Measuring hot topics in sciences

Since the $h$-index\(^1\) was introduced in 2005, it has been mainly applied to evaluate scientists\(^2,3\). However, this application does not in itself lead to improvements or progress in academic research. A search in Web of Science (WoS) reveals that there are different $h$-indices on different topics, which provide meaningful metrics in scientific studies for measuring hot topics\(^4,5\).

Following the definition of original $h$-index, we can define a hot-degree for measuring the hot topics (including titles, subjects, keywords and so on) in sciences, as follows: ‘The hot-degree of a topic is equal to $H$, if $H$ is the largest number of publications each with citations at least equal to $H$ on the study of the topic’.

Simply, $H$ of a topic means that there are $H$ publications at most, each with $H$ citations at least on the topic. Some values of $H$ for various topics are: Cosmos, 62; Dark energy, 222; Dark matter, 292; Higgs, 204; Quantum gravity, 172; Human brain, 516 and Human genome, 457 (data from 3 November 2012, WoS).

Also, $H$ can be searched on titles. In WoS, title $H$ will be smaller than corresponding topic $H$, because topic indexing covers title in the database.

Obviously, each topic (including title, subject, etc.) has its unique $H$ and is database-related. Different databases will give different results as there are differences in their data indexing, so that there are different numbers of $H$ values in databases WoS and Scopus. However, as the $h$-index is robust, the differences of $H$ will be small when we search by similar paths. For example, when we search the same words on ‘title’ in WoS and on ‘article title’ in Scopus, the results of their $H$ values, separately as $H$-WoS and $H$-Scopus, are shown in Figure 1 (WoS contains SCI, SSCI and AHCI; Scopus includes all subject areas; both cover all years).

Although different databases give different results, their highly cited papers get most coherence. (We list the top 1 papers in the physical sciences searched by titles in Appendix 1, in which most records are in coherence, with one exception.)

The $H$ is also time-related. Let $P$ be total publications and $C$ total citations, then we have similar dynamic formula\(^6\) of $h$-index as follows, in the Lotkaian informetric system\(^7\). If there exists a power law relation between $P$ and $C$ (Heaps law or Herdan law), it is generally correct\(^8\).

$$H = ((1 - a')^{1/a})P^{1/a},$$

where $a$ is the ageing rate and $\alpha$ is the exponent of Lotka’s law. According to eq. (1) for definite $a$ and $\alpha$, $H$ will increase quickly in the beginning and become almost stable in the end.

Concluding, the hot-degree $H$ could become an effective indicator for measuring hot topics. Higher the $H$ value of a topic, more will be the studies on the topic and more will be its impact. With $H$ values, we may investigate the hot topic distribution of research in the academic field, which would benefit topic selection and could stimulate truly scientific studies and academic progress.

Appendix 1. Top 1 papers in the physical sciences

Title 1: Cosmos

Title 2: Dark energy

Title 3: Dark matter

Title 4: Higgs


Title 5: Quantum gravity


Fred Y. Ye
School of Information Management, Nanjing University, Nanjing 210 093, China e-mail: yye@nju.edu.cn

Figure 1. Different databases give different $H$ values.