

Anonymous Social Network Ecological Status and Governance Strategies

Zhi-Ye LIN, Yong-Jing WANG

*Department of Creative and Design, Fuzhou University ZhiCheng College, Fuzhou, China
zylin91@163.com, yunkim.wang@hotmail.com*

Abstract—Anonymous social network has attracted ever-increasing attention from research community in recent years, however, it also brings some complications to ours daily life. We want to clarify the status of anonymous social network and put forward some governance strategies. In this paper we focus our attention on analyzing shortcomings of current regulatory strategies by comparing and analyzing the data from anonymous social network application Secret and social network application Sina Weibo. Combining with the characteristics of information dissemination in anonymous social network, we mainly put forward an governance strategy based on geographical location.

1 Introduction

Anonymous social network is attracting more and more people's attention and use, Secret and Whisper have become popular applications in young people. But the current laws and managements are based on real-name system[1], how to monitor and guide the negative information effectively in anonymous social network has become an urgent problem.

2 Ecological Status

To clarify the status of anonymous social network, we crawled data of the anonymous social application Secret from October 25, 2015 to January 24, 2016, including 1179 hottest-messages and 1200 ordinary messages. And we obtained the Sina data through the official APIs from May 3, 2014 to May 11, 2014, including 500 hottest-microblogs and 84168 ordinary-microblogs. We use the constructed emotional dictionary[2] to analyze text emotional tendencies and text themes of the anonymous networks by comparing and analyzing the data from anonymous social network application Secret and social network application Sina Weibo, and the result as shown in table I.

TABLE I. TEXTUAL EMOTIONAL TENDENCY AND THEMES IN SECRET AND SINA WEIBO

LIWC Dictionary	Secret (%)	Sina Weibo (%)	Example Words
Positive emotion	10.2	29.6	hope, love
Negative emotion	39.8	16.8	anger, hate
Sadness	19.6	9.0	sad, cry, grief
Anger	16.4	4.0	why, kill, anger
Sexual	16.3	0.7	love, incest

From table I, we can figure out that the textual emotional tendency and themes of messages are much different in anonymous social network and real-name social network, i.e., social network. Especially in pornographic, nearly 16.3% of text is related to sexual in Secret, in contrast, only 0.7% in Sina Weibo, over staggering 20 times, and other negative emotions are not optimistic too. What interesting is, the proportion of the positive emotion and negative emotion is opposite in domestic and foreign applications. The information characteristics of foreign anonymous social application Whisper and social application Twitter are shown as table II[3], the ecological status is much better.

TABLE II. TEXTUAL EMOTIONAL tendency AND THEMES IN WHISPER AND TWITTER

LIWC Dictionary	Whisper (%)	Twitter (%)	Example Words
Positive emotion	42.1	33.9	love, nice, sweet
Negative emotion	28.8	16.1	hurt, ugly
Sadness	6.5	3.3	grief, sad
Anger	14.1	8.7	anger, kill
Sexual	19.8	9.1	love, incest

From table II, we can figure out that the foreign anonymous social application (Whisper), 42.1% text have positive emotions, only 28.8% of the emotions are negative. The overall ecology tends to be positive, which is totally different in Secret. We can not deny the existence of cultural differences, but we believe that the main reason is resulted from the anonymous social application platform itself. A lack of correct guides will lead to more people to generate negative information, which exacerbate the negative environment of the network. And we will conduct a experiment to confirm the opinion in our further research. The Secret is caught in such a whirlpool, which is a basic characteristic of the current domestic anonymous social application.

3 Regulatory Incompatibility

In social networks, the identity of the network and the reality are eventually unified. However, the anonymous social networks broke the unity of identity. In anonymous social networks, people have a virtual identity which has nothing to do with reality. Therefore, the information regulatory strategies of the real-name system appear incompatible, and we concluded several reasons.

3.1 The absence of identity

The society have established a better legal system for the possible problems under the real-name network, e.g., people who forward defamation messages 500 times will be judged. But the premise of the law is that we can identify everyone which is just missing in anonymous social networks. What we should know is that the law restricts people's identity not people themselves actually.

3.2 The lack of opinion leader

We can intervene or guide the information dissemination through critical nodes such as opinion leaders under social networks, because these critical nodes have huge node entry levels[4], which means numbers of people will accept the critical node's message. In anonymous social networks,

however, the relationship between users tends to be more equal, there is no critical nodes to guide the information dissemination. In other words, the network topology in social networks is different with anonymous social networks.

3.3 Misguidance of anonymous social platforms

The anonymous social networks are at the stage of development, so that they have lowered its user access requirements to acquire continued expansion of user volume and influence. Also, anonymous social network, i.e., Secret, is not strict in messages audit, there is still a large amount of negative information in the information that is reviewed, as shown in table I.

4 The Characteristics of Information Dissemination in Anonymous Social Networks

4.1 User access threshold is at low level

The anonymous social networks are at the stage of development have lowered its user access requirements to acquire continued expansion of user volume and influence. Anonymous social application, Secret can successfully register and publish information on the platform without having to import any valid information. Another anonymous application WuMi, which requires users to import contacts to establish initial relationships, but whether the contacts are imported does not affect the normal use.

4.2 Users tend to be equal

Anonymous social networks have blurred the identities of users, so that users tend to be equal to each other, and there is no opinion leaders in anonymous social networks. Users connect to each other by locations, interests, etc., there is no significant difference between out-degree and in-degree. Anonymous social network topology is more like a classical network model.

4.3 The information audit is not strict

Anonymous social network applications such as Secret audits information by filtering through sensitive words. However, the current information review is clearly not rigorous enough. There is still a lot of illegal information in the anonymous platform, such as pornography and insulting words.

5 Governance Strategy Based on Location

The audit mechanism of anonymous social network information are keyword filtering and semantic recognition. And the library of sensitive words is usually a default

keyword library. But the question is, the sensitive words in different region are different and changing over time. Therefore, we should select the corresponding sensitive keyword library based on location.

5.1 Sensitive word library selection based on location

Establish a sensitive word library based on geographical location: Although the cyberspace of anonymous social network is completely separated from the real world, location is the only parameter that can reflect user characteristics. There are two types of geography, one is the specific function place, i.e., hospital, bar etc. The other is the location of region, i.e., city, block etc.

For geographical location L , its sensitive word library can be expressed as a set,

$$L = \{k_1, k_2, \dots, k_n\} \quad (1)$$

k_n is the sensitive word, but these sensitive words do not appear on average, different keywords have different word frequency.

2) *Select the sensitive word library*: We have set up corresponding sensitive word libraries for all geographical locations and functional places. So how to select the corresponding sensitive library is the key of information audit. Typically users have a limited range of daily activities actually, for user U , the set of his daily geographical locations can be expressed as S ,

$$S = \{L_1, L_2, \dots, L_U\} \quad (2)$$

When user U is posting an anonymous message, the relationship between the user and the sensitive word library is T ,

$$T = \frac{m_1}{l} + m_2c \quad (3)$$

m_1, m_2 are non-zero coefficient, and l indicates the distance between the posting position and the center of sensitive word library, c means the frequency of the sensitive word. The process of selecting the sensitive word library is actually the process of maximizing of T , i.e.,

$$G = \max(T_1, T_2, \dots, T_U) \quad (4)$$

The relationship between user and sensitive word library is shown in Fig. 1.

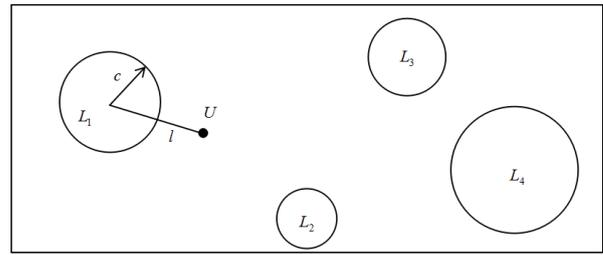


Figure 1. The relationship between user and library

5.2 Information dissemination scope control

Negative information, such as rumors, pornographic, etc., we hope to control the scope of transmission and reduce the effect to other users. For positive information, we hope to increase the spread of scope to effect more users, so that we can guide the overall network environment. The information dissemination scope should be related to the following factors: the quality of the information, the attention of information and the poster's characteristics.

1) *The quality of the information*: Information quality refers to the text emotional tendencies, keywords and themes. As judging the credibility of the information[5], we also need to refine the keywords. For message M , we refine the keywords and get set K_M ,

$$K_M = \{km_1, km_2, \dots, km_i\} \quad (5)$$

km_i is the keyword for message M , similarly, we can refine the emotion collection E_M . Therefore, the quality of information can be expressed as Q_M ,

$$Q_M = q_1 \sum K_M + q_2 \sum E_M \quad (6)$$

q_1, q_2 are non-zero coefficient.

2) *The attention of the information*: The attention means that users are interested in the information, the more people interested, the more attention the information draws. The information's attention attenuates at a certain rate over time and spread distance.

The attention of information M can be expressed as H_M ,

$$H_M = h_1x - h_2t - h_3d \quad (7)$$

h_1, h_2, h_3 are non-zero coefficient, x is the number of people who are interested, t is the time since the information posted, d is the degree of information spreads. We set threshold δ to judge whether people still interested in the information.

If $H_M \geq \delta$, means the information M still has the value of spreading, it should be kept in transmission queue. Otherwise, it means that the information has no value of dissemination and should be removed from the transmission queue.

3) *The user's characteristics*: User characteristics can be easily represented by user accounts[6]. The main factors include the quality of posts, the number of posts and the number of responses. So we can assign an initial value to each new user and adjust the value based on the user's usual usage. The users' posts and responses can be statistically calculated, and the more user posts and responses, the higher the value of the user. So we can represent the user's value by V_U .

Above all, the information dissemination scope can be expressed as

$$D_M = k_1 Q_M + k_2 H_M + k_3 V_U + D_0 \quad (8)$$

k_1, k_2, k_3 are non-zero coefficient, D_0 indicates the initial scope of information dissemination. The scope of information varies with the information quality, the attention of information and user characteristics.

6 Other Strategies

As the network environment manager, the anonymous social network platform has the responsibility and obligation to purify the network ecology. Here are some strategies.

- Anonymous platform should promote the propagation of moral and clarify the freedoms and responsibilities of users.
- Through new user entrance examination, making user understand the moral restricts.
- Using keywords filtering combined with artificial audit mechanism.
- Establishing a reward and punishment mechanism.
- Setting up emotional guides.

Conclusion

In this paper, we conducted three conclusions. First, the ecological status of social network and anonymous social network are much different in some cases, especially in negative information. Second, although we have established a better legal system for the possible problems under the real-name network, but anonymous social network lacks the basic features of real names, which makes current governance strategies show incompatibility. Third, we have conducted a comprehensive study to explore and derive the formula of the dissemination of information scope in anonymous social network. We believe the results are helpful for future researchers and related network holders. And we will do some experiment to verify our conclusions in further research.

References

- [1] P. Hangyu, "Real-name system: dealing with irresponsible microblogs," *Public Communication of Science & Technology*, 2012(13): 176-185.
- [2] Wang D L, "User character mining and analysis based on Chinese microblogging," *Beijing University Of Posts And Telecommunications*, 2015: 36-41.
- [3] Qazvinian, et al, "Rumor has it: Identifying misinformation in microblogs," *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*, 2011: 1589-1599.
- [4] W. Xindong, et al, "Analysis on the influence of online social network," *Chinese Journal of Computers*, 2014(04): 735-752.
- [5] Castillo C, et al, "Information credibility on Twitter," *proceedings of the 20th International Conference on World Wide Web*. New York: ACL, 2011: 675-684.
- [6] Wang G, et al, "Whispers in the dark: analysis of an anonymous social network," *Proceedings of the 2014 Conference on Internet Measurement Conference*. ACM, 2014: 137-150.