Is the Article Crucial to My Research? Evaluating Task-Oriented Impacts of Scientific Articles in Information Seeking

Jiepu Jiang, Daqing He, Shuguang Han, Wei Jeng
School of Information Sciences, University of Pittsburgh
jiepu.jiang@gmail.com, dah44@pitt.edu, shh69@pitt.edu, wej9@pitt.edu

---

Abstract

We propose a new aspect of evaluating scientific articles: crucialness, which refers to the state of articles being not only useful, but also scarce (difficult to be found or identified by scientists). Compared with the popularity-based metrics, crucialness may be a better metric in helping scientists' information seeking and use because it identifies scientists' difficulties in information seeking and reveals the crucial articles that may help scientists succeed in research. Some preliminary results are presented and discussed.

Keywords: crucialness; article evaluation; bibliometrics; information seeking; task.

Introduction

An abundance of metrics evaluate scientific resources based on their popularities, which can be measured by the number of citations (Garfield, 1972), the download statistics (Bollen, Van de Sompel, Hagberg, & Chute, 2009) and web 2.0 data (Jiang, He, & Ni, 2011; Priem & Hemminger, 2010). Studies (Wang & White, 1999) found these metrics can influence scholars' information seeking and use. However, it remains unclear to what extent these metrics can help scientists and contribute to scientific discovery.

We argue that the popularity-based metrics may fail to solve two essential issues in scientists' information seeking. First, the popular resources identified by these metrics (such as the highly cited articles and the journals with high impact factors) are usually well-known in the community. In practice, it is less likely that scientists need to find these resources, unless the scientists are new to the field. Instead, some useful but unpopular articles may deserve more attention and should be revealed to the community. Second, although an article's popularity indicates high usage of the article, it is not necessarily that the article will be beneficial to the scientists who use it. In addition to popularity, we need to consider the actual effectiveness of using an article.

Assuming some important and useful articles are naturally difficult to be found, while some elite scientists have better abilities of finding these articles. Therefore, there should exist such a period of time that those articles are not noticed by the majority of scientists, while attract some elite scientists' attention. We refer to such state of articles as ***crucialness***, which implies not only high quality but also ***scarcity*** of the articles. Scarcity is a temporary state that the article is used limitedly and biasedly in the community. These articles are crucial to scientists' research because they solve the majority of scientists' difficulties in information seeking and shorten the gap between elite scientists and others. The crucial articles may finally turn into popular ones, as other scientists will gradually know these articles because of the elite scientists' references to these articles and recommendations. Thus, revealing these articles can also speed up knowledge sharing among scientific community.

In this paper, we propose a calculable metric indicating the crucialness. Section 2 discusses the metric. Section 3 reports some preliminary results.

A Metric for Crucialness

A General Framework for Task-Oriented Metrics

We start with introducing a framework for evaluating the contribution of resources to the completion of a task. Without loss of generality, let X be the task of interest that makes users search and use resource R. We assume the task will be either completed successfully or not (X and $\bar{X}$).

We define $\dfrac{P(R|X)}{P(R)}$ as a metric that indicates the crucialness of the resource R for users' successful completion of the task X: $P(R|X)$ is the usage rate of R when users have successfully completed X; $P(R)$ is the overall usage rate of R when users are trying to complete X (no matter X is successfully completed or not).

$$\frac{P(R|X)}{P(R)} \propto \frac{P(R|X) \cdot P(X)}{P(R)} = P(X|R) \tag{1}$$

As in Eq(1), applying Bayes' theorem, one can notice that the proposed metric is proportional to $P(X|R)$, the probability that X can be completed successfully given we know R has been used. Once we select the task X, $P(X)$ is a constant and can be ignored when comparing different articles. If we know using R is a factor influencing X's success, the metric indicates the contribution of R for the successful completion of task X: $\dfrac{P(R|X)}{P(R)} > 1$ indicates that using R is positively associated with the successful completion of X.

Calculating Crucialness from Citation

Based on the framework introduced in the above section, we can come to a metric of crucialness by defining X and R as follows:

R　　　　An article used by scientists.

X　　　　Scientists write and publish new articles.

To make the metric calculable, we make the following settings: X is counted as successfully completed if the published article is highly cited in the domain (referred to as **HC articles**); we say R is used for completing X if R is cited by X. Practically, in a dataset, we select the top k% cited articles of a domain each year as the *HC* articles. We use R and X to characterize a task scenario that is ubiquitous among scientists: finding crucial articles that are helpful for conducting highly influential research (and consequently publishing highly influential and cited articles). Here we only consider the case that we have publication and citation data several years later than R so that the metric is easily calculable. The in-time prediction of the metric is left as a future work.

The crucialness metric is calculated as Eq(2): $\dfrac{n_{HC}}{N_{HC}}$ and $\dfrac{n}{N}$ are the maximum likelihood estimations for $P(R|X)$ and $P(R)$.

$$crucialness(R) = \frac{P(R|X)}{P(R)} = \frac{\frac{n_{HC}}{N_{HC}}}{\frac{n}{N}} \tag{2}$$

n_{HC}　　　The number of times R being cited by HC articles.
N_{HC}　　　The total number of HC articles.
n　　　　The number of times R being cited.
N　　　　The total number of articles.

Why is Eq(2) a metric indicating the crucialness of R? In a static dataset, when comparing the crucialness values of different articles, N_{HC} and N are constant and can be ignored in comparison. Thus, the metric in Eq(2) is proportional to $\dfrac{n_{HC}}{n}$, as in Eq(3). This makes certain connections between Eq(2) and our definition of crucialness. n_{HC} and n are citations of R by HC articles and normal ones. If we consider authors of HC articles as an approximation of the elite scientists and authors of normal articles as the group of normal scientists, n_{HC} and n can roughly indicate the number of elite scientists and normal scientists found and used R in their research.

$$crucialness(R) = \frac{P(R|X)}{P(R)} = \frac{\frac{n_{HC}}{N_{HC}}}{\frac{n}{N}} \propto \frac{n_{HC}}{n} \approx \frac{the\ number\ of\ elite\ scientists\ used\ R}{the\ number\ of\ scientists\ used\ R} \tag{3}$$

The following two assumptions (H_a and H_b) make the crucialness metric, literally defined based on conditional probabilities, indicative of certain causal relationship between "using R" and "the successful completion of X".

H_a   Authors of *HC* articles (the elite scientists) and others differ in their abilities of assessing the values of scientific articles.

H_b   Authors of *HC* articles (the elite scientists) and others differ in their abilities of finding useful scientific articles.

We assume that finding and assessing useful articles are difficult for the majority of scientists and require expertise, and the authors of *HC* articles (the elite scientists) do have better such expertise. H_a and H_b assume differences of scientists in their abilities of assessing and finding information, respectively. Let us also assume that scholars can freely search for articles and use articles; then, H_a and H_b result in the differences of article usage between authors of HC articles and the normal ones.

If we take H_a, n_{HC} and n indicate the judged quality of articles by scientists of different assessing abilities. In such case, articles with high crucialness values are those recognized high by elite scientists while low by normal scientists. If we take H_b, n_{HC} and n indicate the difficulties of finding the articles by scientists of different information seeking abilities. If the crucialness value > 1, a higher crucialness value suggests a larger gap between elite scientists and others in finding the articles. If either H_a or H_b is true, the metric indicates the crucialness of articles defined in our paper.

Some Properties of the Metric

The crucialness metric has many good properties that are beneficial for scholars' seeking and assessing of academic information:

(1) As discussed above, the metric is proportional to $P(X|R)$. With proper assumptions, the metric measures the contribution of articles to scholars' success in research.

(2) As it is defined in the form of the ratio of two probabilities, i.e. $P(R|X)$ and $P(R)$, the metric is scaled by the productivity of the communities interested in R. Thus, articles from fields of different levels of citation rates can have comparable crucialness values (for example, articles from history and physics). This unique property makes the crucialness metric helpful for scholars when related articles come from multiple topics that have different levels of citation rates. In such cases, articles identified by popularities may be biased to those from topics with high citation rates.

Unsolved Issues

Although the crucialness metric has numerous useful properties, in this initial work, it is limited in its calculation methods.

(1) We solely use citation frequency as the indicators for HC articles.

(2) We solely use "citing" as evidences of whether scientists used one article or not.

(3) Apparently, we need to wait several years until we can aggregate enough citations to calculate the crucialness value.

(4) The two probabilities in Eq(2) are estimated by maximum likelihood estimation, which can be inaccurate if a small sample size is given (e.g. for articles with very low citation frequency). Such technical issues make it only possible, at current stage, to calculate the crucialness value for highly cited articles.

Solutions to these issues, especially effective in-time prediction of the crucialness values, may finally make the metric suitable for practical use. These issues are left as further works.

Preliminary Results

We select 39 journals from library and information science field from web of knowledge (WoK) "Information Science/Library Science" (IS/LS) category. In JCR 2009, the WoK IS/LS field includes 65 journals; we remove the following journals: (1) journals from information system domain; (2) journals from domain specific informatics (e.g. GIS, medical informatics). We remove these journals because the topics and articles related to these journals may largely use articles from other disciplines, which can result in inaccurate calculation of the crucialness values in our experiment. Articles for the 39 journals from 1954

to 2011 are collected. Only articles with "DT" field in WoK database equals to "article" or "proceeding article" are used in our study (other articles may include editorials and book reviews etc.).

In calculation of crucialness values, k is set to 20 (so that the top 20% cited articles each year will be selected as HC articles). We further normalize the citation frequency by topics. We calculate topic model of articles using LDA (we use both articles' titles and abstracts). We tested for different numbers of topics to be extracted, and find 50 topics are appropriate for our study: fewer topics (e.g. 20) is usually not enough to identify fine-grained topics, and more topics (e.g. 100) will lead to insufficient sample size for our study. We further manually analyze and label the 50 topics, of which 24 are further excluded from our study because: (1) the topics are representative of fields in information system or domain specific informatics (although we removed journals in these topics, there are journals that publish articles of multiple topics); (2) topics with lots of common words and cannot be labeled reasonably; (3) the smallest 5 topics in size. Articles from the remained 26 topics are mainly related to: search and information retrieval; information seeking and use; scientometrics and bibliometrics; librarianship; LIS and education; digital library; semantic web; and other small topics.

Table 1 shows citations and crucialness values for the top 10 cited articles from 1995 to 2010 in the dataset. As showed in the table, highly popular (cited) articles do differ a lot in their crucialness values, which may suggest their different values for scientists in information seeking. Some of the widely cited articles have even lower than 1 crucialness values, indicating these articles are popular but may not contribute to scientists' success in research.

Our study suggests not all popular articles should be presented to scientists, and argues the usefulness of current popularity-based metrics in helping scientists' information seeking and use. As we discussed, the proposed metric has many good properties related to scientists' information seeking and use. However, the metric is not yet mature enough to be applied to practical systems due to the lack of in-time calculation methods, which will be the focus of our further studies.

Table 1. Crucialness values and citations of top 10 cited articles from 1995 to 2010.

Cite	Crucialness	Year	Author(s)	Title
654	1.67	1995	S. Taylor; P. A. Todd	Understanding Information Technology Usage: A Test of Competing Models
628	0.38	2003	V. Venkatesh; M. G. Morris; G. B. Davis; F. D. Davis	User Acceptance of Information Technology: Toward a Unified View
441	1.21	1995	D. R. Compeau; C. A. Higgins	Computer Self-Efficacy: Development of a Measure and Initial Test
360	1.79	1995	D. L. Goodhue; R. L. Thompson	Task-Technology Fit and Individual Performance
351	2.26	2000	B. J. Jansen; A. Spink; T. Saracevic	Real Life, Real Users, and Real Needs: a Study and Analysis of User Queries on the Web
348	0.49	2003	W. H. DeLone; E. R. McLean	The DeLone and McLean Model of Information Systems Success: A Ten-Year Update
338	1.14	1999	D. W. Bates; J. M. Teich; J. Lee; D. Seger; G. J. Kuperman; N. Ma'Luf; D. Boyle; L. Leape	The Impact of Computerized Physician Order Entry on Medication Error Prevention
325	1.38	1999	E. Karahanna; D. W. Straub; N. L. Chervany	Information Technology Adoption across Time: A Cross-sectional Comparison of Pre-adoption and Post-adoption Beliefs
310	1.38	1999	H. K. Klein; M. D. Myers	A Set of Principles for Conducting and Evaluating Interpretive Field Studies in Information Systems
274	1.17	2000	R. Agarwal; E. Karahanna	Time Flies When You're Having Fun: Cognitive Absorption and Beliefs about Information Technology Usage

References

Bollen, J., Van de Sompel, H., Hagberg, A., & Chute, R. (2009). A Principal Component Analysis of 39 Scientific Impact Measures. *PLoS ONE*, *4*(6), e6022.

Garfield, E. (1972). Citation analysis as a tool in journal evaluation. *Science*, *178*(4060), 471–479.

Jiang, J., He, D., & Ni, C. (2011). Social reference: aggregating online usage of scientific literature in CiteULike for clustering academic resources. *Proceedings of the 11th annual international ACM/IEEE joint conference on Digital libraries (JCDL'11)* (pp. 401–402).

Priem, J., & Hemminger, B. H. (2010). Scientometrics 2.0: New metrics of scholarly impact on the social Web. *First Monday*, *15*.

Wang, P., & White, M. D. (1999). A cognitive model of document use during a research project. Study II. Decisions at the reading and citing stages. *Journal of the American Society for Information Science*, *50*(2), 98–114.