Qualitative and Quantitative Methods in Libraries (QQML) 5: 221- 234, 2016

New methods to quantify an individual's scientific research output based on h-index and various factors

Weifan Zhao, Wenjun Liu, and Ruiyu Zhang

College of mathematics and statistics, Nanjing University of Information Science and Technology, Nanjing 210044, China

Abstract: Hirsch proposed h-index in 2005. To overcome some limitations of h-index in practical application, many scholars have proposed a series of derivative indexes. So far, derivative indexes designed by the scholars are mostly based on the citations and total number of papers. Although they are simple to use, they have certain oneness on the data source. To overcome this situation, two methods are used in this paper to analyze various factors. Firstly, we use the principal component analysis method to analyze original data such as the total number of papers, citation and so on, and then draw a comprehensive evaluation result of Z to quantify scientific research output, which avoid the oneness on the factor. Then, based on the "Gold priority" rule and h_a-index proposed by Xu, we introduce a new comprehensive scientific research evaluation index, named as h_f-index, to enhance the degree of differentiation and sensitivity of h-index. For two methods above, examples are given respectively.

Keywords: h-index, principal component analysis, h_f-index

1. Introduction

Jorge Hirsch (2005), an American physicist at the University of California, San Diego, designed a new evaluation index (h-index) in 2005 to characterize the scientific output of a researcher by measuring the impact of the scientist's publications in terms of the received citations.. As soon as h-index was proposed, h-index obtained the widespread attention of the international scientific community and became an international hotspot issue in the field of scientific metrology and science evaluation. Moreover, it is widely used to quantify the research personnel, scientific research institution, science and technology periodical. As a combination of publications and paper quality, it not

Received: 23.6.2015 Accepted: 21.12.2015 ISSN 2241-1925

only possesses some advantages that the single index incomparable, but also has its own difficulty to overcome in the practical application.

So far, many scholars proposed a series of derivative index to enhance the academic evaluation of accuracy for h-index from different aspects. Braun (2006) applied h-index to measure academic periodicals. Egghe, the famous science metrologist, thought that h-index should have reflected these highly cited papers (2006), and thus he designed g-index in 2006. At that same year, Kosmulski (2006) putted forward h_2-index, which focused on highly cited papers. In 2007, in et al (2007)] introduced the concept of "h-core", in which the amount of citation frequency is located in the top h. In addition, to overcome the lack of the degree of sensitivity, differentiation and volatility of h-index, they also introduced A-index, R-index and AR-index. However, h-index, as a performance evaluation method, its score values often gathered in the low score. To overcome this limitation, Xuemei Zhang (2007) putted forward h_m-index. In 2008, Wu (2008) proposed W-index, which inherits the advantage of h-index such as simplification and easy to understand characteristics. W-index paid more attention to the influence of the highly cited papers while h-index only calculated the number of highly cited papers. h_t-index, proposed by Anderson (2008) in the same year, took total citations into account to overcome the lack of differentiation of the h-index. In 2009, Ye (2009) introduced f-index, which was a new academic ranking index inspired by h-index. The index was linked to the subject, it combined quantity and quality, and suitable for disciplines, countries, institutions, periodicals, and Multidimensional academic scholars. In the same year, Xu (2009) proposed a more extensive scope of h_a-index under the enlightenment of the "Gold priority" rule, the index considered the size of the core of performance, measured the strength of the core of performance, improved the accuracy of index, and implemented the comparability between the scholars with the same h-index. In 2010, Alonso (2010) designed hg-index, used in the measure to the scientific research output, which not only considered highly cited papers, but also significantly reduced the effects of the highly cited papers. So, it can achieve a balance between most high quality papers and highly cited papers of the author.

So far, the indexes proposed by the scholars are mostly based on citations and the total number of papers. Although they are simple and convenient to use, it is difficult to know which index is more reasonable to quantify some scholars' scientific research output in many cases.

This paper aims to adopt some mathematical methods to analyze a variety of factors. Firstly, principal component analysis will be adopted to analyze various factors. Then, considering that there is a little complexity and complicated operation to apply the first method, a new index named as h_f-index was proposed under the enlightenment of "Gold priority" rule and h_a-index. It's shown that the proposed new index could improve the degree of differentiation and sensitivity of h-index in a certain extent, and can be simple to use.

This work will not only overcome some defects of the existing index in a certain extent, but also avoid the oneness when use citation frequency and the total number of papers as the main basis to quantify an individual's scientific research output, thereby making the evaluation more fair and reasonable.

2. A new method based on Principal Component analysis with multi-factor

2.1. Principal component analysis

Principal component analysis (1901) was introduced for non-random variables by Karl Pearson in 1901. And then in 1933, Hotelling (1933) extend it to the random vectors.

Generally, a few new variables which are less than the number of original variables can explain variation of most of the data was selected, called principal component and a comprehensive index to explain the data. Thus, the principal component analysis is actually a dimension reduction method. It has the following advantages: First, it used a few comprehensive variables to replace the original multiple variables by dimension reduction techniques, most of these comprehensive variables focused on the most information about the original variables. Secondly, it scored by calculating the comprehensive principal component function, scientifically evaluated objective economic phenomena. Again, it focused on the comprehensive evaluation of information contributions influence on the application.

2.2. Method based on Principal Component analysis with multi-factor

Since h-index was proposed, h-index was widely used to evaluate the researchers, research institutions and scientific journals. Given the many shortcomings of the h-index, scholars proposed a series of derivative index. Due to the big difference in rankings between h-index and other indexes, characteristics of the paper should be taken into account for more concrete analysis. Then it could be more just and reasonable when we measure the study performance of scientific researchers.

For most of the evaluation index, researchers proposed them mostly based on citations. These indexes had certain limitations when calculating. What's more, it was a single source of data, ignored the characteristics of the paper. In this section, we would analyze a variety of factors (Issued amount, total citations and download frequency, etc.) and using some mathematical methods to get a comprehensive evaluation value. This approach can to some extent avoid the oneness of using citations as the main basis of evaluation. As the principal component analysis method can transform the number of correlation highly variables into each independent or irrelevant variables, we will adopt principal component analysis method to analyze the various factors.

2.3. Examples of application

2.3.1. The database

The subsequent analysis is based on 20 well-known scholars' papers cited situation (in descending order of h-index). The data comes from the paper which

entitled "h-index research review and the empirical statistical analysis" (2009) by Wei and Song. It is collected from 1998~2008 Chinese journal full-text database (see Appendix).

2.3.2. The data analysis

Observing from the data in the Appendix, we notice that h-index and other indicators (total citations, issued amount, etc.) is not a simple linear correlation. Thus the index cannot substitute each other. To test this idea, the Pearson correlation coefficients between six indicators were obtained, and the results are as follows (see Table 1).

Table 1. The Correlation Matrix

	Issued amount	h-index	Total citations	The average citation rate	Download frequency	Browse frequency
Issued amount	1.0000	0.3220	0.3983	-0.5422	0.5997	0.4583
h-index	0.3220	1.0000	0.8920	0.0860	0.3988	0.3138
Total citations	0.3983	0.8920	1.0000	0.2466	0.4120	0.3174
The average citation rate	-0.5422	0.0860	0.2466	1.0000	-0.3071	-0.2814
Download frequency	0.5997	0.3988	0.4120	-0.3071	1.0000	0.9721
Browse frequency	0.4583	0.3138	0.3174	-0.2814	0.9721	1.0000

From Table 1, we can know the presence of some correlation between some indicators, such as the correlation between h-index and total frequency is 0.8920, the correlation is strong. If we don't deal with the data, that is using directly, there would be a certain cross-resistance on information aspect, thus affecting the objectivity of comprehensive evaluation value. Considering the advantages of principal component analysis method, to transfer a plurality of indicators into less number, we use it to quantify comprehensively in this case.

Principal component analysis is carried out on the six evaluation indexes, the characteristic root and its contribution rate of the correlation coefficient matrix are in Table 2.

Table 2. The Characteristic Root and its Contribution

No	Eigenvalue	Contribution rate (%)	Cumulative contribution rate (%)
1	3.1172	51.9539	51.9539
2	1.6649	27.7485	79.7024
3	0.8169	13.6152	93.3176
4	0.3366	5.6099	98.9276
5	0.0542	0.9028	99.8303
6	0.0102	0.1697	100.0000

According to the data in Table 2, the cumulative contribution rate of the first three characteristic value is 93%, more than 90%, which shows that the effect of the principal component analysis is very good. So the first three principal components are selected for scientific research and comprehensive evaluation below.

The coefficient in front of the standardized variable is the corresponding eigenvector of standardization sample principal component, and then principal component function of three standardization sample can be obtained.

$$y_1 = 0.4260x_1^* + 0.3881x_2^* + 0.3919x_3^* - 0.1841x_4^* + 0.5115x_5^* + 0.4678x_6^*$$

$$y_2 = -0.2383x_1^* + 0.4765x_2^* + 0.5260x_3^* + 0.6111x_4^* - 0.1689x_5^* - 0.1938x_6^*$$

$$y_3 = -0.4835x_1^* - 0.2505x_2^* - 0.1894x_3^* + 0.4544x_4^* + 0.3986x_5^* + 0.5498x_6^*$$

With three principal component contribution rate for weight respectively, we reconstruct the principal components Comprehensive Assessment Model:

$$Z = 0.5195y_1 + 0.2775y_2 + 0.1362y_3$$

Plugging each principal component value into the above formula for computing, comprehensive evaluation value and ranking results of each scholar are shown in Table 3.

Table 3. 30 scholars' comprehensive evaluation value and ranking results

Name	Suning You	Huo Ren	Canhua Chen	Fuman Gong	Hong Xiao
Z	2.8113	0.6054	0.3233	-0.1277	0.3239
Rank	1	7	6	2	5
Name	Yiwen Shao	Daming Zhu	Weiguo Xu	Fan Tao	Lijun Yang
Z	1.2235	1.7913	-0.4524	0.2713	-0.1310
Rank	3	9	4	14	10
Name	Qianghui Ran	Chengfu Wu	Yuan Yao	Xingyong Zhang	Zuoxin Zhou

Z	-0.3793	-0.3674	-0.5348	-0.1294	-0.8193
Rank	12	11	8	13	19
Name	Chucai Xiong	Yulin Cai	Wenshen Xu	Meixiang Zhu	Tiecheng Jin
Z	-1.0874	-0.9078	-0.8882	-0.7303	-0.7948
Rank	20	15	18	17	16

2.3.3. Results analysis

Through Pearson correlation analysis between Z and other factors, the correlation coefficient Matrix can be acquired (see Table 4).

Table 4. The Correlation Coefficient between Z and other factor

Issued amount	h-index	Total citations	The average citation rate	Download frequency	Browse frequency
0.5302	0.8271	0.8632	0.0350	0.8008	0.7348

According to the Table 4, the correlation coefficient between Z and above factors is not very large, it have not reach 90%. Observing the results in Table 1, the correlation coefficient between h-index and issued amount is 0.3220, while the correlation coefficient between h-index and total citation is 0.8920. What's more, there is a relatively strong correlation between Z and h-index while the correlation coefficient is 0.8271 in Table 4. However, the correlation coefficient between Z and issued amount is 0.5302, while the correlation coefficient between Z and total citation is 0.8632. The comparison shows that the comprehensive evaluation value Z to a certain extent weakened the contribution of citation, so as to emphasize the contribution of other factors. However, considering that there is a little complexity and complicated operation to apply this method, we will introduce a new index in the next section.

3. A new index to measure: h_f -index

3.1. "Gold priority" rule

Olympic gold medal table originates in the 1924 Olympic Games in Paris, mainly to sort and compare the level of national and regional sports. This ranking method is popular all over the world. It is the highest level of medal - gold medal in the number as the key results sorted. At the time of sorting, the first comparison is the number of gold medals, an equal number of gold medals leads to the comparison of the silver number. If the numbers of silver are the same as well, then the number of bronze can be compared, this is in turn. In other words, try to use the highest level of sports achievement to reflect the standard of sports in this country. This ranking rule is named "Gold priority" rule.

"Gold priority" rule is used to measure various national and regional contribution to the Olympic sports. Although this ranking method is not formal, but the delegations and audiences all over the world give the gold medal table a

high degree of attention, gold medal table and therefore becomes the focus of the audience during the competition. Ranking as this method, during the 2004 Athens Olympic Games, Chinese delegation got 32 gold medals, the medal total to 63, while Russia got 27 gold medals, the medal total to 92, from the perspective on the total number of medals, China lags behind that of Russia, but as a result of China's gold medals ahead of Russia, China eventually was second in the medal table. Remember the 2008 Beijing Olympic Games, the Chinese delegation got 51 gold medals, a total of 100 medals, although 10 medals behind the United States in the total number. But by the leading number of gold medals, China won the gold medal table champion.

3.2. The rationality of "Gold priority" rule applied to the evaluation

"Gold priority" rule can be applied to improve the current evaluation system which owns quite a lot of defects. First, it is in line with the meaning of evaluation which can contribute to the academic level of the papers which published by the scientific researcher. Second, it determines the content of evaluation. The key is the true level of scholars in research contribution. And other factors can only in the same situation consider as the secondary factors, such as "silver" and "bronze" factor. Third, it can truly and objectively show the overall strength of the best athletes from each delegation. Compared with those evaluation approaches that excessively pursuit "objective quantitative evaluation", this approach is more fair and reasonable.

3.3. The application of "Gold priority" rule to the evaluation

Wei Bu, Chinese academy of the social sciences, and others have presided over a research project, mainly about the value evaluation of social science achievement, evaluation methods that research finally established, showing the characteristics of the "Gold priority" rule. In 2006, inspired by the Olympic gold medal table ranking rule, Qiu et al. mentioned in "The thoughts for "gold priority" rule applies to research talent evaluation" (2006) that the academic level of researchers should depend on the researchers' optimal results. This method can guide the scientific researchers pay more attention to the quality of academic papers. In 2009, Xu proposed a new modified index: h_a -index, which was based on the "Gold priority" rule on the initiative. h_a -index, on the basis of the "gold priority" rule, expanded the scope of application of the h-index, was more sensitive than h-index, with higher precision, so that the scholars with same h-index could be a deeper comparison.

In this section, as a way of implementation for the "Gold priority" rule, a new method is proposed based on the idea from Qiu et al. From another perspective, it can also be regard as a correction for the h_a -index.

3.4. h_f -index

If quantifying an individual's scientific research output in accordance with the "Gold priority" rule directly, the value of each factor must be separately

calculated, and then one by one to compare. This requires multiple times, and the process is relatively complicated. In view of this situation, this paper puts forward h_f-index, which retain all properties and advantages of the "Gold priority" rule and simplified it into a single value, then scientific research rank can be based on the size of the h_f-index.

3.4.1. Min-Max Normalization

1) Min-Max Normalization: $X^* = \frac{X - m}{M - m}$.

Wherein, X represents the original data, m represents the maximum value among the original data, M represents the minimum value among the original data.

2) If not specified, X^* reserves two decimal places in this article.

3.4.2. The meaning of h_f-index

The definition of h_f-index: according to "Gold priority" rule, regard h-index as the gold medal, the score of average citation rate as the silver medal, the score of average download frequency as bronze medal, a new index is concluded.

The computational formula of h_f-index is:

$$h_f = H + \frac{1}{100} AR + \frac{1}{100^2} AD$$

Wherein, h represents the h-index, round represents the data rounded, Ar represents an average citation rate, Ad represents the thesis average download frequency.

Capital letter H represents the h-index with Min-Max Normalization. Similarly, Capital letter AR, AD represents the Ar, Ad with Min-Max Normalization respectively.

Comparative method: Unlike h_a-index, h_f-index can directly compare size, the greater the h_f-index indicates a larger influence of the scholar, and then the greater the order will be.

3.4.3. The applicability of h_f-index

h_f-index presented in Section 3.4.1 also has certain limitation. Since the index retains six decimal places, h_f-index is likely to be equal in the high number of scholars being evaluated. In this case, we can expand the number of decimal places from six to eight or more decimal digits. Expanding method is: In view of the single factor except "gold medal", changing percentile scoring system with a thousand-point scale even higher median score, the higher the digits, the more accurate ranking the single factor. Computation formula is as follows.

The computation formula with 6-bit decimal is as follows.

$$h_f = H + \frac{1}{100} AR + \frac{1}{100^2} AD$$

The computation formula with $3N$ -bit decimal is as follows:

$$h_f = h + \frac{1}{10^N} AR + \frac{1}{10^{2N}} AD$$

3.4.4. Promotion of the h_f -index

Section 3.4.1 and Section 3.4.2 mainly tell the scientific evaluation method with three factors, this method is to promote, and the formula can take more factors into consideration after promoting. It is important to point out the importance and the sequence of each factor in the process of promoting.

The scientific evaluation method with two factors is as follows.

$$h_{2f} = H + \frac{1}{100} AR(h)$$

Wherein, h represents the h-index, Ar(h) represents the average citation rate of "h-core" papers, $AR(h)$ indicates the Ar(h) with Min-Max Normalization among persons evaluated.

In the process of comparison to h_a -index, the first comparison with descending order is the integer part of h_a -index, which is the h-index. And then compare the decimal part of h_a -index with ascending order. This ranking method is easy to cause confusion, whereas h_{2f} -index can be directly compared according to the size of the values, so this method can also be seen as a correction of h_{2f} -index.

The scientific evaluation method with four factors is as follows.

$$h_{4f} = H + \frac{1}{100} C + \frac{1}{100^2} N + \frac{1}{100^3} D$$

Wherein, h represents the h-index, c represents the citation, n expresses the total number of papers, and d represents download frequency. Of course, four factors selected here is not fixed, they can be changed according to the actual situation.

The scientific evaluation method with N factors is as follows.

$$h_{Nf} = X(1) + \frac{1}{100} X(2) + \cdots + \frac{1}{100^{i-1}} X(i) + \cdots + \frac{1}{100^{N-1}} X(N)$$

Wherein, $X(i)$ represents the i-th factor in the process of scientific research evaluation, which is ranked i-th according to the importance of factors.

3.5. Practice of the h_f -index

3.5.1. Data source

We use the same data as in Section 2.3.1, which can be got in the Appendix.

3.5.2. Data analysis

We work out the h_f-index of 20 scholars according to the data in the Appendix. In order to show h-index more intuitive and accurate, comprehensive evaluation value Z, h_f-index and rankings, the results are as follows (see Table 5).

Table 5. The h-index. Z. h_f-index and ranking results of 20 scholars

Name	h-index	Z	hf-index	Rank for hf
Suning You	11	2.8113	1.006515	1
Huo Ren	6	0.6054	0.380900	4
Canhua Chen	6	0.3233	0.384914	2
Fuman Gong	6	-0.1277	0.381406	3
Hong Xiao	5	0.3239	0.260010	5
Yiwen Shao	5	1.2235	0.250000	9
Daming Zhu	5	1.7913	0.251943	7
Weiguo Xu	5	-0.4524	0.252300	6
Fan Tao	5	0.2713	0.250514	8
Lijun Yang	4	-0.1310	0.131508	14
Qianghui Ran	4	-0.3793	0.136436	11
Chengfu Wu	4	-0.3674	0.131114	15
Yuan Yao	4	-0.5348	0.130401	16
Xingyong Zhang	4	-0.1294	0.138900	10
Zuoxin Zhou	4	-0.8193	0.133610	12
Chucai Xiong	4	-1.0874	0.131901	13
Yulin Cai	3	-0.9078	0.005504	17
Wenshen Xu	3	-0.8882	0.005209	18
Meixiang Zhu	3	-0.7303	0.004835	19
Tiecheng Jin	3	-0.7948	0.001718	20

Observe the h-index column in Table 5, in 20 scholars, in addition to Suning You who located in the first place and with a h-index of 11, much higher than other scholars, the rest of the 19 scholars, each h-index of the scholar is not unique, and even seven scholars have h-index of 4 (more than a third of people have equal h-index), in this case, it is obviously unreasonable to use the h-index to evaluate, so an index with higher sensitivity and differentiation is very necessary. And then observed the h_f-index column in Table 5, by definition,

the integer part of h_f -index maintaining the h-index, so the decimal part has only carried on the further comparison, the effect of this approach is obvious, each h_f -index is unique, ranking results are therefore rather unique, won't appear name repetition.

3.5.3. The relation between h-index, comprehensive evaluation value Z and h_f -index

Respectively calculate the Pearson correlation coefficient between h-index, comprehensive evaluation value Z and h_f -index, and the results as follows: The correlation coefficient between h-index and Z is 0.8271, the correlation coefficient between Z and h_f -index is 0.8240, and the correlation between the h-index and h_f -index is up to 0.9999. Therefore, there is a relatively significant correlation between the h-index and Z, while h-index and h_f -index is particularly significant.

3.5.4. Difference analysis between comprehensive evaluation value Z and h_f -index

The correlation coefficient between Z and the h_f -index reached 0.8240, but compare with the correlation coefficient between h-index and h_f -index of 0.9999, there is still a big difference, and the reason is as follows.

There is a different thought between comprehensive evaluation values Z and h_f - index. h_f -index is an amendment to the h-index, enables the scholar with same h-index to more in-depth comparison, and improves the degree of differentiation for the h-index. In addition, because the h-index, the thesis average citation and thesis average download frequency are considered in the h_f -index, this will to a certain extent enhance the h-index linked to other factors. As for the comprehensive evaluation value Z, proposed to avoid the oneness on numerical factors, it comprehensively considers various factors, is more complicated in the calculation process, and there is a big difference with the h_f -index.

4. Conclusions

h-index, from the date of birth, has attracted the attention of the international scientific community. Many scholars study the discussion of the h-index with great interest, hoping to get some results. The current paper mainly concentrated on two aspects: first, to explore scientificity and practicability of the h-index from all aspects, so as to expand the applicable scope of the h-index; Second, to modify the h-index and class h-index and then put forward some new derivative index, thereby improving the evaluation system.

In the process of discussion and research, people need to recognize that the h-index is not omnipotent, it still exists many defects. When we apply h-index and class h-index to quantify an individual's scientific research output, the advantages and disadvantages as well as prerequisites of these indexes are worth noting. When evaluating scholars, comprehensive evaluation method with multifaceted, multi-factor and multi-index should be taken into consideration, so as to make evaluation more efficient, fair and reasonable.

To make comprehensive assessment on scientific research achievements of scholars more multifaceted, multi-factor, multi-indexes, we has used two methods to quantify an individual's research achievements in a comprehensive analysis. The first method is to analyze a variety of factors in the original data, taking the various factor's contribution to the scientific research evaluation into account. The second method is to propose a new comprehensive research evaluation h_f-index, which can avoid the complex calculations when evaluating with some mathematical thoughts, and enhance the degree of differentiation and sensitivity of h-index, easy to use.

h-index is an indicator in the field of literature metrology, and many achievements in theoretical and empirical research have been launched. But it is still not perfect in many places and need a further research and argumentation.

References

Hirsch, J. E., (2005). An index to quantify an individual's scientific research output. Proceedings of the National Academy of Sciences of the USA, 102 (46): 16569-16572.

Braun, T, Glaznzel, W, Schubert, A, (2006). A Hirsch-type index for journals. Scientometrics,69 (1): 169-173.

Egghe, L, (2006). Theory and practise of the g-index. Scientometrics, 69 (1): 131–152.

Marek, K, (2006). A new Hirsch-type index saves time and works equally well as the original h-index. ISSI Newsletter, 2 (3): 4-6.

Jin, B.H., Liang, L.M., (2007), Ronald Rousseau and Leo Egghe. The R- and AR-indices: Complementing the h-index. Chinese Science Bulletin, 52 (6): 855-863.

Zhang, X.M., (2007). h_m-index——Correction of h index. Library and information service, 51 (10): 116-118 (In Chinese).

Qiang, Wu, (2008). The w-index: A significant improvement of the h-index. Access date 10.04.2016 available at http://arxiv.org/abs/0805.4650.

Anderson, T.R., Hankin, R.K.S., Killworth, P.D., (2008). Beyond the Durfee square: Enhancing the h-index to score total publication output. Scientometrics, 76 (3): 577–588.

Ye, Y, (2009). A new index of academic sort —— f-index analysis. Information learned journal, 28 (1):142-149 (In Chinese).

Xu, X.J. (2009). h_a-index：Correction of h-index —— the revelation of the "gold priority" rule. Journal of information theory and practice, 11 (32): 8-12 (In Chinese).

Alonso, S, Cabrerizo, F. J., Herrera-Viedma, E, Herrera, F, (2010). hg-index: a new index to characterize the scientific output of researchers based on the h- and g-indices. Scientometrics, 82:391–400.

Pearson, K, (1901). On lines and planes of closest fit to systems of points in space. Philosophical Magazine, 1901, 2 (6):559-572.

Hotelling, H, (1933). Analysis of a complex of statistical variables into principal components. Journal of Educational Psychology, 24 (6): 417–441.
Wei, R.B., Song, G, (2009). h-index research review and the empirical statistical analysis. Chinese Journal of Scientific and Technical Periodicals, 20 (2): 220-224 (In Chinese).
Qiu, J.P., Zhu, S.Q., Liu, Y, (2006). The thoughts for "gold priority" rule applies to research talent evaluation. Knowledge of library and information science, 4: 101-104 (In Chinese).

Appendix

20 well-known scholars' papers cited situation

	Issued amount	h-index	Total citations	The average citation rate	Download frequency	Average download frequency	Browse frequency
Suning You	36	11	258	7.17	729	20.2500	699
Huo Ren	66	6	122	1.85	540	8.1818	425
Canhua Chen	20	6	113	5.65	375	18.7500	270
Fuman Gong	27	6	63	2.33	331	12.2593	322
Hong Xiao	13	5	137	10.54	211	16.2308	195
Yiwen Shao	117	5	121	1.03	952	8.1368	841
Daming Zhu	32	5	91	2.84	1345	42.0313	1700
Weiguo Xu	23	5	75	3.26	180	7.8261	139
Fan Tao	37	5	56	1.51	704	19.0270	729
Lijun Yang	32	4	80	2.5	444	13.8750	467
Qianghui Ran	8	4	57	7.13	295	36.8750	298
Chengfu Wu	25	4	53	2.12	484	19.3600	382
Yuan Yao	36	4	52	1.44	309	8.5833	364
Xingyong Zhang	5	4	47	9.4	441	88.2000	414

Zuoxin Zhou	10	4	45	4.5	157	15.7000	127
Chucai Xiong	10	4	28	2.8	88	8.8000	101
Yulin Cai	9	3	56	6.22	101	11.2222	98
Wensh en Xu	9	3	54	6	132	14.6667	120
Meixia ng Zhu	8	3	45	5.63	287	35.8750	260
Tieche ng Jin	15	3	39	2.6	336	22.4000	337