An evaluation method of government micro-blogging service quality based on user behavior analysis

Suozhu Wang\textsuperscript{1*}, Tianhao Guo\textsuperscript{1} and LinQi Gao\textsuperscript{2*}

\textsuperscript{1}School of Management, Capital Normal University, Beijing, 100048, China
\textsuperscript{2}School of Management, Tianjin Normal University, Tianjin, 300387, China
\textsuperscript{*}Corresponding author’s e-mail: wsz_wsz@163.com, gaolqi@hotmail.com

Abstract: As a new channel and way for the government to provide public services, how to accurately evaluate the service quality is the key to its healthy and rapid development. The government micro-blogging service quality evaluation index system is constructed and the calculation methods of the evaluation indexes are given in this paper. It contains four dimensions, namely, government micro-blogging construction quality, government micro-blogging functional service, user interaction behaviour and benefit-cost. On this basis, an evaluation process and method of government micro-blogging service quality that integrates subjective and objective indexes is proposed based on the idea of fuzzy comprehensive evaluation in this paper.

1. Introduction
With the popularization and application of information technology and mobile Internet, the development of government micro-blogging has made unprecedented progress. Government micro-blogging has become a new channel and way for the government to provide public information services and public information products because of its low technical requirements for users, ease of operation, and low content restrictions. According to the statistical data, there are 179932 government micro-bloggings in China\textsuperscript{[1]}, but the actual application effect has not reached the expected effect. The general problems of government micro-blogging are the phenomenon of empty shell microblogs, poor service, poor interaction between the government and the people, etc., which leads to its low actual benefit. However, the root of these problems lies in the lack of a correct understanding of the objectives of the construction of government micro-blogging, and the lack of a set of scientific methods to measure the quality of government micro-blogging. Therefore, how to evaluate the service quality of government micro-blogging accurately and comprehensively is the key to promote its rapid and steady development.

With the increasing influence of government micro-blogging, its research has become the focus of the academic community. From the perspective of domestic and foreign research, the research on government micro-blogging mainly focuses on the propagation mechanism of microblogs, social function of microblogs, the role of microblogs in democratic supervision, influence mechanism of microblogs on public policy, and how to improve the image of government and officials by government micro-blogging\textsuperscript{[2]}. In foreign countries, the United States has provided government services on Twitter to provide new communication tools for the interaction between the government and the people since 2006; the United Kingdom promulgated "Twitter use guide" for government departments, which stipulates the purpose and evaluation method of government micro-blogging, risk
management in the use process, effective management and use of microblog, etc. in 2009. In China, because the government micro-blogging is developed from the Internet industry, the initial research on government micro-blogging is mostly from the perspective of practicality, in order to find out the current status of the development of government micro-blogging, the paper puts forward the main problems faced by the development of government micro-blogging and corresponding policy suggestions[4]. At present, the academic research on government micro-blogging is still in its infancy, mainly focusing on the recognition and application of microblogs, and has not yet formed a systematic theoretical framework, and the literature on service quality evaluation of government micro-blogging is still a minority.

Research on the evaluation of the service quality of government micro-blogging is conducive to the government to understand and master the current development of government micro-blogging, and find out the shortcomings, so as to objectively make improvements and make it truly serve the people, so that it can truly serve the people. Because government micro-blogging is a new political media, relevant research at home and abroad is still in development, and has not formed a systematic theory and method to evaluate the service quality of government micro-blogging. There are only several papers about the evaluation of the service quality of government micro-blogging in China[4][5]. The research results in this area are still a minority, and new theories and methods need to be put forward for further research. The evaluation of the service quality of government micro-blogging involves the influence of multiple index factors. How to build a reasonable index system and the measurement method of the index is the key to make the evaluation scientific and reasonable. Therefore, a simple evaluation index system from the aspects of the construction quality of government micro-blogging, user behavior analysis, government micro-blogging function, and benefit-cost is built in this paper. The model of government micro-blogging service quality evaluation based on fuzzy comprehensive evaluation is also given in this paper, which provides new ideas and methods for the comprehensive evaluation of government micro-blogging service quality.

2. Government micro-blogging service quality evaluation index system

2.1. Construction of evaluation index system

The goal of government micro-blogging is to provide various government services for the public, which is a main feature different from other microblogs. The service quality of government micro-blogging refers to the difference between the service level perceived by users and the service level expected by users[6]. Therefore, the evaluation of government micro-blogging service quality should not only analyze the results of microblog service and user perception, but also consider the behavior of users. Users interact with government micro-blogging through follow, retweet, comment, and other usage behaviors, which indirectly reflect the service quality of government micro-blogging. If the service quality of a government micro-blogging is good, users are willing to use the platform, otherwise, if the service quality of a government micro-blogging is poor, it is not attractive to users. Therefore, the analysis of the service quality of government micro-blogging is a systematic project involving many factors. When constructing the index system, not only should it be designed from the perspective of the actual functions of government departments and the superiority that government microblogs can bring, but also the practical problems such as the availability of data and the size of workload in the evaluation of government micro-blogging should be considered. Therefore, an evaluation index system with strong operability is constructed in this paper. It includes the construction quality of government micro-blogging, the function of government micro-blogging, user interaction behavior, and benefit-cost. As shown in Table 1.

| Table1. Government micro-blogging service quality evaluation index system. |
|---------------------------------------------------------------|
| First grade index | Secondary index | Index data acquisition path |
| Construction quality (U$_1$) | Interface design friendliness (U$_{11}$) | Expert evaluation mark |
2.2. Description of evaluation index

(1) Interface design friendliness (U_{11}). This indicator reflects the rationality of government micro-blogging interface design, the unity of content and form, the scientificity of color matching, and the simplicity and convenience of operation. The design of government micro-blogging interface design is an important way to show the government’s image and introduce the government functional services. Therefore, to evaluate the service quality of government micro-blogging must first evaluate the design of the interface. Because the interface design is subjective evaluation, there is no uniform and formal standard to measure, so it is generally obtained by the method of hierarchical comprehensive scoring by experts in the industry.

(2) Page link accuracy rate (U_{12}). This indicator reflects the frequency of correct response of microblog page links. In the design of microblog, there will be some unreasonable structure or link errors, which make users report errors on the page and affect users' reading. Therefore, the higher the accuracy, the better the construction quality of microblog. This indicator can be calculated with the aid of traffic analysis tools (such as WebTrends, Click-Tracks, etc.). Let E_p denote the number of pages with errors and T_p denote the total number of pages visited, then U_{12} can be calculated by the following formula.

$$U_{12} = \frac{T_p - E_p}{T_p}$$  \hspace{1cm} (1)

(3) Convenience and flexibility of operation (U_{13}). Refers to the ease of use and flexibility of microblog functions. This indicator is the subjective evaluation content, and there is no uniform and formal standard to measure it. Therefore, it is generally obtained by the method of hierarchical comprehensive scoring by experts in the industry.

(4) Microblog update frequency (U_{21}). This indicator can be reflected by the average number of daily microblogs published in the statistical period, reflecting the update frequency of information.
resources of government micro-blogging. Providing government information is the primary function of a government micro-blogging. The government should provide a large number of news and government information every day, so that the people can grasp the latest trend in time. Therefore, the faster the update frequency of microblog is, the timelier the information it provides, and the more valuable it can be used. Let $T_w$ denote the total number of microblog postings on the $D_1$ to $D_2$ day, then $U_{21}$ can be calculated by the following formula.

$$U_{21} = T_w / (D_2 - D_1)$$ (2)

(5) Microblog private message response rate ($U_{22}$). This indicator refers to the proportion of original microblogs in all microblogs published by government micro-blogging in the statistical period, reflecting the innovation and activity of government departments in using microblogs to ask for politics. In a certain statistical period, let $T_w$ denote the total number of microblogs published, $F_w$ denote the number of retweeted microblogs, then $U_{22}$ can be calculated by the following formula.

$$U_{22} = (T_w - F_w) / T_w$$ (3)

(6) Microblog comment response rate ($U_{23}$). This indicator refers to the rate of government micro-blogging users actively responding to comments in the statistical period, reflecting the activity of interaction between government departments and the public. In a certain statistical period, let $C_w$ denote the number of microblog comments received by government departments, $R_w$ denote the total number of replies, then $U_{23}$ can be calculated by the following formula.

$$U_{23} = R_w / C_w$$ (4)

(7) Microblog private letter response rate ($U_{24}$). This indicator refers to the proportion of private letters sent by government departments to other users in the statistical period, reflecting the affinity between government departments and the public. In a certain statistical period, let $L_w$ denote the number of private letters received by government departments, $L_R$ denote the total number of private letters replied, then $U_{24}$ can be calculated by the following formula.

$$U_{24} = L_w / L_R$$ (5)

(8) Fans/followers ratio ($U_{31}$). This indicator refers to the proportion of fans of government micro-blogging in other users' attention, reflecting the degree of the public's active attention to government agencies. If the service quality of government micro-blogging is good, the number of its fans will be more. On the contrary, if the service quality of government micro-blogging is poor, the number of its fans will be small. In a statistical period, let $F_o$ denote the number of followers of government micro-blogging, $F_o$ denote the number of users it focuses on, then $U_{31}$ can be calculated by the following formula.

$$U_{31} = F_o / F_o$$ (6)

(9) Average daily retweeting rate of microblog ($U_{32}$). This indicator refers to the average number of times that microblog is retweeted on the government micro-blogging every day. Generally speaking, the first factor to examine whether the content provided by the government micro-blogging is attractive is to see the number of times it is retweeted. If its service quality is close to the public demand, its daily average retweeting times must be very high. Let $T$ denote a statistical period, the set of all published microblogs is $w = \{w_1, w_2, \ldots, w_n\}$, and $F_{w_i}$ denote the times of $w_i$ ($i=1,2,\ldots,n$). Then $U_{32}$ can be calculated by the following formula.

$$U_{32} = \sum_{i=1}^{n} F_{w_i} / nT$$ (7)

(10) Average daily commenting rate of microblog ($U_{33}$). This indicator refers to the average number of times that microblogs on the government micro-blogging are commented every day, reflecting the degree of users' attention to the government micro-blogging. Generally speaking, the content provided by the government micro-blogging can stimulate the public's interest, and the public
will actively participate in retweeting and commenting. Accordingly, the daily average number of comments on microblogs must be very high. Suppose that in the statistical period $T$, the set of all microblogs published is $w = \{w_1, w_2, \ldots, w_n\}$, and $C_{w_i}$ is used to express the times of $w_i$ ($i=1,2,\ldots,n$) being commented, then $U_{33}$ can be calculated by the following formula.

\[ U_{33} = \frac{\sum_{i=1}^{n} C_{w_i}}{nT} \quad (8) \]

(11) Average daily collection rate of microblog ($U_{34}$). This indicator refers to the average number of daily collection of microblogs on the government micro-blogging, reflecting the practicability and available value of the information provided by the government micro-blogging. If the resources provided by government departments are practical, the public will take the initiative to collect them. Suppose that in the statistical period $T$, the set of all published microblogs is set as $w = \{w_1, w_2, \ldots, w_n\}$, and $C_{w_i}$ is used to express the times of $w_i$ ($i=1,2,\ldots,n$) being collected, then $U_{34}$ can be calculated by the following formula.

\[ U_{34} = \frac{\sum_{i=1}^{n} C_{w_i}}{nT} \quad (9) \]

(12) Average daily praising rate of microblog ($U_{35}$). This indicator refers to the average number of times that microblogs on the government micro-blogging are praised every day, reflecting netizens' recognition of the services provided by the government micro-blogging. If the service provided by the government micro-blogging can effectively solve the problems of public concern, it will be highly recognized by netizens. Suppose that in the statistical period $T$, the set of all published microblogs is set as $w = \{w_1, w_2, \ldots, w_n\}$, and $L_{w_i}$ is used to express the times of $w_i$ ($i=1,2,\ldots,n$) being praised, then $U_{35}$ can be calculated by the following formula.

\[ U_{35} = \frac{\sum_{i=1}^{n} L_{w_i}}{nT} \quad (10) \]

(13) The ratio of average daily cost savings to management and maintenance costs ($U_{41}$). This indicator refers to the ratio between the cost saving of government departments and the operation and maintenance cost of microblog after using government micro-blogging. This indicator can be calculated through financial statistics. Because one of the purposes of using government micro-blogging is to save costs, facilitate office work and improve performance, so the quality of efficiency is another important standard to measure the quality of government micro-blogging.

3. Comprehensive evaluation process based on user behavior analysis

Generally speaking, the amount of data used for evaluation is large, especially the data from web logs. Therefore, before comprehensive evaluation, data preparation and data analysis are required. The comprehensive evaluation process of the service quality of government micro-blogging is shown in Figure 1.
3.1. Data collection and analysis
The task of this stage is to collect evaluation data from different aspects, including web log data, user behavior data, expert grading results data, and financial statistics data.

(1) Web log data. The data in the web log is very complex, there is no specific model to describe it, and there are often data missing and errors, so it is impossible to obtain operable data directly from it. In order to ensure the accuracy and integrity of the data, it is necessary to carry out data cleaning, data integration, data transformation, data reduction[7] and other preprocessing work. A standard data set is obtained, which is convenient for the next step of data analysis.

(2) User behavior data collection. Use the web crawler tool to collect the user behavior data such as the number of fans, the number of retweets, the number of comments, etc. on the government microblogging, and carry out standardized processing to prepare for the subsequent calculation of the value of relevant indicators.

(3) Other data preparation. a) Use the hierarchical comprehensive scoring template to obtain the expert scoring results \( U_{11} \) and \( U_{13} \), \( (U_{11}, \ U_{13} \subseteq [0, 100]) \); b) The financial department calculates and gives the benefit indicator \( U_{41} \).

3.2. Data analysis
In this stage, the user behavior data set obtained from the preprocessed web data and crawler is analyzed and sorted into the data format needed to calculate each indicator. According to the formula (2-1), (2-2), (2-3), (2-4), (2-5), (2-6), (2-7), (2-8), (2-9), (2-10) the values of evaluation indexes \( U_{12}, U_{21}, U_{22}, U_{23}, U_{24}, U_{31}, U_{32}, U_{33}, U_{34}, \) and \( U_{35} \) are calculated respectively, and the results of relevant indexes are obtained.

3.3. Comprehensive evaluation
The main task of this stage is to integrate the subjective evaluation data and objective evaluation data.
to obtain the comprehensive evaluation value of the service quality of government micro-blogging. The specific steps are as follows.

1) Dimensionless. In the above calculated index values, their dimensions are different, and the difference between the values is also large. Therefore, it is not appropriate to directly carry out weighted average, and it has no practical significance. Therefore, these values need to be dimensionless processing, that is, the values are all mapped into the interval [0,1] to obtain an evaluation score, and then these scores are worth the comprehensive score of the government micro-blogging. The specific method is as follows.

For the evaluation index \( u_{ij} \), let its domain be \( d_{ij} \in [m_i, M_i] \), where \( m_i \) and \( M_i \) are the minimum and maximum values of the evaluation index \( u_{ij} \) respectively. Then the formula can be used to normalize each index value.

\[
    r_i = \begin{cases} 
        1 & u_{ij} > M_i \\
        \frac{u_{ij} - m_i}{M_i - m_i} & m_i \leq u_{ij} \leq M_i \\
        0 & u_{ij} < m_i 
    \end{cases} \tag{11}
\]

2) Determine the comment set. The comment set is a description of the merits and demerits of evaluation objects. The comment set cannot be too much or too small. Too many comment sets are not easy to master the standard. Too few comment sets are not conducive to distinguishing the quality of evaluation. Suppose that the comment set of government micro-blogging service quality is as follows: \( Y = \{ \text{very poor, poor, medium, good, excellent} \} = \{ Y_1, Y_2, Y_3, Y_4, Y_5 \} \).

3) Determine the weight set of evaluation factors. Because each index factor has different influence on the evaluation, a factor importance set \( A = \{ q_1, q_2, \ldots, q_n \} \), \( q_i \in [0,1] \), \( \sum_{i=1}^{n} q_i = 1 \) is given according to its importance. The weights can be set by the adjacent index comparison method, Delphi method, statistical method, etc.\cite{8}.

4) Comprehensive score. Through the above steps, the evaluation values with different dimensions are dimensionless, then weighted average is carried out, and the final evaluation result expression \( S = \sum_{i=1}^{n} q_i \cdot r_i \) is obtained. Then, the final evaluation value \( S \) of the government micro-blogging service quality is classified according to the numerical value, which is the final evaluation result. The specific classification can be determined by the actual situation, for example.

(i) \( 0 \leq S < 0.2 \), the rating \( Y_1 \) is given as “very poor”;  
(ii) \( 0.2 \leq S < 0.4 \), the rating \( Y_2 \) is given as “poor”;  
(iii) \( 0.4 \leq S < 0.6 \), the rating \( Y_3 \) is given as “medium”;  
(iv) \( 0.6 \leq S < 0.8 \), the rating \( Y_4 \) is given as “good”;  
(v) \( 0.8 \leq S < 1.0 \), the rating \( Y_5 \) is given as “excellent”.

4. Conclusion

The evaluation of the service quality of government micro-blogging is a new research topic. At present, a unified evaluation index system has not been formed. How to build a scientific and reasonable index system and evaluation methods are worthy of in-depth study. In terms of the collection of evaluation index data, if only the user questionnaire or expert grading is used to obtain the index data, such evaluation results may lack of objectivity. Therefore, in order to accurately and comprehensively evaluate the service quality of government micro-blogging, the evaluation indicators must be diverse, and cannot be determined by a single method, and the combination of subjective and objective indicators is needed to be adopted. Based on this, a comprehensive evaluation method of the service quality of government micro-blogging is proposed in this paper. In the design of the index system, the evaluation indexes are divided into two categories: objective and subjective. The subjective indexes are obtained according to the expert grading method, and the objective indexes are collected by means of Web log analysis, web crawler tools, statistical tools and financial statistics.
Finally, the evaluation results of the service quality of government micro-blogging are obtained by synthesizing these indexes. Of course, in the field of government microblog service quality evaluation research, an exploratory study is made in this paper, and a large number of examples are still needed to verify the effectiveness and rationality of the method proposed in this paper.

**Acknowledgements**

This research was supported by National Natural Science Foundation of China (grant No. 61371177), and the Scientific Research Foundation of Beijing Municipal Education Commission (grant No.KM201810028021).

**Reference**

[1] People’s Network Data Public Opinion Center., (2019)2019 Government Affairs Index Weibo Impact Report. http://yuqing.people.com.cn/n1/2020/0117/c209043-31553643.html.

[2] Wang, F.S. (2013) A Summary of Research on Chinese Government Affairs Weibo. E-government, 9:27-34.

[3] Xiong, J., Liu, A.(2013) A Summary of Research on Government Affairs Weibo. Knowledge Economy, 3:73-73.

[4] Bao, M.L., Liu, R. (2015) The Study on Government Microblog Service Quality Evaluation Index System. Journal of Modern Information, 35(9):93-97+110.

[5] Zou, K., Bao, M.L. (2016) An Empirical Study on Government Microblog Services Public Satisfaction Index Model. Journal of Xiangtan University (Philosophy and Social Sciences), 40(1):75-80.

[6] Hu, Y. C., Wang, J. H., Hung, L. P. (2012) Evaluating Microblogging e-Service Quality using ANP. Journal of Multi-Criteria Decision Analysis, 19:89-111.

[7] Chen, A., Chen, N., Zhou, L.X. (2006) Data mining technology and application. Science Press, Beijing.

[8] Hu, Y.H., He, S.H. (2000) Comprehensive evaluation method. Science Press, Beijing.