data for this study, which is why the author chose Weibo as a representative product of social media platforms.

Except for the increase users of Weibo, in China, the growth of mobile search is also extremely fast. As of June 2012, the number of mobile search users reached 260 million, an increase of 17.3% over 2011. It is expected that mobile search will continue to grow at a high rate in the coming years because of the increasing demand of users to find information anytime and anywhere, and mobile search is more timely and convenient than desktop search. In terms of the Chinese market, Baidu accounted for 71.10% of the Chinese search engine market in

February 2021, which is the largest share [2]. Considering of user usage and product market share, Baidu, as the representative search engine in this study, has certain representativeness and research value.

In previous studies, most scholars have looked at specific media, such as social media represented by Twitter and Facebook, to explore the forms of user interaction and the direction of information dissemination on the platform. Some scholars have also analyzed the motivation of users to generate interactive behaviors in different arenas using qualitative analysis. Almost no one noticed the differences in user comments on different types of online platforms, which aroused the author's curiosity.

To conclude, this paper enriches the framework of user review diversity in different platforms to a certain extent. Meanwhile, this paper delves into the reasons for the characteristics of user behavior under different opinion fields from a qualitative perspective.

## 2. LITERATURE REVIEW

The author first collected past research and found that most of the studies focused on the impact of online media user reviews on advertising and marketing, or on predicting health communication based on search engine data. Little attention has been paid to the comparison between social media and search engines, either in terms of user habits or user interactions.

In a previous study, Hsueh, Yogeewaran and Malinen explored the issue of "civility" of content posted by anonymous users in online reviews. The study analyzed the attitudes of professional journalists and forum participants during interactive sessions through text analysis and interviews. The study demonstrates that anonymity in online forums provides users with the opportunity for emotional expression. This is because users believe that anonymity helps them to hide their true identity well. It has led to a lot of incivility in the comments of anonymous users and a clear lack of calmness and politeness among users [3].

Walther et al., suggested that in the new media era, user comments have become much more influential on the audience than the detail information provided by the producers. This is the reason why the authors wanted to find out the influence of implicit messages or attitudes in users' comments on other users' participation in interactions [4].

Zhao and Li suggest that consumers increasingly value online product reviews as a pre-purchase reference to guide their shopping behavior. The study constructed a model based on consumer review situations for analyzing the relationship between user reviews and consumer perceptions in online shopping scenarios, and from this, inferred that the influencing factors of each dimension in

online reviews are positively related to consumer perceptions [5].

Also focusing on the impact of user reviews on consumers in the online shopping scenario, Rashid and Zeeshan focused on the comparison of traditional and online media advertising and marketing dimensions [6]. The study mapped a network to quantify the effect of advertising. It also evaluates user attitudes towards different brands of cell phones based on close to 3,000 video user comments and responses, which provides support for advertising video marketing and cell phone category user comment management.

One of the few studies around the correlation of search engine and social media hits is also based on health communication issues, as exemplified by the 2015 Study of the high correlation between the spread of Middle East Respiratory Syndrome (MERS) and Google search and Twitter trends in South Korea, Shin et. al. evaluated the use of a digital surveillance system based on web search and social media data to monitor this MERS outbreak likelihood [7]. They used Google Trends and Topsy to identify data for daily relevant information. Spearman correlation analysis was then used to examine the correlation between the data. Analysis of the results with four keyword searches as text revealed a high correlation between Google searches and Twitter results and the number of cases using the keywords mentioned above. This study demonstrates the potential of using digital surveillance systems to monitor MERS outbreaks. It is also a comparison and contribution of search engines and social media in synchronizing health communication.

Silverstovs and Wochner examined the correlation between popular search topics in Google search and topics that users actually follow. The study used the representation of real-world economy activity reflected in Google search as an example, based on Google Knowledge Graph technology to accurately measure the volume of relevant search queries [8]. And to explore the similarity between the search data of Swiss tourism demand and the real-world situation, the study assesses the information value of the data as strong, semi-strong or weak based on unbiasedness and efficiency considerations in a Mincer-Zarnowitz-type regression model. It also found that the search-based tourism demand forecasts in the Swiss tourism demand example are, on average, highly accurate generalizations of reality. This suggests that search-based metrics can serve as a valuable real-time complement to economic policy guidance.

In general, there are several shortcomings in past studies:

1. No direct comparison of user reviews under different APPs.
2. No sentiment analysis of user comments under different topics.

3. No analysis of results using quantitative and qualitative analysis.

**3. DATA CRAWLING**

**3.1. Trending Topics**

Baidu and Weibo track the most frequently occurring phrases, words, and hashtags mentioned by the users and publish them under the heading "Popular Trends." Baidu uses user search frequency to complete the generation of popular search lists. On social media platforms, hashtags are a convention between Weibo users to create and follow discussion specific topics by adding "#" characters in front of the word and at the end of the word. It has been commonly recognized that the popular trends are a good representative of the most concerning issues [9], so the author decided to crawl them for this study. The author uses the query string to get the daily trends on Weibo and Baidu from September 20 to September 26. In order to avoid the impact of different times for collecting data on the results, the author began to collect data at 10 o'clock every evening. Finally, 350 popular trends on Weibo and 210 trends on Baidu were collected within 7 days.

**3.2. Comments Collection**

The author collected 10 hashtags that were exactly the same on Baidu and Weibo on the same day. Based on the same topic, search for the result on the two platforms, and collect comments response to the most popular microblog or the hottest article. Through crawling data, the maximum number of comments on Weibo can reach more than 27,000, and the maximum number of comment data on Baidu is 190. In order to make the data comparison more representative, the author selected the top 20 popular comments as representatives for subsequent analysis. During the collection process, this research collected the full text, author, time of writing, as well as reviewer, time of the comment, and other information.

**4. RESULTS**

**4.1 Trends similarities on the same day of Two platforms.**

After completing the first step of collecting the trends, the author compared the similarity of the trends lists on the two platforms on the same day. Firstly, extract keywords from each hashtag. Then use the cosine similarity to calculate the keywords to get the similarity between the terms. The results are shown in Table 1. The red numbers mean it is higher than the average similarity of 30%, and the green data is lower than the average. Only on September 25th, the trends similarity between Baidu and Weibo exceeded 50%. In other words, most of the

time, what users search on Baidu is not similar to what users search on Weibo.

**Table 1.** Similarity between the terms.

Date	Similarity
20th Sep.	0.14
21st Sep.	0.39
22nd Sep.	0.24
23rd Sep.	0.14
24th Sep.	0.37
25th Sep.	0.55
26th Sep.	0.24

The author believes that the large difference between the keywords of the two platforms' trends is due to the characteristics of different APPs. Search engine users usually hope to obtain direct and accurate answers through search instead of participating in discussions on the topic. However, social media platforms such as Weibo use the user's past data for analysis and then push relevant news to the user [10]. In this case, the user is actually passively receiving information. What's more, because social media platforms accurately publish users' favorited topics or content, users will be more willing to spend longer on the platform. This is why, compared to Baidu, Weibo has richer categories of trends.

**4.2 Terms similarities of One platform within one week.**

The author conducted a comparative analysis of the hot search terms on the same platform within a week by hierarchical clustering. Similar terms are grouped together. Among all the hashtags, the same label with a similarity greater than 0.2 will be classified into one category. The results indicate that the trends on the two platforms hardly repeated within a week. It indicates that users are hardly remain their interests on a same topic within one week.

After collecting data for the entire week, the author compared the trends throughout the week to see if they are related or similar on a particular platform. Therefore, the author tried to categorize these hashtags and view the frequency of keywords in blocks. The data in the appendix shows that although several similar keywords appear repeatedly within a week, they point to different content. For example, the repeated occurrence of words related to the Chinese National Games is actually because different players won the championship in the week.

**4.3 Sentiment Analysis of Two platforms.**

The sentiment analysis was conducted by using the snownlp package to output the sentiment value of each

comment. The following pie charts indicate there is almost no difference in the distribution of positive attitudes on Baidu and Weibo. Although more negative comments can be found in Baidu's user reviews, because the number of Baidu's comments is relatively small, the impact can be said to be minimal. In fact, the author originally believed that the filter bubbles created by social media platforms would reduce the number of people with different attitudes in the comments. However, the result shows that the information filtering system of social media does not have too much influence on users' attitudes, and compared with search engines, social media platforms' comments do not lack diversity.

## 5. DISCUSSION

Conjecture 1: The number of user comments on Weibo is greater than the number of user comments on Baidu for the same topic.

Based on the web data crawled, the author found that the number of comments from Weibo users on the same topic can be up to 2500 times higher than that of Baidu users. Due to the more flexible form of interaction on Weibo, including likes, retweets, favorites and other behaviors, the actual number of users involved in topic-specific discussions is much higher than the number of Baidu users. It verified the author's suspicion and prompted me to wonder about the reasons for this phenomenon. After the interviews, the author found that while respondents were less likely to participate in discussions in Weibo, especially under popular Weibo topics, they were even less likely to leave their opinions under content found by Baidu searches. This is mainly because most respondents use search engines as a tool to obtain direct information rather than a platform to generate communication or interaction with others. Thus, the author believes that most users with higher education levels are less willing to express their opinions on hot topics on the Internet, and if they do want to express their opinions, they will not choose to discuss them under the results presented by search engines.

Conjecture 2: Microblog users are more consistent than Baidu users in commenting on the same topic.

Since most of the social media software represented by Weibo uses algorithms to select content recommendations for users, it contributes to a certain extent to the greater influence of filter bubble. Therefore, the author conjecture that users who share the same opinion on a specific topic are more likely to appear on Weibo at the same time and influence other users, creating a silent spiral effect leading to a more consistent attitude in the comment section. On the contrary, Baidu as a search engine, users mostly want to get more objective or truthful answers, which in turn motivates users to constantly put forward their own opinions. In

theory, this will make the attitude of Baidu's comment section richer and more people holding different views.

Conjecture 3: Among different topic types, the number of negative comments from users of political topics is greater on both platforms.

With the advent of the Internet era, new media communication has changed the original mode of information dissemination, and users are able to obtain information from more perspectives. The once more sensitive political topics are also openly available in cyberspace. To a certain extent, the emergence of political topics confirms that Chinese netizens' political participation is gradually increasing. The author believe that political topics are related to the livelihood issues of Chinese people, and such topics that are closely related to citizens' lives are more likely to trigger users' attention and interaction. Based on the 560 popular topics collected, the author found that about 38% of the topics are related to political topics, and in such topics, the number of user comments is small compared to other showbiz celebrity topics, but the sentiment analysis is complicated. Not only are there positive, negative or neutral attitudes in the interactive discussions caused by such topics, but at the same time there are some users who take a self-deprecating approach to share their views. For example, under the topic of an official's corruption and bribery, the trending comment was interestingly "I can't even dream of such a large amount". This comment appeared in the top comments on Weibo, bringing to the forefront the function of social media as an entertaining pastime tool. During the interviews, most users said they were more critical in their comments when it came to politically sensitive topics. In particular, one civil servant interviewee was very cautious or even wished to avoid talking about such topics. It is concluded that compared with entertainment topics, netizens are less likely to make strong negative comments and choose carefully whether to participate in discussions on political topics or not.

## 6. CONCLUSION

This paper focuses on the diversity of user comments in two mainstream information sharing and interaction platforms in China. Based on this study, comment diversity refers to users' negative, positive or neutral attitudes. The author crawl the popular search lists on both Weibo and Baidu, the representatives of social media platform and the search engines, and combine them with popular comments to perform text analysis and sentiment analysis. Firstly, from the perspective of the top trending lists, there is little intersection of user attention between search engines and social media. The top search lists on the same platform change every day, and users' attentions are different every day. Secondly, from the viewpoint of comment diversity, more than half of the user comments in social media platforms and in search engines show positive attitudes. Although there are more

negative comments from users under search engines, they are not representative because the sample size is too small compared to social media platforms. A more generalized qualitative analysis would require subsequent scholars to collect more information for analytical argumentation. This quantitative study fills in the attitudinal comparison of user comments in different fields and provides a direction for subsequent research. On the other hand, from the perspective of qualitative research, the author shallowly analyze users' usage of different mobile products through interviews. This analysis corroborates that users clearly know the characteristics of different information platforms and explains why the number of user comments within search engines is low but the attitude distribution is largely consistent with that of Weibo.

First, from the perspective of the research subject, this paper tries to find the difference between user comments in two types of fields, search engines and social media. Due to the limitation of time and data processing volume, this paper only investigates the two most popular platforms in China as an example. In fact, there are many mobile applications with user comments, including news APPs and other integrated media APPs, and scholars can try to compare and analyze from other application products in future studies. Second, in terms of research content, although this paper distinguishes three types of sentiment analysis for user comments, the complexity of information encoding and decoding is difficult to be fully processed and analyzed by algorithms. In the case of sufficient time, the author suggest that the part of user sentiment analysis can be completed by two or more scholars with experience in sentiment analysis and information decoding. This can effectively increase the credibility of the text analysis and positively influence the final analysis results. Finally, from the perspective of research methods, although this paper involves both quantitative and qualitative analysis methods, the small sample size of data due to time constraints largely affects the accuracy of the quantitative analysis results. At the same time, the small number of interviews and the concentration of age groups also caused the problem of insufficient depth of analysis and the lack of generalizability of the qualitative analysis results.

## REFERENCES

- [1] Carr, C. T., & Hayes, R. A., Social media: Defining, developing, and divining. *Atlantic journal of communication*, 2015, 23(1), pp.46-65.
- [2] Weibo User Development Report, 2021. <https://finance.sina.com.cn/tech/2021-03-12/doc-ikntiak9143019.shtml>
- [3] Hsueh, M., Yogeewaran, K., & Malinen, S., Leave your comment below: Can biased online comments influence our own prejudicial attitudes and behaviors?, *Human communication research*, 2015, 41(4), pp.557-576.
- [4] Walther, J. B., DeAndrea, D., Kim, J., & Anthony, J. C., *The influence of online*, 2010.
- [5] Zhao, J., & Li, Y., Influence of emotional expression in online comments on consumers' perception. *Journal of Ambient Intelligence and Humanized Computing*, 2021, pp.1-10.
- [6] Rashid, Y., & Zeeshan, M., Customer attitude towards online ads of smartphone brands::A netnographic analysis of user generated comments on YouTube, *Journal of Management Sciences*, 2018, 5(2), pp.40-64.
- [7] Shin, S. Y., Seo, D. W., An, J., Kwak, H., Kim, S. H., Gwack, J., & Jo, M. W., High correlation of Middle East respiratory syndrome spread with Google search and Twitter trends in Korea, *Scientific reports*, 2016, 6(1), pp.1-7.
- [8] Siliverstovs, B., & Wochner, D. S., Google Trends and reality: Do the proportions match? Appraising the informational value of online search behavior: Evidence from Swiss tourism regions, *Journal of Economic Behavior & Organization*, 2018, 145, pp.1-23.
- [9] Zhang, L., Zhao, J., & Xu, K., Who creates trends in online social media: The crowd or opinion leaders?, *Journal of Computer-Mediated Communication*, 2016, 21(1), pp.1-16.
- [10] Lewandowski, D., Web searching, search engines and Information Retrieval, *Information Services & Use*, 2005, 25(3-4), pp.137-147.