

## Evaluating Research and Patenting Performance Using Elites: A Preliminary Classification Scheme

Chung-Huei Kuan, Ta-Chan Chiang

Graduate Institute of Patent Research, National Taiwan University of Science and Technology, Taipei, Taiwan, ROC

**Abstract**--Using a set of elite publications as representatives to an entity's research output is a common practice in bibliometrics. There are however few studies using such concept of *elites* for patenting performance evaluation. This paper gathered a number of elite-based bibliometric approaches and organized them into a simple classification scheme so as to observe the various approaches in a systematic manner. According to the scheme, the various elite-based methods can be categorized into those using individual entities' elite sets and those using a combined elite set. These two major categories can be further divided into those using fixed, variable, and *h*-type thresholds, and those calculating size-, citation-, and contribution-based indices for assessment. This classification scheme provides us hints about possible directions of designing elite-based research and patenting performance evaluation methods.

### I. BACKGROUND

Even though there can never be a perfect performance evaluation method that satisfies all involved parties and required demands, performance evaluation is inevitable in our lives so that decision can be made and resource can be allocated accordingly.

The discipline *bibliometrics* provides a wealth of tools, indicators, and methodologies for the evaluation of researchers, journals, institutions, or nations (i.e., entities) in terms of their research publications (hereinafter, *research performance evaluation*). Among them, a type of approaches is to use a set of *elites* as representatives and the entities are evaluated in terms of their elites. For example, the number of highly cited papers (HCPs) as a measure to a researcher's performance is a form of using elites (cf. [1][41]). The famous *h*-index [24] is another example where the so-called *h*-core [42] is the set of elites of the entity's research output.

Vinkler [48, 49] introduced the *elite set* concept in determining the eminence of scientific journals. Vinkler proposed several ways to determine a journal's elite set such as using the Lotka law [33], the citation rate, the *h*-index, or the  $(10 \log P) - 10$  most highly cited publications where  $P$  is the total number of publications.

The idea of using a few outstanding publications (i.e., elites) to represent the entire research output would be a reasonable choice when most publications can only produce few citations. In other words, using elites would be best when entities have significant *skewedness* in terms of the distribution of citations among their publications. One of the greatest benefits of using elites is that the scales (i.e., sizes) of entities are of little influence. When only the elites are considered, small institutes can still outperform large institutes.

Another type of performance evaluation is the evaluation of patent assignees in terms of their patents (hereafter, *patenting performance evaluation*). This evaluation can help us, for example, to gauge the technological strength of a competitor, to assess the technological position of a company within a technology sector. Patent data, which are structurally organized and substantially objective, are well recognized as a viable source of technological intelligence for various competitor analysis and technology management tasks such as tracing knowledge diffusion [2, 10, 11], strategic planning [4, 17, 32], technology analysis [35, 36], technological forecasting [5, 16, 32], finding relationship among companies and industries [25], and even providing assessment for various aspects of merger and acquisition [8].

Narin, Carpenter, and Woolf [38] suggested that the number of patents and patent citations can be interpreted as productivity and quality of technological performance. Narin [37] then pioneered the term *patent bibliometric* and suggested that the traditional bibliometric approaches are applicable to patent data as well. Even though there were concerns about such adaption (cf. [34]), a large number of methods are borrowed from bibliometrics and applied to patent data with or without modification (cf. [16][39][46][53]). For example, the *h*-index, even though originally designed to evaluate researchers, is extended to evaluate patent assignees by Guan and Gao [22], Kuan, Huang, and Chen [29, 30, 31], etc.

The concept of using elites for performance evaluation should be applicable to the patent assignees as well as the patent portfolios of patent assignees possess even more skewed distribution of citations. Narin in his pioneering work [37] pointed out that the distribution of citations among patents is similar to the Lotka's Law [33] of research publications, but even more skewed. Hall [23] found that, among 1995 U.S. patents, only 0.01% of them has more than 100 citations, and one fourth of them has no citation at all. Similarly, Silverberg and Verspagen [45] used patents from European Patent Office (EPO) and U.S. Patent and Trademark Office (USPTO) and found that the distribution of citations among patents has a "fat tail" (i.e., most patents are of few or no citation).

We noticed that there are few studies using elites for patenting performance evaluation. In this study, we therefore gathered a number of elite-based bibliometric approaches and organized them in a systematic way before jumping immediately into the design of a new evaluation method for patent assignees using elites. By doing so we are hoping to obtain hints about possible directions of designing elite-based research and patenting performance evaluation methods.

II. NOTATION AND GEOMETRY

The rank-citation curve [51] is a valuable tool for graphically illustrating an entity’s distribution of citations among its publications and for understanding various elite-based methods.

Let an entity has produced  $N$  publications and these publications are sorted in descending order of their respective citation counts into an ordered list  $\{P_1, P_2, \dots, P_{N-1}, P_N\}$ , where  $P_i$  is the publication ranked at the  $i$ th place,  $C(P_i) \geq 0$  is the citation count of  $P_i$ , and  $C(P_i) \geq C(P_j)$  if  $i \leq j$ . The rank-citation curve is obtained by plotting and connecting the points  $(i, C(P_i))$ ,  $1 \leq i \leq N$ , together in a two dimensional coordinate system where the horizontal axis is the rank of the publications and the vertical axis is the citation count. For example, Fig. 1 shows the rank-citation curves of the top 100 assignees having the greatest numbers of U.S. patents granted in the year 2009 where only the 500 most frequently cited patents of their portfolios are shown for simplicity’s sake.

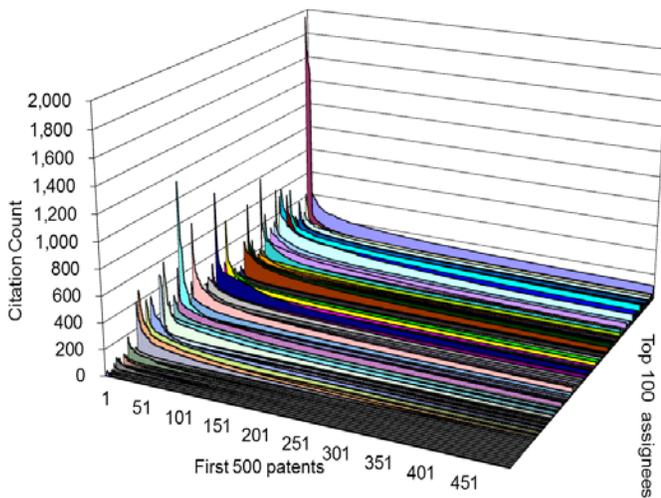


Fig. 1: The rank-citation curves of the top 100 U.S. assignees of year 2009.

The rank-citation curve provides a systematic view to various elite-based evaluation methods. For example, the  $\pi$  index [47] is an elite-based evaluation method for researchers where a researcher has total  $N$  publications and the set of elite publications  $ES$  of the researcher is top  $\sqrt{N}$  most frequently cited publications:

$$ES = \{P_i | i \leq \sqrt{N}\}.$$

The  $\pi$  index is then calculated as the sum of the citation counts of the elites:

$$\pi = \sum_{i \in ES} C(P_i).$$

The  $\pi$  index can be conveniently understood using the rank-citation curve and is equal to the shaded area of Fig. 2.

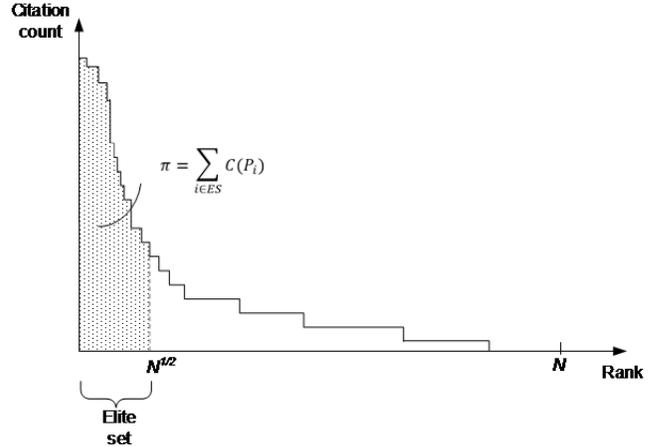


Fig. 2: A fictitious entity’s rank citation curve and its  $\pi$  index.

For another example, an entity has  $h$ -index  $n$  if it has at least  $n$  publications, each receiving at least  $n$  citations. Therefore the elite set  $ES$  can be expressed as follows:

$$ES = \{P_1, P_2, \dots, P_n | C(P_n) \geq n, C(P_{n+1}) \leq n\}.$$

The  $h$ -index  $n$  is equal to the size of the elite set  $ES$ :

$$h = |ES| = n.$$

As shown in Fig. 3 and according to the above definition of the  $h$ -index, the intersection point  $(n, n)$  between the rank-citation curve and the line  $y=x$  determines an entity’s  $h$ -index and therefore its elite set. Fig. 3 depicts two rank-citation curves of fictitious entities A and B with  $h$ -indices  $n_A$  and  $n_B$ , respectively.

The various types of elite-based methods are best understood by picturing them using rank-citation curves. However, for brevity’s sake, they are not included in the following discussion.

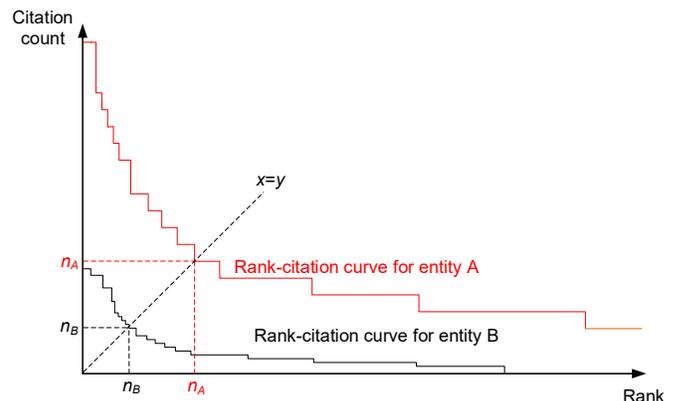


Fig. 3: Two fictitious entities’ rank citation curves and their  $h$ -indices.

III. CLASSIFICATION SCHEME

The various elite-based methods can be categorized into two major species. One specie contains those methods where,

for a set of  $M$  entities  $E_1, E_2, \dots, E_M$ , their *individual elite sets*  $ES_i$  are determined first, and then these elite sets  $ES_i$  are compared against each other. The other species are those that a *combined elite set* is determined from the publications of the  $M$  entities first. Then, these entities are assessed by how they *contribute* to the combined elite set. Most elite-based methods belong to the former and only a few belongs to the latter.

From another point of view, we can see that an elite-based evaluation method, whether it is an individual-elite-set or combined-elite-set method, involves two essential ingredients: (1) a *threshold* in determining the elites; and (2) an *index* calculated from the elites.

Some prior studies proposed a threshold but didn't specify a related index. These studies are still considered in this paper. In addition, since the various thresholds and indices should be applicable to both individual-elite-set and combined-elite-set methods, in the following we mainly use individual-elite-set methods as examples.

#### A. Fixed Thresholds

Glanzel and Schubert [21], in determining the highly cited publications, proposed that a publication is considered highly cited (i.e., an elite) if (a) it has received at least  $c$  citations; and (b) it has received at least  $k\bar{C}$  citations where  $c, k$  are constants, and  $\bar{C}$  is the average citations of all publications. Glanzel and Schubert gave us a hint that the various elite-based evaluation methods can be divided into two categories: (1) threshold is fixed at a specific constant (hereinafter *fixed threshold*); and (2) threshold is varied according to some feature of the evaluated or related entities (hereinafter *variable threshold*).

The fixed threshold methods can be further divided into two sub-categories where the constant is related to: (1) the ranks of publications; or (2) the citations of publications. The first sub-category is therefore referred to as *rank-based fixed threshold*, which can be expressed as:

$$ES_j = \{P_{ij} | i \leq k\},$$

whereas the second one is referred to as *citation-based fixed threshold*, which can be expressed as:

$$ES_j = \{P_{ij} | C(P_{ij}) \geq c\}.$$

As shown in Fig. 4, the rank-based fixed threshold is a vertical line at  $k$  along the rank axis and those publications to the left of the line are considered as elites. The citation-based fixed threshold is a horizontal line at  $c$  along the citation count axis, and those publications having citation counts above the line are considered as elite.

A number of studies using the rank-based fixed threshold methods are as follows. Garfield [20] considered the 100 most frequently cited life science publications published in 1975 as elites. Similarly, Frogel [19] selected the first, the first 50, and the first 100 most frequently cited astronomy publications as elites. Ryan and Woodall [44] applied the same concept to statistics publications with the rank threshold set at 25. Patsopoulos, Ioannidis, and Analatos [40] chose 30

as the rank threshold for medicine-related publications.

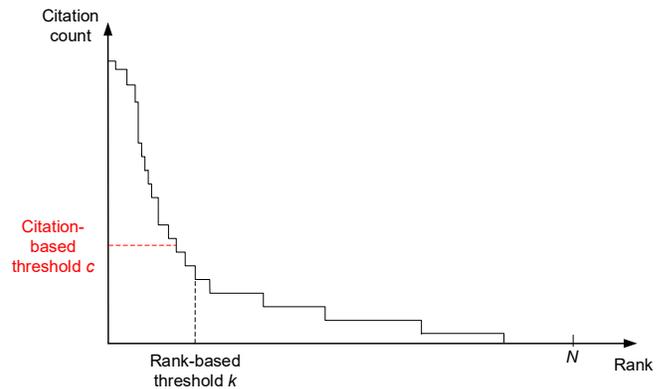


Fig. 4: Rank- and citation-based thresholds.

Some studies using citation-based fixed thresholds are as follows. Plomp [41] considered a researcher's elite publications are those receiving at least 25 citations. The i10 and i100 indices of Google Scholar<sup>1</sup> uses fixed citation thresholds 10 and 100. Blessinger and Hryca [6] used 10 and 50 citations as criteria to generate two groups of elite publications. Garfield [20] set the fixed citation threshold at 10.

A greatest advantage of fixed thresholds is that the elites of all evaluated entities are extracted using a uniform criterion and the elite sets can be compared on a common ground. However, the choice and justification to a particular fixed threshold is more complicated and difficult. For example, some disciplines usually produce and accumulate large numbers of publications and citations in a shorter period of time. Therefore a smaller rank threshold or a greater citation threshold is used for these disciplines.

#### B. Variable Thresholds

The variable threshold methods have thresholds varied from entity to entity. There are also *rank-based variable thresholds* where a general form can be expressed as:

$$ES_j = \{P_{ij} | i \leq f(N_j)\}.$$

The various rank- and citation-based variable threshold methods differ in the functions  $f$ 's used. The Highly cited papers, Hot papers, ESI most cited papers, etc. of Thomson Reuters<sup>2</sup> uses variable rank thresholds with functions  $0.01\%N_j, 0.1\%N_j, 1\%N_j$ , respectively. Similarly, Fernandez-Alles and Ramos-Rodríguez [18] used a function  $1.45\%N_j$ . In finding the elite researchers, as mentioned earlier, the  $\pi$ -index of Vinkler [47] considered only the top  $\sqrt{N_j}$  most frequently cited publications of the evaluated researchers. Vinkler's another  $\pi_v$  index [48] is for evaluating journals, and each

<sup>1</sup> Google Inc. (n.d.). *Google scholar citations open to all*. Retrieved January 28, 2014 from <http://googlescholar.blogspot.ca/2011/11/google-scholar-citations-open-to-all.html>.

<sup>2</sup> Thomson Reuters. (n.d.). *Essential science indicator*. Retrieved October 6, 2014, from <http://thomsonreuters.com/essential-science-indicators/>.

journal is assessed by its  $(10 \log N_j) - 10$  most frequently cited publications.

There are however few *citation-based variable threshold* methods. One example is that an entity  $j$ 's elite publications are most frequently cited publications jointly producing a certain percentage of the entity  $j$ 's all citations  $C_j$ . Another example (cf. [21]) is that the elite publications are those receiving at least  $k\bar{C}$  citations where  $\bar{C}$  is the average number of citations. For these methods using variable citation thresholds, their elite set can be expressed generally as:

$$ES_j = \{P_{ij} | C(P_{ij}) \geq f(C_j)\}.$$

We can see that variable thresholds can be very flexible and are adaptable to different entities. Each individual entity therefore can have their own elite sets and small entities are not overwhelmed by large entities. The disadvantage however is that, as there is not a single uniform criterion, one entity's elite may be mediocre to another entity.

### C. *h-Type Thresholds*

The *h-index* is also an elite-based indicator. Its threshold is neither rank-based nor citation-base, but a combination of the two. We therefore grouped *h-index* and other similar indicators into a separate category and referred to their thresholds as *h-type thresholds*.

Since its introduction, the *h-index* has quickly become a de facto indicator for research performance evaluation, as evident from the significant number of related articles and its adoption by on-line databases such as Scopus and Web of Science. The popularity of the *h-index* arises mainly out of its claimed characteristic in capturing both productivity (i.e., the number of publications published or patents granted) and impact (i.e., the citations received by published publications or granted patents) in a single number (cf. [12][24][42][43]).

*h-Index* has been mostly criticized for being insensitive to publications' excessive citations above the *h-index* and, as such, a large number of so-called *h-type indices* were proposed to address this issue and to replace or augment the original *h-index*. Some examples of these *h-type indices* are the *g-index* [13, 14], the *h<sup>(2)</sup>-index* [28], the *A-, R-, AR-indices* [26, 27], the *m-index* [7], the *e-index* [52], the *hg-index* [3], the *q<sup>2</sup>-index* [9], and the *w-index* [50]. A thorough review and comparison of these *h-type indices* can be found in Egghe [15].

The *h-type indices* can be roughly categorized as those aiming to replace the original *h-index* (e.g., the *g-, hg-, w-index*), and those aiming to supplement the original *h-index* (e.g., the *e-, A-, R-index*). For the latter, the thresholds of their elite sets are the same as the *h-index*. For the former, however, the thresholds are their respective indices. Using the *g-index* as example, an entity  $j$ 's elite set  $ES_j$  can be expressed as follows:

$$ES_j = \left\{ P_{1j}, P_{2j}, \dots, P_{nj} \mid \sum_{i \leq n} C(P_{ij}) \geq n^2, \sum_{i \leq (n+1)} C(P_{ij}) \leq n^2 \right\}$$

*h-Type thresholds* provides a uniform approach similar to the fixed thresholds. However the obtained thresholds can still vary from entity to entity. Therefore *h-type thresholds* share similar advantages as fixed thresholds (e.g., entities can be compared on a common ground) and disadvantages as the variable thresholds (e.g., one entity's elite may be mediocre to another entity).

### D. *Index*

After the entities' individual elite sets or combined elite set is determined based on one of the thresholds described above, an *index* for each entity is calculated from the elite set(s) so that these entities can be evaluated in accordance with their respective indices.

For methods belonging to the individual-elite-set species, we noticed the following indices from prior studies:

- (1) *Size-based index* where an entity  $j$ 's index  $I_j$  is determined by the number of elites (i.e.,  $|ES_j|$ ) or more generally  $f(|ES_j|)$ . Plomp (1990), the i10 and i100 indices of Google Scholar, *h-index*, etc. used  $|ES_j|$  as their indices. Another example is Leiden Ranking<sup>3</sup> by Leiden University where an institute's Impact Factor involves the proportion of top 10% publications, meaning the proportion of the top 10% most frequently cited publications (i.e.,  $|ES_j|$ ) to the institute's total publications (i.e.,  $N_j$ ). Therefore, using our notation, this index is calculated as  $|ES_j|/N_j$ .
- (2) *Citation-based index* where an entity  $j$ 's index  $I_j$  is determined by the citations received by the elites and therefore is the "area" under a segment of the rank-citation curve above the elite set (i.e.,  $\sum_{i \in ES_j} C(P_{ij})$ ) or more generally  $f(\sum_{i \in ES_j} C(P_{ij}))$ . For example, Vinkler's  $\pi$  index and  $\pi_v$  index [47, 48] are calculated as  $0.01 \sum_{i \in ES_j} C(P_{ij})$ .

These indices inherit the advantages and disadvantages of the thresholds used. For example, for indices from elite sets determined using variable thresholds, these indices may lack a common ground for comparison.

There are very few studies involving indices for the combined-elite-set species. Kuan, Huang, and Chen [31] provided two approaches. They combined the publications of all evaluated institutes together, and determined a combined elite set using *h-index*. Then, an institute is assessed by (1) how many publications of the combined elite set are produced from the institute; or (2) how many citations of the combined elite set's total citations are received by the institute. For example, there are 5 institutes to be evaluated, and the 5 institutes have produced totally 1,000 publications. The combined elite set of the 1,000 publications are the 100 most frequently cited publications. Among the 100 elite publications, 50 are from institute 1, 30 are from institute 2,

<sup>3</sup> Leiden ranking, Retrieved October 6, 2014 from <http://www.leidenranking.com/>.

10 are from institutes 3 and 4, respectively, and none is from institute 5. Then these institutes contributes 50%, 30%, 10%, 10%, and 0% of the elite publications. Kuan, Huang, and Chen referred to these indices as the *contribution ratios* of these institutes to the combined elite set, and these institutes are assessed by their respective contribution ratios.

Using our notation, there are  $M$  entities with an combined elite set  $ES$  receiving citations  $C(ES) = \sum_{i \in ES} C(P_i)$ , and

$$ES = es_1 \cup es_2 \cup es_2 \cup \dots \cup es_M$$

where  $es_j$  is a subset of the combined elite set  $ES$  produced from an entity  $j$ . Then, the entity  $j$  is evaluated using one of the following indices:

$$I_j = \frac{|es_j|}{|ES|} \text{ or}$$

$$I_j = \frac{C(es_j)}{C(ES)}$$

Kuan, Huang, and Chen pointed out that the former, share-of-size one is less discriminating as many institutes may have identical indices and cannot be differentiated. On the other hand, the latter, share-of-citation one is more sensitive but favors those entities producing exceptionally high citations. These contribution-based indices are determined from a single combined elite set and entities are assessed on a common ground.

#### IV. SUMMARY

This study arises out of the belief that elite-based method should be suitable for patenting performance evaluation. However, after surveying a number of prior studies, there are few researches related to this topic. We therefore gathered a limited set of related studies from bibliometrics and established a classification scheme so that the various elite-based methods can be observed within a consistent framework as shown in Fig. 5.

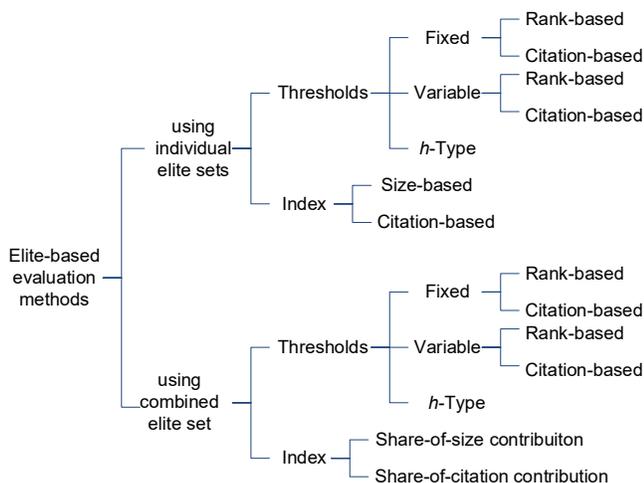


Fig. 5: A classification scheme for elite-based evaluation methods.

This classification scheme provides us hints about possible directions of designing elite-based research and patenting performance evaluation methods. For example, using the classification scheme, we observed that there are few combined-elite-set methods and these methods only use  $h$ -index as thresholds. Therefore a possible research or patenting performance evaluation method would be one using a type of variable thresholds, such as the average number of citations proposed by Glanzel and Schubert [21] to determine the combined elite set, together with the contribution-based indices.

Due to limited time, our literature review is limited to earlier studies and is definitely far from perfect. Currently we are investing more effort to gather more recent studies so as to refine and improve the above classification scheme so that it will be more useful to related researchers.

#### ACKNOWLEDGEMENTS

This study is funded by the Ministry of Science and Technology, Taiwan, R.O.C., under Grant No. MOST 104-2221-E-011-050.

#### REFERENCES

- [1] Aksnes, D. W; "Characteristics of highly cited papers," *Res. Eval.*, vol. 12, pp. 159-170, 2003.
- [2] Almeida, P.; "Knowledge sourcing by foreign multinationals: Patent citation analysis in the U.S. semiconductor industry," *Strategic Manage. J.*, vol. 17, pp. 155-165, 1996.
- [3] Alonso, S., F. J. Cabrerizo, E. Herrera-Viedma and F. Herrera, "hg-Index: A new index to characterize the scientific output of researchers based on the h- and g-indices," *Scientometrics*, vol. 82, pp. 391-400, 2010.
- [4] Ashton, W. B. and R. K. Sen, "Using patent information in technology business planning-I," *Res. Technol. Manage.*, vol. 31, pp. 42-46, 1988.
- [5] Basberg, B. L.; "Patents and the measurement of technological change: A survey of the literature," *Res. Policy*, vol. 16, pp. 131-141, 1987.
- [6] Blessinger, K. and P. Hrycaj, "Highly cited articles in library and information science: An analysis of content and authorship trends," *Libr. Infor. Sci. Res.*, vol. 32, pp. 156-162, 2010.
- [7] Bornmann, L., R. Mutz and H.-D. Daniel, "The h index research output measurement: Two approaches to enhance its accuracy," *J. Informetrics*, vol. 4, pp. 407-414, 2010.
- [8] Breitzman, A. and P. Thomas, "Using patent citation to target/value M&A candidates," *Res. Technol. Manage.*, vol. 45, pp. 28-36, 2002.
- [9] Cabrerizo, F. J., S. Alonso, E. Herrera-Viedma and F. Herrera, "q<sup>2</sup>-index: Quantitative and qualitative evaluation based on the number and impact of papers in the Hirsch core," *J. Informetrics*, vol. 4, pp. 23-28, 2010.
- [10] Chakrabarti, A. K., I. Dror and N. Eakabuse, "Interorganizational transfer of knowledge: An analysis of patent citations of a defense firm," *IEEE T. Eng. Manage.*, vol. 40, pp. 91-94, 1993.
- [11] Chen, C. and D. Hicks, "Tracing knowledge diffusion," *Scientometrics*, vol. 59, pp. 199-211, 2004.
- [12] Costas, R. and M. Bordons, "The h-index: Advantages, limitations and its relation with other bibliometric indicators at the micro level," *J. Informetrics*, vol. 1, pp. 193-203, 2007.
- [13] Egghe, L.; "An improvement of the h-index: The g-index," *ISSI Newsletter*, vol. 2, pp. 8-9, 2006.
- [14] Egghe, L.; "Theory and practise of the g-index," *Scientometrics*, vol. 69, pp. 131-152, 2006.
- [15] Egghe, L.; "The Hirsch-index and related impact measures," *Annu. Rev.*

- Inform. Sci.*, vol. 44, pp. 65-114, 2010.
- [16] Ernst, H.; "The use of patent data for technological forecasting: The diffusion of CNC-technology in the machine tool industry," *Small Bus. Econ.*, vol. 9, pp. 361-381, 1987.
- [17] Ernst, H.; "Patent information for strategic technology management," *World Patent Infor.*, vol. 25, pp. 233-242, 2003.
- [18] Fernandez-Alles, M. and A. Ramos-Rodríguez, "Intellectual structure of human resources management research: A bibliometric analysis of the journal Human Resource Management, 1985-2005," *J. Am. Soc. Inf. Sci. Tec.*, vol. 60, pp. 161-175, 2009.
- [19] Frogel, J. A.; "Astronomy's greatest hits: The 100 most cited papers in each year of the first decade of the 21st century (2000-2009)," *Publ. Astron. Soc. Pac.*, vol. 122, pp. 1214-1235, 2010.
- [20] Garfield, E.; "1975 life sciences articles highly cited in 1975," *Curr. Contents*, vol. 15, pp. 5-9, 1976.
- [21] Glanzel, W. and A. Schubert, "Some facts and figures on highly cited papers in the sciences, 1981-1985," *Scientometrics*, vol. 25, pp. 373-380, 1992.
- [22] Guan, J. C. and X. Gao, "Exploring the h-index at patent level," *J. Am. Soc. Inf. Sci. Tec.*, vol. 60, pp. 35-40, 2009.
- [23] Hall, B. H. and I. F. S. London, "Patent data as indicators," In *WIPO Conference Proceedings*, 2004.
- [24] Hirsch, J. E.; "An index to quantify an individual's scientific research output," *Proceedings of the National Academy of Sciences of United States of America*, vol. 102, pp. 16569-16572, 2005.
- [25] Huang, M.-H., L.-Y. Chiang and D.-Z. Chen, "Constructing a patent citation map using bibliographic coupling: A study of Taiwan's high-tech companies," *Scientometrics*, vol. 58, pp. 489-506, 2003.
- [26] Jin, B.; "The AR-index: Complementing the h-index," *ISSI Newsletter*, vol. 3, p. 6, 2007.
- [27] Jin, B., L. Liang, R. Rousseau and L. Egghe, "The R-and AR-indices: Complementing the h-index," *Chinese Science Bulletin*, vol. 52, pp. 855-863, 2007.
- [28] Kosmulski, M.; "A new Hirsch-type index saves time and works equally well as the original h-index," *ISSI Newsletter*, vol. 2, pp. 4-6, 2006.
- [29] Kuan, C.-H., M.-H. Huang and D.-Z. Chen, "Ranking patent assignee performance by h-index and shape descriptors," *J. Informetrics*, vol. 5, pp. 303-312, 2011.
- [30] Kuan, C.-H., M.-H. Huang and D.-Z. Chen, "Positioning research and innovation performance using shape centroids of h-core and h-tail," *J. Informetrics*, vol. 5, pp. 515-528, 2011.
- [31] Kuan, C.-H., M.-H. Huang and D.-Z. Chen, "Cross-field evaluation of publications of research institutes using their contributions to the fields' MVPs determined by h-index," *J. Informetrics*, vol. 7, pp. 455-468, 2013.
- [32] Liu, S.-J. and J. Shyu, "Strategic planning for technology development with patent analysis," *Int. J. Technol. Manage.*, vol. 13, pp. 661-680, 1997.
- [33] Lotka, A. J.; "The frequency distribution of scientific productivity," *J. Washington Academy Sci.*, vol. 16, pp. 317-323, 1926.
- [34] Meyer, M.; "What is special about patent citations? Differences between scientific and patent citations," *Scientometrics*, vol. 49, pp. 93-123, 2000.
- [35] Mogee, M. E.; "Using patent data for technology analysis and planning," *Res. Technol. Manage.*, vol. 34, pp. 43-49, 1991.
- [36] Mogee, M. E. and R.G. Kolar, "International patent analysis as a tool for corporate technology analysis and planning," *Technol. Anal. Strateg.*, vol. 6, pp. 485-503, 1994.
- [37] Narin, F.; "Patent bibliometrics," *Scientometrics*, vol. 30, pp. 147-155, 1994.
- [38] Narin, F., M. P. Carpenter and P. Woolf, "Technological performance assessments based on patents and patent citations," *IEEE T. Eng. Manage.*, EM- vol. 31, pp. 172-183, 1984.
- [39] Narin, F., E. Noma and R. Perry, "Patents as indicators of corporate technological strength," *Res. Policy*, vol. 16, pp. 143-155, 1987.
- [40] Patsopoulos, N. A., J. P. Ioannidis and A. A. Analatos, "Origin and funding of the most frequently cited papers in medicine: database analysis," *BMJ*, vol. 332, pp. 1061-1064, 2006.
- [41] Plomp, R.; "The significance of the number of highly cited papers as an indicator of scientific prolificacy," *Scientometrics*, vol. 19, pp. 185-197, 1990.
- [42] Rousseau, R.; "New developments related to the Hirsch index," *Science Focus*, vol. 1, pp. 23-25, 2006 (in Chinese). An English translation is available on-line at <http://eprints.rclis.org/6376/>.
- [43] Rousseau, R.; "Reflections on recent developments of the h-index and h-type indices," *COLLNET J. Sci. Inf. Manage.*, vol. 2, pp. 1-8, 2008.
- [44] Ryan, T. P. and W. H. Woodall, "The most-cited statistical papers," *J. Appl. Stat.*, vol. 32, pp. 461-474, 2005.
- [45] Silverberg, G. and B. Verspagen, "The size distribution of innovations revisited: An application of extreme value statistics to citation and value measures of patent significance," *J. Econometrics*, vol. 139, pp. 318-339, 2007.
- [46] Thomas, P.; "A relationship between technology indicators and stock market performance," *Scientometrics*, vol. 51, pp. 319-333, 2001.
- [47] Vinkler, P.; "The  $\pi$ -index: A new indicator for assessing scientific impact," *J. Inf. Sci.*, vol. 35, pp. 602-612, 2009.
- [48] Vinkler, P.; "The  $\pi_v$ -index: a new indicator to characterize the impact of journals," *Scientometrics*, vol. 82, pp. 461-475, 2010.
- [49] Vinkler, P.; "Application of the distribution of citations among publications in scientometric evaluations," *J. Am. Soc. Inf. Sci. Tec.*, vol. 62, pp. 1963-1978, 2011.
- [50] Wu, Q.; "The w-index: A measure to assess scientific impact by focusing on widely cited papers," *J. Am. Soc. Inf. Sci. Tec.*, vol. 61, pp. 609-614, 2010.
- [51] Ye, F. Y. and R. Rousseau, "Probing the h-core: an investigation of the tail-core ratio for rank distributions," *Scientometrics*, vol. 84, pp. 431-439, 2010.
- [52] Zhang, C.-T.; "The e-index, complementing the h-index for excess citations," *PLoS ONE*, vol. 4, e5429, 2009.
- [53] Zucker, L. G. and M. R. Darby, "California's inventive activity: Patent indicators of quantity, quality and organizational origins," 1999.