Research on Influence Evaluation and Application of University Weibo Users

Tong Li
School of Management
Harbin Institute of Technology
Harbin, China
litong900722@163.com

Guang Yu*
School of Management
Harbin Institute of Technology
Harbin, China
yug@hit.edu.cn

Yongtian Yu
School of Management
Harbin Institute of Technology
Harbin, China
381417821@qq.com

Abstract—In the rapid development of information age, it is an unprecedented challenge and opportunity for universities. As a new social network, Weibo provide a way to increase public impression and acknowledge to universities, it has not only a fast and convenient way to release and transfer news, but also a big scale of users. This article studied on the evaluation methods of microblog influence. We choose three methods to measure the microblog influence, include the average retweet, the H-index, and the Principal Component Analysis. We calculate separately ten sample university official Weibo users' influence base on these three method to explain their differences. We also studied on the application value of microblog influence. Further, combining the microblog influence with the quality of enrollment work, we analyzed the H-university's microblog influence and its new students, find that the average of microblog's retweet number will affect the quality of enrollment.

Keywords—University Weibo users, microblog influence evaluation, enrollment

I. INTRODUCTION

As a new type of social media, people can freely publish and obtain information through microblog, thus greatly improving the efficiency of information dissemination. In China, Weibo has penetrated into every aspect of life and become an indispensable and effective "assistant": Government can publish information and listen to public opinions through microblog, enterprises can survey users' needs and provide customer service through microblog, Weibo has become a popular choice for various organizations to communicate with the outside world. It has also been clearly pointed out in the measures on information disclosure of university issued by decree No. 29 of the ministry of education that, according to their actual conditions, universities should disclose information of all kinds of schools through various media inside and outside the university [1].

Therefore, in order to adapt to the development of society, universities have opened official microblog one after another, and the user groups of university microblog have mushroomed rapidly. As a new platform for information communication in universities, Weibo has become an important bridge between the school and the public, especially in crisis public relations, external publicity, education teaching and other aspects. Weibo has become an indispensable channel for universities to communicate with teachers and students, the examinees and the public [2,3].

The information dissemination through the Weibo platform can make hot events spread quickly, immediately attract the attention of the public, and be widely discussed. In the process of information dissemination, there is a very important internal driving force -- influence. That is to say, the influence on Weibo platform is the embodiment of a user's "status". By measuring the influence of Weibo users, the application of microblog can be further expanded. In other words, the more influential the users are, the more authority they will have and the more important they will play in the process of information dissemination [4,5]. The influence of microblog comes from the users' aggregation and integration of information resources. Therefore, the study on the influence of university's microblog will be conducive to strengthening the information integration of university's microblog, thus further improving the public's cognition and understanding of university.

In this paper, the influence of university official Weibo users is measured by three different methods, so as to further study the influence evaluation model, and to have an intuitive and quantitative understanding of the influence of university Weibo users. At the same time, universities can improve their social awareness through the Weibo platform, which plays a significant role in publicizing information, building image and improving reputation. The research on the application value of microblog influence can realize the promotion effect of Weibo platform on university management.

II. REVIEW

A. Research Status of Microblog Influence

"Influence" refers to the ability to act on the relevant subjects behaviors. In the Weibo platform, it can be regarded as the change of thoughts or behaviors caused by one user after certain information is released [6,7]. Influence is often achieved through communication.

The research on microblog influence evaluation methods can be summarized into the following three types: (1)Focusing on network topology structure. In such methods, PageRank algorithm is of great significance. PageRank is a kind of web page sorting algorithm based on link analysis, and its main method is to calculate the number of times a web page is linked to get the importance of the web page [8]. According to the characteristics of Weibo platform, it will be to evaluate the influence of users by analyzing their fans (quantity and quality). This kind of method only considers the following relationship

between users, but few factors are referenced. (2)Focusing on user interaction behavior. Since PageRank method can't fully reflect the characteristics of the interaction and communication between Weibo users from the perspective of network structure, many scholars have improved on this basis and proposed a new influence evaluation model. User influence can be seen as the driving role of users in the process of information dissemination. Thus, the dissemination characteristics can be considered as measurement indicators, including speed, range, distance, etc., and then the influence of users can be calculated through statistical analysis[9]. (3)Focusing on user attributes. Meeyoung Cha and Hamed Haddadi put forward to in-degree (the number of fans), retweets, mentions three index as evaluation factor of influence, then compared the values of the three indicators in pairs by spearman rank correlation coefficient, found that the more users are mentioned, the more they are retweeted, but the correlation between in-degree and the other two indicators is not significant.

B. Research Status of University Microblog

Discussion about university microblog already exists, Li Gangcun pointed out the definition of university microblog classification: university microblog can be divided into narrow sense and broad sense, narrow microblog refers to the only official one designated by university and generalized microblog refers to the campus microblog, such as department microblog, community microblog, and so on[10].

Scholars have also done a lot of work on the research of university microblog content. But the research on the influence of university microblog is not very thorough. Jiang Qiyan considers that the evaluation criteria for the influence of university microblog mainly include coverage (number of active fans), activity (number of microblog posts), communication ability (number of retweeted, commented, mentioned by effective fans), according to the characteristics of university microblog, the major determinants of the university microblog influence included the update frequency of microblog, the richness of information, and the personalized characteristics [11].

Based on this, scholars have made relevant researches on the influence of university microblog, such as evaluation criteria and influencing factors. However, most of them focus on the analysis and discussion level, and there are few empirical studies based on large amounts of data.

III. RESEARCH ON THE EVALUATION METHOD OF MICROBLOG INFLUENCE

A. Data

As mentioned above, university microblog can be divided into narrow sense and broad sense. The research object of this paper is the official microblog in a narrow sense, it is the only official microblog that represents the highest authority of the university. In this study, 10 universities will be selected as research samples, and the data base is from Sina Weibo platform of 10 sample universities from 00:00 on April 1, 2014 to 24:00 on March 31, 2015.

B. Evaluation of Microblog Influence Based on Average Retweet

The average retweet method is to divide the total number of microblog retweeted by the total number of microblog retweeted during a certain period by a user. This law mainly considers the advantages and disadvantages of microblog content and the quality of fans. The more a user retweets, the more he can get the sympathy of his fans, the more influence he will have, and vice versa.

The calculation formula of the average retweet is:

$$Inf(u) = \frac{\sum_{s \in S_u} R(s)}{|S_u|} \tag{1}$$

Inf(u) represents the influence score of user u, *R(s)* represents the retweet number of microblog s, *Su* is the collection of all microblogs of user u.

C. Evaluation of microblog influence based on the Principal Component Analysis

Principal component analysis (PCA) is a classical feature extraction method, which comprehensively reflects the main information contained in the original variables by finding an appropriate linear transformation rule and using several variables with large variances. Considering that microblog is a new type of social media, it can be considered from its unique information dissemination mechanism to find the relevant indicators.

We analyze several main functions of Weibo platform from the perspective of interactive information dissemination mode: post function, attention function, retweet function, and comment function. Based on the main functions of the Weibo platform, we have selected five main factors to build an evaluation index system for the influence of university microblog: the number of microblog, the number of fans, the number of followers, the number of retweet and the number of comments.

Through the correlation analysis of the collected data and calculate the contribution rate of each principal component, it can be seen that the cumulative contribution rate of the first two components accounts for 87% of the total variance. According to the principle that the cumulative contribution rate is greater than 85%, the first two feature values are selected as the main components. The initial factor loading matrix (as shown in Table I) reflects the correlation coefficient between the principal component and the original variable, the magnitude of the load on the principal component indicates the extent to which the principal component can explain the original variable.

Further, by dividing the data of the initial load matrix by the square root of the eigenvalues corresponding to each principal component, the corresponding coefficients of each index in the two main principal components can be calculated as shown in Table II.

TABLE I. PCA - INITIAL FACTOR LOADING MATRIX

	component	
	1	*2*
microblog	0.956	0.032
fans	0.204	0.866
followers	-0.149	0.869
cumulative retweets	0.971	-0.139
cumulative comments	0.949	0.061

TABLE II. PCA - THE COMPOSITION RELATION OF MAIN COMPONENT INDEXES

Indicators	Component 1	Component 2
microblog	0.569087947	0.025887392
fans	0.121437177	0.700577548
followers	-0.088696762	0.703004491
cumulative retweets	0.578017151	-0.112448359
cumulative comments	0.564920985	0.049347841

Above all, the expression of microblog influence is as follows:

$$F = \frac{2.822}{2.822 + 1.528}F_1 + \frac{1.528}{2.822 + 1.528}F_2 \qquad (2)$$

D. Evaluation of Microblog Influence Based on H-index

H-index is a bibliometrics index proposed by Hirsch, which was originally applied to academic achievements of scholars [12]. The h-index method refers to the descending order of the number of citations of a scholar's published papers (with the same number of cited papers have been given a different serial number), when and only when the number of citations of the first h paper is at least h, and the number of citations of the second h+1 paper is less than h+1, then the h index of the scholar is h. It can be said that the h-index evaluation method not only gives attention to the quantity and quality of academic achievements, but also has the characteristics of simple operation.

With fully considering the portability of this method, it is found that H-index method is suitable to measure the effectiveness of information dissemination of Weibo users. We can compare the number of retweets and comments of microblog to the number of citations of academic papers. According to the information dissemination characteristics of the Weibo platform, we transplanted H-index into two indicators, namely: retweet H-index and comment H- index.

E. Experimental Results and Consistency Analysis

The results of microblog influence of 10 sample universities obtained by the three methods are listed in Table III below:

TABLE III. RESULTS OF THREE EVALUATION METHODS

Rank	average retweet	PCA	H-index
1	Peking University	Peking University	Peking University
2	Shanghai Jiaotong University	Zhejiang University	Shanghai Jiaotong University
3	Zhejiang University	Shanghai Jiaotong University	Zhejiang University
4	Fudan University	Fudan University	Fudan University
5	University of Science and Technology of China	Xi 'an Jiaotong University	Harbin engineering University
6	Xi 'an Jiaotong University	University of Science and Technology of China	Tsinghua University
7	Harbin Institute of Technology	Harbin Institute of Technology	Harbin Institute of Technology
8	Tsinghua University	Tsinghua University	Xi 'an Jiaotong University
9	Nanjing University	Nanjing University	University of Science and Technology of China
10	Harbin engineering University	Harbin engineering University	Nanjing University

Because the above three evaluation methods respectively calculate the influence of microblog from different aspects, each method has different characteristics. According to the rank of microblog influence of 10 universities as shown in Fig. 1, it can be observed that the three methods have similar trends. Moreover, the overlap between the average retweet method and the H-index method is very large, while the principal component analysis method is slightly different from these two methods.
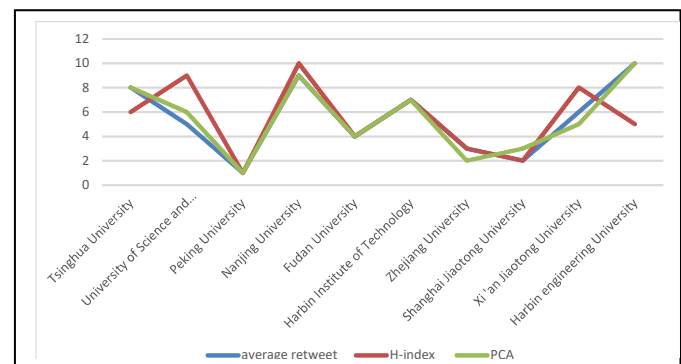

Fig. 1. Ranking trends of 10 sample universities by three methods

In order to further analyze the three methods, we used Kendall rank correlation coefficient to compare and analyze the experimental results. Kendall coefficient, also known as harmony coefficient and consistency coefficient, is a method to represent the degree of correlation between multi-level variables. This method is generally suitable to the data obtained by rating method, that is, by asking K judges to evaluate N things, or one judge to evaluate N things in K times

successively, then get a data form that is hierarchical. The final correlation coefficient results are as follows:

TABLE IV.　　Correlation of the Three Evaluation Methods

	average retweet	PCA	H-index
average retweet	1.0	0.5111	0.9111
PCA	0.5111	1.0	0.5111
H-index	0.9111	0.5111	1.0

As shown in the Table IV, the correlation coefficient between the results of the average retweet and the results of the H-index is 0.9111, indicates that the results obtained by these two methods are very similar, and there is a certain difference between the results of principal component analysis and those obtained by the other two methods. Therefore, we can analyze: the average retweet method mainly measures the ratio between the retweet number and the total microblog number, so as to reflect the quality of microblog content published by a Weibo user and whether it is enough to attract the attention of fans; the measurement of H-index method is also focused on the number of retweet and also includes the number of comments, this method examines whether the microblog content has received more attention from the perspective of quantity and quality; the principle of principal component analysis is to comprehensively investigate the microblog influence by taking two main components as the representative of the five factors. Therefore, we can find that, although the results obtained from the three methods are similar to some extent, different results are obtained through the three methods due to their respective evaluation emphases.

Based on this, we cannot arbitrarily decide which method is more effective and reasonable in calculating the microblog influence. That is to say, we cannot comment on the accuracy of the three methods in calculating the microblog influence from the above results. For the official university microblog, first of all, the university is a non-profit organization, different from the commercial marketing microblog, a unified ranking cannot explain substantive problems; Secondly, the formation factors and the process of are complex and changeable, and not all cases can use a single result. Therefore, in terms of the evaluation of microblog influence of the official university microblog, we can have an in-depth understanding of all aspects of its constituent factors and analyze them separately, which is more meaningful than a unified ranking. The purpose of this paper is to objectively study and analyze the evaluation methods of microblog influence and evaluate the university microblog influence from various aspects, it is expected to analyze and solve relevant university management affairs in the process of studying the application value, aiming at the quantitative microblog influence. We need to take into account the relatively comprehensive results of microblog influence to further explore its application value.

IV. RESEARCH ON THE APPLICATION VALUE OF UNIVERSITY MICROBLOG INFLUENCE

A. Overview of University Enrollment

The functions of a university mainly include personnel training, scientific research and service to the society. Talents are an important driving force for the university to carry out all kinds of work. The quality of students is an important factor in talents training, while the enrollment work has become an important part. In the process of filling in the application of university, the examinee does not only consider the score and the admission score line, but also makes an in-depth comparison of all kinds of information about the university. There are many channels for collecting information about a university, but nowadays the Internet information is mixed, so it is not easy to get useful information. The Weibo platform can provide a large amount of data, and has become a new channel for people to get information. It may provide more comprehensive and objective information, which can be said to be one of the important channels for students and parents to get information about universities.

Therefore, we will explore the correlation between the influence of university microblog and the quality of enrollment.

B. The Correlation Analysis between the Influence of University Microblog and the Enrollment Work

In this section, we choose H-university as the case study sample, and will focus on the analysis of the correlation between its microblog influence and enrollment quality. The total score of university entrance examination has changed in the past five years due to the different situation of university entrance examination policies in different provinces and cities. Therefore, in the process of analysis, we decided to define the variable representing the enrollment quality as the difference between the admission score line and the key score line of the province (city) in that year, and finally calculates the average difference of each province to represent the enrollment quality of H-university in that year as shown in Table V.

We have collected all the official microblog data of H-university from the date of opening (November 24th, 2011) to June 30th, 2014. According to the entry time period of university entrance examination in relevant years, we divided the microblog data by the time cut-off point, so as to calculate the microblog influence in each year (as shown in Table VI).

Next, the correlation analysis between the results of the three microblog influences and the quality of the students is conducted successively.

According to the results in the Table VII, only the correlation analysis Sig value between the microblog influence calculated by the average retweet method and the quality of the students is less than 0.05, it is considered that the result is significant and the correlation coefficient is 0.969, indicating that it is very relevant. Therefore, we can see that there is a correlation between the quality of students and the microblog influence, and the higher the average retweet rate is, the higher the quality of students will be in that year.

Universities can propagandize themselves and attract more people to pay attention to universities through the Weibo platform. And the higher the microblog influence, the better the effect of its publicity, the more people can attract more attention to the university, thus increasing the number of applicants. According to the results obtained, in terms of the enrollment work, we can effectively dissemination the official microblog information of universities, that is, improve the retweet rate of their microblog, and enable more people to receive microblog information, so as to influence people's cognition of universities, and further play the role of Weibo platform in publicity in the enrollment work.

TABLE V.   THE AVERAGE DIFFERENCE BETWEEN THE ADMISSION SCORE LINE AND THE KEY SCORE LINE IN EACH PROVINCE (CITY) OF H-UNIVERSITY

province /city	2011	2012	2013	2014
Average difference in five years	90.37	100.81	108.42	112.78

TABLE VI.   THE MICROBLOG INFLUENCE OF H-UNIVERSITY FROM 2010 TO 2014

Year	average retweet	PCA	H-index
2011	11.41025641	50	-2.14006285
2012	16.09385965	93	-0.34846686
2013	26.56222548	129	1.959202325
2014	27.49619903	115	0.529310239

TABLE VII.   THE CORRELATION BETWEEN THE MICROBLOG INFLUENCE AND THE QUALITY OF STUDENTS OF H-UNIVERSITY

	average retweet	PCA	H-index
Pearson Correlation	0.969*	0.936	0.863
Sig. (2-tailed)	0.031	0.064	0.137
N	4	4	4

V. CONCLUSIONS

This paper analyzes the microblog data of 10 universities, and evaluate microblog influence by three methods of these 10 universities respectively. The results of microblog influence are obtained through ranking; then to analyze the results by Kendall rank correlation coefficient, and the conclusion was drawn that the average retweet method and H-index method had a high consistency, while the correlation between the principal component analysis method and these two methods was not very significant.

By studying the correlation between the university microblog influence and the quality of students, we find that the retweet rate of university microblog is highly correlated with the quality of students.

In future studies, the university microblog influence can also be correlated with other work in university management, such as the student employment, student satisfaction, and gathering alumni resources and so on. It is of great significance to optimize university management to build microblog influence to assist university management affairs, so as to provide more channels and solutions for university managers to solve problems in the information era.

REFERENCES

[1] "The measures on information disclosure of university issued by decree No. 29 of the ministry of education," http://www.ahedu.gov.cn/162/view/5076, April 2015.

[2] Rosmala D., "Study of Social Networking usage in Higher Education Environment," Procedia-Social and Behavioral Sciences, no. 67, pp. 156-166, 2012.

[3] Bicen H, Cavus N, "Twitter usage habits of undergraduate students," Procedia-Social and Behavioral Sciences, no. 46, pp. 335-339, 2012.

[4] Tiago Rodriguesy, Meeyoung Cha, "Characterizing User Behavior in Online Social Networks," IMC, pp. 54-58, 2009.

[5] Dejin Zhao, Mary Beth Rosson, "How and Why People Twitter: The Role that Micro-blogging Plays in Informal Communication at Work," GROUP of ACM, pp. 247-249, 2009.

[6] Sun J, Tang J, "A survey of models and algorithms for social influence analysis," Social Network Data Analytics, Springer US, pp. 177-214, 2011.

[7] Zhou Y, Chen K, Song L, et al, "Feature Analysis of Spammers in Social Networks with Active Honeypots: A Case Study of Chinese Microblogging Networks," Advances in Social Networks Analysis and Mining(ASONAM), IEEE/ACM International Conference on. IEEE, pp. 728-729, 2012.

[8] Kamvar S, "Extrapolation Methods for Accelerating PageRank computations," USA: Stanford University, 2003.

[9] Zhang Y, Li X, Wang T W, "Identifying Influencers in Online Social Networks: The Role of Tie Strength," International Journal of Intelligent Information Technologies (IJIIT), vol. 9, no. 1, pp. 1-20, 2013.

[10] Li Gangcun, "A brief analysis of the dissemination and operation of university microblogs," News Research, 02, pp. 67-68, 2012.

[11] Jiang Qiyan, "The Influence of the Official Microblogs in Higher Institutions and Promotion Strategies," Journal of Mudanjiang College of Education, no. 05, 75-76, 2012.

[12] Hirsch J E, "An index to quantify an individual's scientific research output," Proceedings of the National academy of Sciences of the United States of America, vol. 102, no. 46, p. 16569, 2005.