Judging Paper more objectively and effectively by Identification Algorithm

Summary

An effective and objective tool for evaluating paper is the amount of citation and the cited papers' prestige. Citations symbolize the conceptual association of scientific ideas as recognized by publishing research authors. In a nutshell, the cited frequencies and the significant degree of citations have being the fundamental indicators.

We create **Identification Algorithm** searching the citation of papers in a large system. By this algorithm, we can not only count the amount of citation, but also identify the path of cited. Under the path of cited and the magnitude, we successive modify the citation evaluate, array the prestige of each paper. In addition, exemplify the simple system upon the problem. Those six arrayed papers, B, F, A, E, D, C is showed as follows. Their score are 0.4946, 0.2070, 0.1148, 0.0919, 0.0459, 0.0458.

In actual model, we discuss the journal impact factor as a quantitative tool for ranking, evaluating, and comparing journals. By analyzing the impact factor calculation from SCI, we consider some factors will affect the impact factor, such as self-citations, title change, and Evaluation cycle. We use the factor as a weight coefficient for each journal. Across the different weight coefficient, we develop the actual model to array the paper objectively. In the conclusions section, we simulate our actual model by computer, arraying the 10 journals and 100 papers. The result shows more objectively and accurate.

In our further discussion, we consider the "web metrics" as a further to judge the papers. In additional, our model using actual citation data, evaluate a gross approximation of the prestige of papers and journal. It shows more objectively and actual.

Introduction

An appropriate topic to launch this new series is perhaps the most rudimentary — the basic concept of citation indexing. Citations symbolize the conceptual association of scientific ideas as recognized by publishing research authors. A case of indexes taking into account, Thomson Scientific's indexes are comprehensive, providing complete coverage of all types of published source items--not just original research papers, review articles, and technical notes but also letters, corrections and retractions, editorials, and other items. Thomson Scientific studies have shown that these latter items are important, have substantial impact, and provide useful links to scientific issues and controversies. [1] Facing with the historical development of citation indexing, it is not difficult to know that the demand of managing information, the condition of growing dissatisfaction and the hope of automotive mechanisms are the three main phase affecting citation indexing. Later on, the viability and efficiency of citation indexing projects is tested by Eugene Garfield and Associates in the early 1960s. Gradually, papers and articles are published along with scientific and economical development and needed to be evaluated with some actual indicators. In a nutshell, the cited frequencies and the significant degree of citations have being the fundamental indicators.

At the time of the project's completion, the government sponsors chose not to subsidize the development of a national citation database, Eugene Garfield was encouraged to move ahead with the private publication of his multidisciplinary citation index as the first edition of the Science Citation Index® (SCI®). [2]That is to say that the advent of the Thomson Scientific citation indexes made it possible to do computer-compiled statistical reports not only on the output of journals but also in terms of citation frequency. And in the '60s we invented the journal "impact factor." After using journal statistical data in-house to compile the Science Citation Index® (SCI®) for many years, Thomson Scientific began to publish Journal Citation Reports® (JCR®)3 in 1975 as part of the SCI and the Social Sciences Citation Index® (SSCI®). [3] As for the evaluation of papers and articles from SCI, the effect on appraising papers is made by total cites, impact factor, immediacy index, article counts and cited half-life and so on. Today, this citation indexes are considered to be one of the most reliable of resources in tracing the development of an idea across the multitude of disciplines that are part of our body of scientific knowledge. [2]

First of all, the simplest complexion is related to our consideration. Based on the given relationship among different papers in the title, these journals are assumed to be self-citation within same journals that are considered as contemporaneous publications. Furthermore, we develop our algorithm and arrange the significance of articles. In the next place, we extend the application of front papers into more papers. And the number of papers is N . These papers belong to journals, the number of which is considered to be M ($M \leq N$). Integrating the time with different impact factors, algorithm can be gained and papers can be disposed in terms of the weightiness degree of N papers.

Basic Assumptions

We make the following assumptions about the preventing Avian Flu pandemic in this paper.

- **Each paper and journal belongs to the same domain.** In a perfect system papers and journals ought to be possible to compare journals with an identical profile. But in fact there rarely are two journals with identical semantic or bibliographic profiles. We believe, however, categorize can make the model more actual and objective.
- **Self-citation can't be taken into account.** Self-citation rate shows only a weak correlation with the impact and subject of a journal and paper. In addition, self-citation has the potential to distort the true role of the title as a participant in the literature of its subject.

Additional assumptions are made to simplify analysis for individual sections. These assumptions will be discussed at the appropriate locations.

Key Terminology

- **Total Cites.** Total cites indicate the total number of times that each journal has been cited by all journals included in the database within the current year.
- **Impact Factor.** The Impact Factor identifies the frequency with which an average article from a journal is cited in a particular year. You can use this number to evaluate or compare a journal's relative importance to others in the same field or see how frequently articles are cited to determine which journals may be better for your collection.
- **Immediacy Index.** The Immediacy Index measures how frequently the average article from a journal is cited within the same year as publication. This number is useful for evaluating journals that publish cutting-edge research.
- **Article Counts.** The number of articles published in a journal in a particular year or years.
- **Cited Half-life.** The cited half-life benchmarks the age of cited articles by showing the number of years back from the current year that account for 50% of the total number of citations to a journal in the current year. This number is useful in making collection management and archiving decisions. A publisher may use this number to adjust editorial policies to compete in different market segments.
- **"Reference" and "citation".** if Paper R contains a bibliographic footnote using and describing Paper C, then R contains a reference to C, and C has a *citation* from R.

Simple Model

Formal Definition

Let A be the citation matrix of article and a_{ij} mean paper i is cited by paper j within involved matrix. To illustrate better the above statement, figure 1 is presented as follows.

Figure 1

The situation that paper i is cited by paper j can be viewed in figure **. We regard the citation is existent as $a_{ij}=1$. Otherwise, the relation is not of existence, which is considered to be $a_{ij}=0$.

Therefore, we establish the matrix A as well as a_{ij} ($a_{ij}=1$ or $a_{ij}=0$). An example of part situation of papers citations is taken by given title. In succession, as the figure 2 shows.

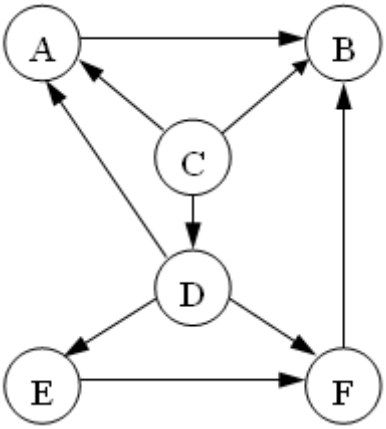

Figure2

According to the direction of front figures, the cited papers matrix A_1 is developed.

$$A_1=\begin{pmatrix}0&1&0&0&0&0\\0&0&0&0&0&0\\1&1&0&1&0&0\\1&0&0&0&1&1\\0&0&0&0&0&1\\0&1&0&0&0&0\end{pmatrix}$$

Having constructed the cited papers matrix, we are now faced with the following two tasks:

- The citations of respective articles are sought in allusion to every article.
- On the basis of different citations of articles, the importance of articles is arranged.

Searching the citation

This issue is treated of the cited nexus of articles. In relation to the complicated connection citations among these articles, we propose cited matrix and identifier matrix to express the cited relationship of articles.

Above all, based on the built cited matrix A_1:

$$A_1 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

The sum of line j in matrix A_1 expresses the number of citing the article j, and the sum of row i expresses the number of articles that paper i cites.
Therefore we can gain the corresponding matrix A_1^n describe the number of cited indirectly relation is $n-1$(cited directly connection is thought as the number of cited indirectly is 0). Such as A_1^2:

$$A_1^2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 2 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

The formula $a_{31} = 1$ of matrix A_1^2 figures that paper C cites indirectly paper A through one access and one time. And the formula $a_{42} = 1$ figures that paper D cites indirectly paper B through two different accesses and one time. These illustrate is showed as following figures3:

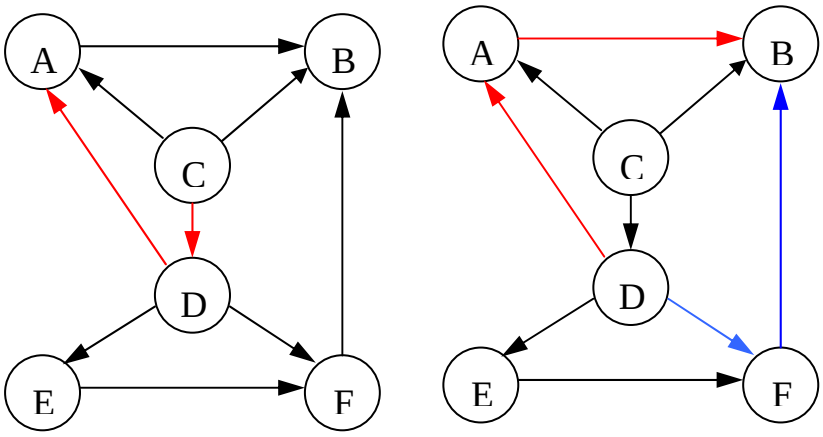

Figure3. The access of citation

In spite of the matrix A_1^n is capable of expressing whether there is $(n-1)$ time of indirect citation connection or not, the access is unable to be apparent. Hence, our paper establishes the corresponding identifier matrix and treats the signals of articles as identifiers.

$$A_2 = \begin{pmatrix} 0 & ab & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ ca & cb & 0 & cd & 0 & 0 \\ da & 0 & 0 & 0 & de & df \\ 0 & 0 & 0 & 0 & 0 & ef \\ 0 & fb & 0 & 0 & 0 & 0 \end{pmatrix}$$

In the above identifier matrix, the formula $a_{12} = ab$ of matrix A_2 shows that paper a is cited by paper b. The identifier matrix can be developed without much time after the signal calculation of MATLAB can be an assistant of computing the outcome. The possible access can be expressed with the corresponding identifier in this identifier matrix. When the corresponding matrix A_1^n is appeared, the matrixes that the number of indirect citation connection and access is $n-1$ are described evidently and legibly. And then work out A_2^2:

$$A_2^2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ cd*da & ca*ab & 0 & 0 & cd*de & cd*df \\ 0 & da*ab+df*fb & 0 & 0 & 0 & de*ef \\ 0 & ef*fb & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Therefore, the description of indirect citation access can be showed. A case of $a_{31} = cd*da$ is taken so as to express that paper c cites directly paper d and paper d also cites directly paper a. That is to say that paper c cites indirectly paper a via paper d. By this token, the formula $da*ab+df*fb$ shows that the former and the latter formula can deliver two different accesses citations.

After computing such a calculation, it is not difficult to attain a cited and related figure of every single article. Paper b is as an instance:

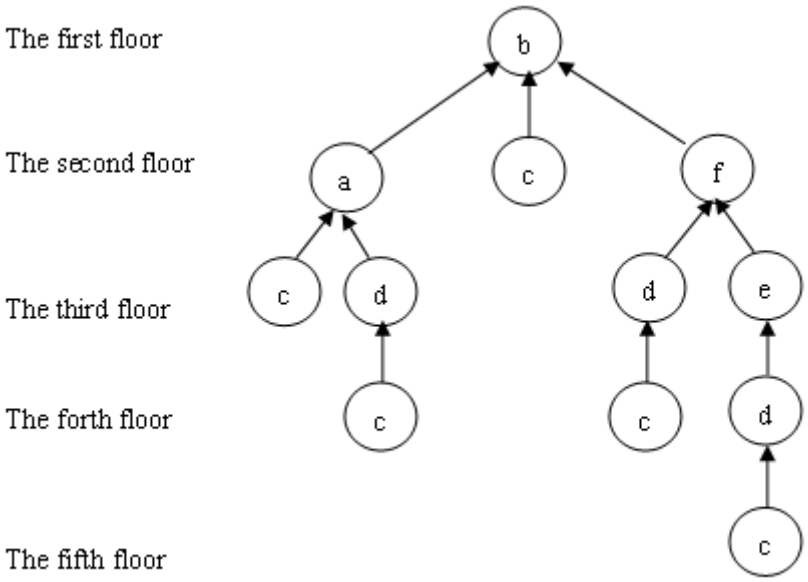

Figure4

Ranking

We obtain the citation relationship by the usage of identifier searching measures in the segment of above papers. So the next destination is the arrangement of importance degree of different papers.

An example of paper b is taken as the above figure. The influence of the significance

degree of paper b is the three paper a、c、f citing paper b and the importance degree of these papers . Evidently, the importance degree of three papers is affected by the paper in the third floor. The rest may be deduced by analogy, the upstairs have an effect on the downstairs through every floor gradually and the bottom floor affects weakly to the up floor.

According with the above figure, the importance degree of every paper is endowed with the initial value $1/6$ in the beginning. According to the above citation figure, then the importance degree of every paper is revised gradually through every floor. Consequently, the importance degree of paper b is gained after revision.

Since the up floor has a weak connection with the lower floor, revised strategies are operated as is following (the importance degree of literature =IDL):

$$floor\ i\ node\ j\ of\ IDL = original\ IDL + \frac{floor\ (i+1)\ sum\ of\ IDL\ citation}{the\ number\ of\ node\ in\ floor\ i}$$

Based on the above-mentioned analysis, we first figure out the revised value of importance degree of every paper. In the next instance, the revised value is not normalized until the revised value hardly has discrepancy compared with the former revised value after operating in the above system. Finally, our revision will be ceased and we gain the final importance degree of every paper. Namely, according to the final importance degree of every paper, the arrangement is attained.

Simulation

The simulation of this system with available the tool MATLAB is obtained that the importance degree of six papers a、b、c、d、e、f is 0.1148 、0.4946、0.0458 、0.0459 、0.0919 、 0.2070, and the arrangement is showed as table 1.

Table1. The outcome of arrangement

rank	Importance degree	paper	Cited situation	Citing situation
1	0.4946	b	a c f	-
2	0.2070	f	d e	b
3	0.1148	a	c d	b
4	0.0919	e	f	d
5	0.0459	d	c	a e f
6	0.0458	c	-	a b d

The importance degree of paper is in accordance with the number of cited papers, which is showed on the above table. It is unlabored to learn about the more number of papers is cited, the more important degree of papers is certain. The table of paper f and paper a are cited by two papers (cited by paper d), and also citing paper b. However, the paper e cited

by the paper f is as much more significant as the paper d cited by the paper a. Consequently, the importance degree of paper f seems to be outstanding compared with the paper a. It is proved that paper cited by important paper behave more effectively to add the importance degree of articles than the one cited by common papers.

Actual model

Considering these articles pertain to different journals under practical situation, the capability of influence is dissimilar according to diverse journals. In this instance, the cited frequencies and the significant degree of citations into consideration are taken is insufficient. The solution of SCI to the problem is the introduction of impact factor. We work out a precise impact factor via settling the citied statement of various journals. So it is certain to gain the prestige of diverse journals.

Considering these analysis, we are confronted with the following challenge:

- Making sure the impact factor of disparate journals;
- Introducing the impact factor into a model to obtain the significant arrangement of articles for sake of according with practice.

Formal Definition

Above all, we classify these N pieces of articles in the light of journals. An assumption of the existence of M pieces of journals is known, and every journal i consists of n_i pieces of paper. It is claimed that every journal is treated as a rooted tree. As showed in the down figure:

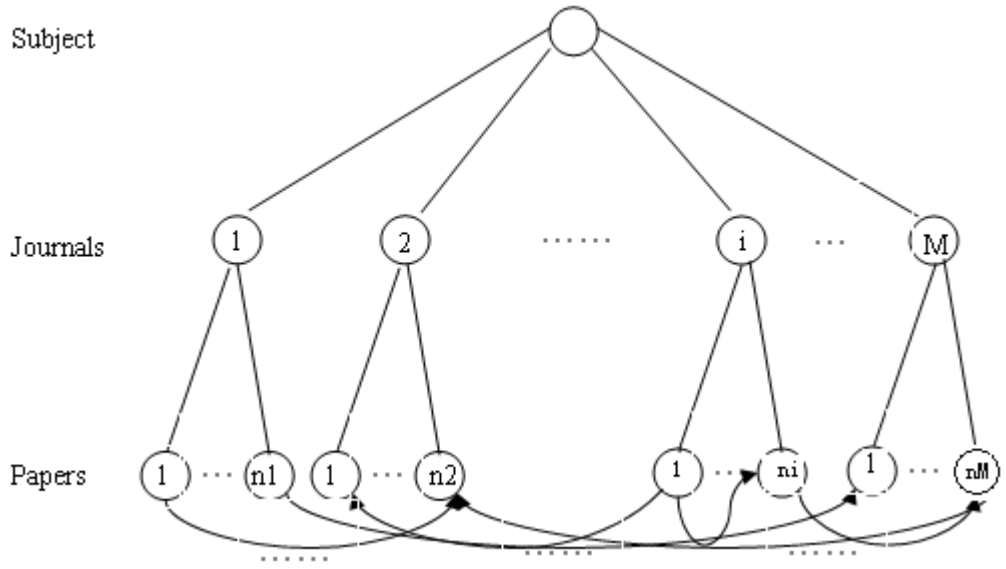

Figure 5 a rooted tree

In the previous figure, we need to analyze the mutual effect connections of respective leaf in the rooted tree through summarizing. Later on, figure out the impact factor of every rooted tree M_j. Furthermore, we endow these impact factors as power values with every

leaf in the rooted tree. Sequentially, we modify the simple model, and gain a better model of matching with reality to arrange the importance degree of papers.

Impact Factor

Conventional Calculation

Journal Citation Reports (JCR) is considered as part of the SCI and the Social Sciences Citation Index (SSCI).[4] The JCR provides quantitative tools for ranking, evaluating, categorizing, and comparing journals[5]. And the impact factor which is one of these can be used to measure the frequency with which the "average article" in a journal has been cited in a particular year or period. The annual *JCR* impact factor is a ratio between citations and recent citable items published. Thus, the impact factor of a journal is calculated by dividing the number of current year citations to the source items published in that journal during the previous two years (see table 2). So, for example, the 1999 impact factor is the citations in 1999 to articles published in 1997 and 1998 divided by the number articles published in 1997 and 1998.

Table 2. Calculation for journal impact factor

letter	meaning
A	total cites in 1992
B	1992 cites to articles published in 1990-91 (this is a subset of A)
C	number of articles published in 1990-91
D	B/C = 1992 impact factor

There are three existing aspects within which it is uncertain to calculate the impact factor effecting journal impact factor among the above strategies, and these aspects are:

- Scientific the cycle of calculation;
- The self-citation of articles;
- The publication of new or old articles

Using the Impact Factor Wisely

Different specialties exhibit different ranges of peak impact. That is why the JCR® provides subject category listings. In this way, journals may be viewed in the context of their specific field. Still, a five-year impact may be more useful to some users and can be calculated by combining the statistical data available from consecutive years of the JCR (see table 3). It is rare to find that the ranking of a journal will change significantly within its designated category unless the journal's influence has indeed changed.

The cited half-life benchmarks the age of cited articles by showing the number of years back from the current year that account for 50% of the total number of citations to a journal in the current year. This number is useful in making collection management and archiving decisions. A publisher may use this number to adjust editorial policies to compete in different market segments.

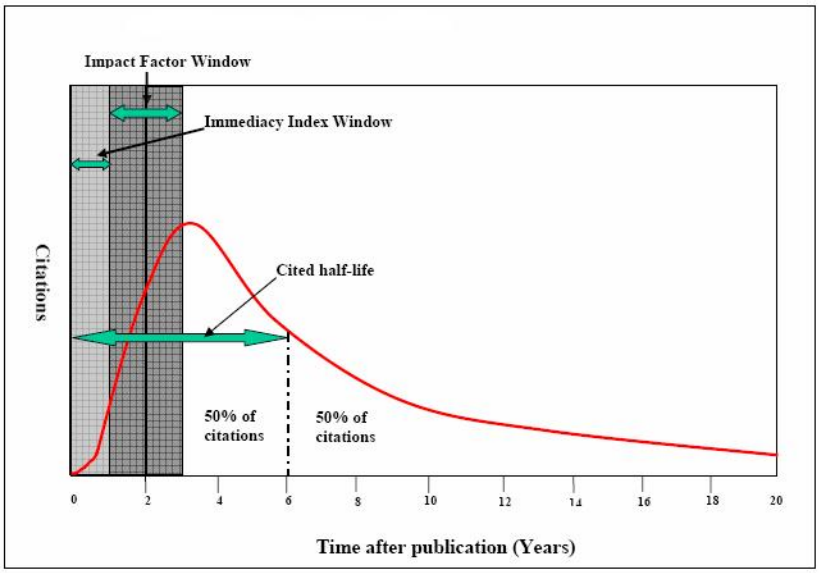

Figure6. Cited half-life

source: http://www.elsevier.com/framework_editors/pdfs/Perspectives1.pdf[5]

It is the number of years that the number of current citations takes to decline to 50% of its initial value (the cited half-life is 6 years in the example given in Figure3). It is a measure of how long articles in a journal continue to be cited after publication. [5]

Table3 Calculation for five-year impact factor:
One year of citations to five years of articles

letter	meaning
A	citations in 1992 to articles published in 1986-91
B	articles published in 1986-91
C	A/B = five-year impact factor

An alternative five-year impact can be calculated based on adding citations in 1987-92 articles published in the same five-year period. And yet another is possible by selecting one or two earlier years as factor "B" above.

The self-citation of articles

An examination of the Cited Journal list for most titles will contain data counting citations from the journal to itself, i.e., instances in which an article published in a journal has cited a previously published article in that same journal. These references are often called "self-citations." We found that self-citation rate shows only a weak correlation with the impact and subject of a journal. There is also a weak correlation between self-citation rate and the size or specificity of the category (categories) assigned to a journal. The removal of self-citations from Impact Factor calculation had little effect on the relative rank of high impact journals. Some journals with lower Impact Factors and rank in category did show more dependence on the contribution of self-citations, but only a small proportion of journals show significant changes in quartile rank following the removal of

self-citations. [4]

Some of the journals listed in the JCR are not citing journals, but are cited-only journals. This is significant when comparing journals by impact factor because the self-citations from a cited-only journal are not included in its impact factor calculation. Self-citations often represent about 13% of the citations that a journal receives. The cited-only journals with impact factors in the JCR Journal Rankings and Subject Category Listing may be ceased or suspended journals, superseded titles, or journals that are covered in the science editions of Current Contents®, but not a citation index.

Users can identify cited-only journals by checking the JCR Citing Journal Listing. Furthermore, users can establish analogous impact factors, (excluding self-citations), for the journals they are evaluating using the data given in the Citing Journal Listing (see table 4).

Table4 .Calculation for impact factor revised to exclude self-citations

letter	meaning
A	citations in 1992 to articles published in 1986-91
B	1992 self-citations to articles published in 1986-91
C	A - B = total citations minus self-citations to recent articles
D	number of articles published 1986-91
E	revised impact factor (C/D)

The publication of new or old articles

A user's knowledge of the content and history of the journal studied is very important for appropriate interpretation of impact factors. Situations such as those mentioned above and others such as title change are very important, and often misunderstood, considerations.

A title change affects the impact factor for two years after the change is made. The old and new titles are not unified unless the titles are in the same position alphabetically. In the first year after the title change, the impact is not available for the new title unless the data for old and new can be unified. In the second year, the impact factor is split. The new title may rank lower than expected and the old title may rank higher than expected because only one year of source data is included in its calculation (see table 5). Title changes for the current year and the previous year are listed in the *JCR*® guide [6].

Table5. Unified 1992 impact factor calculation for title change

letter		meaning
A		1992 citations to articles published in 1986-91 (a1 + a2)
	A1	those for new title
	A2	those for superseded title
B		number of articles published in 1986-91 (B1 + B2)
	B1	those for new title
	B2	those for superseded title
C		unified impact factor (A/B)
	C1	A1/B1 = factor for the new title
	C2	A2/B2 = factor for the superseded title

On the basis of the front analysis of scientific the cycle of calculation, the self-citation of articles, the publications of new or old articles and so on, the formulation of figuring out impact factor is developed.

The impact factor of Journals in y year:

$$\lambda_y = \frac{A_y + \frac{A_{y-1}}{2} + \ldots + \frac{A_{y-5}}{6}}{B_y + \frac{B_{y-1}}{2} + \ldots + \frac{B_{y-5}}{6}} = \frac{\sum_{i=0}^{5} \frac{A_{y-i}}{i+1}}{\sum_{j=0}^{5} \frac{B_{y-j}}{j+1}}$$

Where

A_i is total citations minus self-citations to articles in a year i

B_i is the number of articles published in a year i.

Implementing Impact Factor objectivity

In the previous segment we come up with an idiographic calculation of the impact factor of different journals according to use for the reference of SCI. In succession, we add the arithmetic of impact factor to simple model so as to construct more practical model. Let the impact factor of every journal be $\lambda_i, i = 1,2,\cdots M$. And the power value of any paper among journal i is all considered λ_i. Firstly, the importance degree of every paper among journal i is endowed with initial value I_0.

$$I_0 = \frac{\lambda_i}{\sum_{i=1}^{M} \lambda_i n_i}$$

We define the coefficient of impact factor of journal i:

$$\alpha_i = \frac{\lambda_i}{\frac{1}{M}\sum_{i=1}^{M} \lambda_i}$$

As for belonging to paper in floor k and node l among journal i, the importance degree of paper is revised for t times (the importance degree of literature =IDL):

$$I_t' = I_{t-1} + \frac{floor(n+1)sum\,of\ IDL\,citations}{the\,number\,of\ node\,in\ floor\,m} \times \alpha_i$$

Then normalize the revision of different paper, and obtain the normalized revision I_t.

In the next instance, the revised value is not normalized until the revised value hardly has discrepancy compared with the former revised value after operating in the above system. Finally, our revision will be ceased and we gain the final importance degree of every

paper. Namely, according to the final importance degree of every paper, the arrangement is attained.

Simulation

To study this model, we run computer simulations and average the results over many simulations. In order to have data with which, we can regress the parameters in our model, we simulate a citation situation three journals and every journal is made of 15 pieces of papers.

In the first place, the impact factor and the coefficient of impact factor about different journals is calculated based on the published papers data and the number of citations (except self-citations) about three kind of journal within six years. The consequence is the following table:

Table6. The impact factor and the coefficient of impact factor

journal	impact factor	coefficient of impact factor
Ⅰ	8.2568	1.3594
Ⅱ	4.4688	0.7357
Ⅲ	5.4962	0.9049

Then we revised the results with the help of modify measures. Eventually, the final importance degree of papers and the ending of arrangement are gained.

Table7. The arrangement of outcome

rank	importance degree	journal	literature	cited statement	citing statement
1	0.17372	Ⅱ	h	b c I k m n o	0
2	0.15861	Ⅰ	c	b f k l	h
3	0.14731	Ⅱ	f	a d j m	c
4	0.10588	Ⅰ	d	e g k m	f
5	0.10563	Ⅰ	a	e	f
6	0.07804	Ⅰ	e	i l m	a d
7	0.06010	Ⅲ	l	b j n	c e
8	0.04461	Ⅰ	b	i j k m	c h l
9	0.03696	Ⅲ	n	g k o	h l
10	0.02796	Ⅱ	j	k o	b f l
11	0.01452	Ⅲ	k	i	b c d h j n
12	0.01202	Ⅱ	i	m	b e h k
13	0.01199	Ⅲ	o	---	h j n
14	0.01199	Ⅲ	m	---	b d e f h i
15	0.00975	Ⅱ	g	---	d n

Conclusions

We create **Identification Algorithm** searching the citation of papers in a large system. By this algorithm, we can not only count the amount of citation, but also identify the path of cited. Under the path of cited and the magnitude, we successive modify the citation evaluate, array the prestige of each paper. In addition, exemplify the simple system upon the problem. Those six arrayed papers, B, F, A, E, D, C is showed as follows. Their score are 0.4946, 0.2070, 0.1148, 0.0919, 0.0459, 0.0458.

In actual model, we discuss the journal impact factor as a quantitative tool for ranking, evaluating, and comparing journals. By analyzing the impact factor calculation from SCI, we consider some factors will affect the impact factor, such as self-citations, title change, and Evaluation cycle. We use the factor as a weight coefficient for each journal. Across the different weight coefficient, we develop the actual model to array the paper objectively. In the conclusions section, we simulate our actual model by computer, arraying the 10 journals and 100 papers. The result shows more objectively and accurate.

Further Discussion

In addition to helping scholars decide which papers to read and which journal to purchase. Journal impact factors are also used by authors to decide where to submit their articles. As a general rule, the journals with high impact factors include the most prestigious. The perception of prestige is a murky subject. Some would equate prestige with high impact. However, some scholars argue that the numerator in the impact-factor calculation is itself even more relevant. Bensman argued that this 2-year total citation count is a better guide to journal significance and cost-effectiveness than is the impact factor. [7] This brings us full circle to the first slide I showed on the most-cited journals.

Journal impact can also be useful in comparing expected and actual citation frequency. Thus, when *ISI* prepares a personal citation report it provides data on the expected citation impact not only for a particular journal but also for a particular year, because impact factors can change from year to year.

The use of journal impact factors instead of actual article citation counts to evaluate individuals is a highly controversial issue. Granting and other policy agencies often wish to bypass the work involved in obtaining actual citation counts for individual articles and authors. And allegedly recently published articles may not have had enough time to be cited, so it is tempting to use the journal impact factor as a surrogate evaluation tool. Presumably the mere acceptance of the paper for publication by a high impact journal is an implied indicator of prestige. Typically, when the author’s recent bibliography is examined, the impact factors of the journals involved are substituted in lieu of the actual citation count. Thus, the impact factor is used to estimate the expected influence of individual papers which is rather dubious

considering the known skewness observed for most journals.

Today so-called “webometrics” are increasingly brought into play, though there is little evidence that this is any better than traditional citation analysis.[7] Web “sitations” may occur a little earlier, but they are not the same as Citations. Thus, one must distinguish between readership or downloading and actual citation in new research papers. But some studies would indicate that web sitation is a harbinger of future citation.

The assumption that the impact of recent articles cannot be evaluated in *SCI* is not universally correct. While there may be several years delay on some topics, papers that achieve high impact are usually cited within months of publication and certainly within a year or so. This pattern of immediacy has enabled *ISI* to identify “hot papers” in its bimonthly publication *Science Watch*. However, full confirmation of high impact is generally obtained 2 years later. *The Scientist magazine* waits up to 2 years to select “hot papers” for commentary by authors. Most of these papers will eventually go on to become “citation classics.”

Evaluation of Solutions

Strengths

Our main model’s strength is its enormous flexibility. All of the data is fully parameterized, so the model can be applied to any situation, with different journals, different structures, and papers. Including all these factors into a single, robust framework, our model enables realistic simulation of citation but remains receptive to almost any modification.

In additional, our model using actual citation data, evaluate a gross approximation of the prestige of papers and journal. It shows more objectively and actual.

Weakness

Although we list the model's comprehensive, as a strength, it is (paradoxically) also the most notable weakness. Our model is a very useful tool for evaluation of papers, but it must be used discreetly. The impact factor should be used with informed peer review. This subjective factor is hard to obtain, and we didn’t consider in our objectively model.

References

[1] The concept of citation indexing: A unique and innovative tool for navigating the research literature. Retrieved 2, 2008 from http://scientific.thomson.com/free/essays/citationindexing/concept/

[2] History of citation indexing. Retrieved 2, 2008 from http://scientific.thomson.com/free/essays/citationindexing/history/

[3] The Thomson scientific impact factor. Retrieved 2, 2008 from http://scientific.thomson.com/free/essays/journalcitationreports/impactfactor/

[4] Journal self-citation in the journal citation reports: Science edition (2002) Retrieved 2, 2008 from *http://scientific.thomson.com/free/essays/journalcitationreports/selfcitation2002*

[5] Mayur Amin Elsevier，Boulevard, Langford Lane, Kidlington, derived October 2007 from the Institute for Scientific Information, Philadelphia, PA, USA. Retrieved 2, 2008 from http://www.elsevier.com/framework_editors/pdfs/Perspectives1.pdf

[6] Using the Thomson scientific impact factor. Retrieved 2, 2008 from *http://scientific.thomson.com/free/essays/journalcitationreports/usingimpactfactor/*

[7] Journal Citation Reports. ISI Web of Knowledge. Retrieved 2, 2008 from www.scientific.thomson.com/tutorials/jcr4

[8] The Agony and the Ecstasy—The History and Meaning of the Journal Impact Factor. Eugene Garfield. International Congress on Peer Review And Biomedical Publication Chicago, September 16, 2005. Retrieved 2, 2008 from www.eugenegarfield.org

Appendix 1: the process of the first problem

```
function [Imp, GI]=com_imp1
% A1 is the abut matrix
%    a  b  c  d  e  f
A1=[ 0  1  0  0  0  0  % a
     0  0  0  0  0  0  % b
     1  1  0  1  0  0  % c
     1  0  0  0  1  1  % d
     0  0  0  0  0  1  % e
     0  1  0  0  0  0];% f

% ==compute the tree matrix Am==
 A1=A1';
 e=1; % e is a flag to judge if AA==0 and dap the loop
 i=1;
 while e~=0
     AA(:,:,i)=A1^i;
```

```
      i=i+1;
      e=sum(sum(A1^i));
 end
BB(:,:,2:i-1)=AA(:,:,1:i-2);
BB=repmat(sum(BB,2),1,6);
BB(BB==0)=1;
AA=AA./BB;
Am=sum(AA,3);
Am=Am+eye(size(Am));


% ==compute the value of Imp==
Imp=1/6*ones(6,1); % initialize the value of the importance
C=Imp;
E=1; % initialize the value of the equation
while E>1e-8 % Dap the loop while the equation<=e-10
    Imp=Am*Imp;% account the new value of importance
    Imp=Imp./sum(Imp);% normalize the Imp
    Imp(find(sum(A1,2)==0))=10;
    Imp(find(sum(A1,2)==0))=min(Imp)-0.000001;
    E=sum((C-Imp).^2); % account the equation
    C=Imp; % renovate the value of the equation
end
Imp=Imp';
Imp=Imp./sum(Imp);

%==========compositor========
[Y,I]=sort(Imp); % sort by priority;
I=fliplr(I);
syms a b c d e f
gradation=[a;b;c;d;e;f]; % initialize the gradation of literatures
gradation=gradation(I);% account the new gradation
GI=[gradation vpa(Imp(I)',4)]
```

Appendix 2: the simulated data of the second problem

Table8 the amount of articles about every journal per year

journal \ year	2007	2006	2005	2004	2003	2002
Ⅰ	134	103	106	128	142	145
Ⅱ	110	119	140	111	115	124
Ⅲ	208	227	226	211	223	228

Table9 the number of citation papers about every journal per year

journal \ year	2007	2006	2005	2004	2003	2002
Ⅰ	1002	1035	1037	1064	1074	1054
Ⅱ	514	531	542	547	512	515
Ⅲ	1135	1220	1217	1290	1239	1224

Table10 the citation statement of every paper

journal		Ⅰ	Ⅰ	Ⅰ	Ⅰ	Ⅰ	Ⅱ	Ⅱ	Ⅱ	Ⅱ	Ⅱ	Ⅲ	Ⅲ	Ⅲ	Ⅲ	Ⅲ
	literature	**a**	**b**	**c**	**d**	**e**	**f**	**g**	**h**	**i**	**j**	**k**	**l**	**m**	**n**	**o**
Ⅰ	**a**	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
Ⅰ	**b**	0	0	1	0	0	0	0	1	0	0	0	1	0	0	0
Ⅰ	**c**	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
Ⅰ	**d**	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
Ⅰ	**e**	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0
Ⅱ	**f**	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Ⅱ	**g**	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0
Ⅱ	**h**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ⅱ	**i**	0	1	0	0	1	0	0	1	0	0	1	0	0	0	0
Ⅱ	**j**	0	1	0	0	0	1	0	0	0	0	0	1	0	0	0
Ⅲ	**k**	0	1	1	1	0	0	0	1	0	1	0	0	0	1	0
Ⅲ	**l**	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0
Ⅲ	**m**	0	1	0	1	1	1	0	1	1	0	0	0	0	0	0
Ⅲ	**n**	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0
Ⅲ	**o**	0	0	0	0	0	0	0	1	0	1	0	0	0	1	0

Note: A_{ij} express whether paper i is cited by paper j or not.

Appendix 3: the process of the third problem

```
function [Important, IGCP]=com_imp2
%year 2007 2006 2005 2004 2003 2002
NO_p=[134  103  106  128  142  145
      110  119  140  111  115  124
      208  227  226  211  223  228];
NO_c=[1002 1035 1037 1064 1074 1054
      514  531  542  547  512  515
      1135 1220 1217 1290 1239 1224];
pc=[1:6]';
pc=1./pc;
la=(NO_p*pc).\(NO_c*pc);
I0=la./((5*ones(1,3))*la);
afa=la./mean(la);
```

```
% A1 is the abut matrix
%   a b c d e f g h i j k l m n o
A1=[0 0 0 0 0 1 0 0 0 0 0 0 0 0 0
    0 0 1 0 0 0 0 1 0 0 0 1 0 0 0
    0 0 0 0 0 0 0 1 0 0 0 0 0 0 0
    0 0 0 0 0 1 0 0 0 0 0 0 0 0 0
    1 0 0 1 0 0 0 0 0 0 0 0 0 0 0
    0 0 1 0 0 0 0 0 0 0 0 0 0 0 0
    0 0 0 1 0 0 0 0 0 0 0 0 0 1 0
    0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
    0 1 0 0 1 0 0 1 0 0 1 0 0 0 0
    0 1 0 0 0 1 0 0 0 0 0 1 0 0 0
    0 1 1 1 0 0 0 1 0 1 0 0 0 1 0
    0 0 1 0 1 0 0 0 0 0 0 0 0 0 0
    0 1 0 1 1 1 0 1 1 0 0 0 0 0 0
    0 0 0 0 0 0 0 1 0 0 0 1 0 0 0
    0 0 0 0 0 0 0 1 0 1 0 0 0 1 0];

% ==compute the tree matrix Am==
 A1=A1';
 e=1; % e is a flag to judge if AA==0 and dap the loop
 i=1;
 while e~=0
     AA(:,:,i)=A1^i;
     i=i+1;
     e=sum(sum(A1^i));
 end
BB(:,:,2:i-1)=AA(:,:,1:i-2);
BB=repmat(sum(BB,2),1,15);
BB(BB==0)=1;
AA=AA./BB;
Am=sum(AA,3);
Am=Am+eye(size(Am));

% ==compute the value of Imp==
% initialize the value of the importance
Imp=[I0(1)*ones(5,1);I0(2)*ones(5,1);I0(3)*ones(5,1)];
AFA=[afa(1)*ones(5,1);afa(2)*ones(5,1);afa(3)*ones(5,1)];
C=Imp;
imp=Imp;
E=1; % initialize the value of the equation
while E>1e-5 % Dap the loop while the equation<=e-10
    Imp=Am*(Imp.*AFA);% account the new value of importance
```

```
    Imp=Imp./sum(Imp);% normalize the Imp
    E=sum((C-Imp).^2); % account the equation
    C=Imp; % renovate the value of the equation
    Imp(find(sum(A1')==0))=imp(find(sum(A1')==0));
end
Imp(find(sum(A1')==0))=imp(find(sum(A1')==0))./6;
Important=Imp;
Important=Important';
Imp=Imp./sum(Imp);
%==========compositor========
[Y,I]=sort(Imp); % sort by priority;
I=flipud(I);
Imp=Imp(I);
syms a b c d e f g h ii jj k l m n o
% initialize the gradation of literatures
gradation=[a;b;c;d;e;f;g;h;ii;jj;k;l;m;n;o];
gradation=gradation(I);% account the new gradation

A1=A1';
syms a b c d e f g h ii jj k l m n o
Ac=[a b c d e f g h ii jj k l m n o];
Ac=repmat(Ac,15,1);
Ac=A1'.*Ac;
Ac=sum(Ac,2);
Ac=Ac(I);
Ap=[a b c d e f g h ii jj k l m n o];
Ap=repmat(Ap,15,1);
Ap=Ap.*A1;
Ap=sum(Ap,2);
Ap=Ap(I);
Journal=[ones(5,1);2*ones(5,1);3*ones(5,1)];
Journal=Journal(I);


IGCP=[vpa(Imp,5) Journal gradation Ac Ap];
```