Journal of Informetrics 5 (2011) 303–312

Contents lists available at ScienceDirect

Journal of Informetrics

journal homepage: www.elsevier.com/locate/joi

Ranking patent assignee performance by *h*-index and shape descriptors

Chung-Huei Kuan[a], Mu-Hsuan Huang[b], Dar-Zen Chen[c,*]

[a] Department of Mechanical Engineering, National Taiwan University, Taipei 10617, Taiwan, ROC
[b] Department of Library and Information Science, National Taiwan University, Taipei 10617, Taiwan, ROC
[c] Department of Mechanical Engineering and Institute of Industrial Engineering, National Taiwan University, Taipei 10617, Taiwan, ROC

ARTICLE INFO

Article history:
Received 1 December 2010
Received in revised form
29 December 2010
Accepted 4 January 2011

Keywords:
h-Index
Rank-citation curve
Geometric interpretation
Shape descriptor
Patentometrics
Patent assignee ranking

ABSTRACT

We propose a geometric interpretation to the ranking of patent assignees by their *h*-indices as indicating the relative positions of their rank-citation curves. We then propose two shape descriptors characterizing the rank-citation curves over the *h*-cores and *h*-tails, respectively. Together with the *h*-indices, the shape descriptors help verifying the geometric relationship among rank-citation curves and the relative performance among the assignees' *h*-cores and *h*-tails. The geometric interpretation and shape descriptors are proven by empirical data to be reliable, accurate, robust, flexible, and insightful, and their application could be extended to research performance evaluation as well.

1. Introduction

The *h*-index (Hirsch, 2005) has become a de facto scientometric indicator for research performance evaluation in various scientific disciplines and has been adopted by databases such as Scopus and Web of Science.

The extension of the *h*-index to patent assignees seems intuitive due to the many analogous features between publications and patents (Meyer, 2000; Meyer & Bhattacharya, 2004). However, to our best knowledge, so far there is only one article (Guan & Gao, 2009) dedicated to the application of *h*-index to patent assignees, where top 20 corporations in the semiconductor sector are evaluated. According to Guan and Gao (2009), there is a poor correlation between the *h*-indices and the patent counts while the ranking by total citation counts roughly agrees with that by the *h*-indices. The paper then concluded arguably that the *h*-index "is indeed an effective indicator for evaluating the technological importance and quality, or impact, for an assignee."

Instead of treating the *h*-index as yet another patentometric indicator, we believe that the claimed characteristic of the *h*-index in capturing both productivity (related to the number of publications published/patent granted) and impact (related to the citations received) in a single yet simple-to-calculate number (Costas & Bordons, 2007; Hirsch, 2005; Rousseau, 2008) suggests that it could function as *a general scale* for ranking assignees' innovation performance where both productivity and impact are taken into consideration.

* Corresponding author at: Department of Mechanical Engineering and Institute of Industrial Engineering, National Taiwan University, No. 1, Sec. 4, Roosevelt Road, Taipei 10617, Taiwan, ROC. Tel.: +886 2 23692178; fax: +886 2 23631755.
E-mail addresses: d98522047@ntu.edu.tw (C.-H. Kuan), mhhuang@ntu.edu.tw (M.-H. Huang), dzchen@ntu.edu.tw (D.-Z. Chen).

1751-1577/$ – see front matter.
doi:10.1016/j.joi.2011.01.002

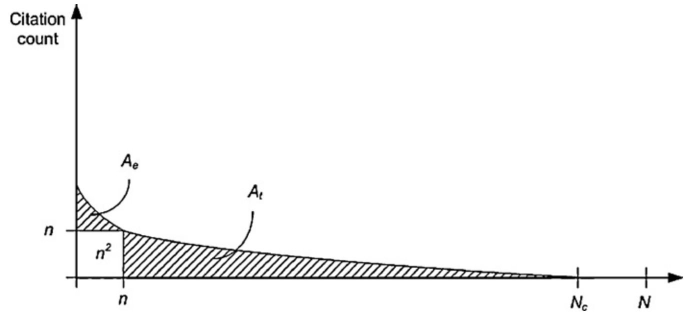

Fig. 1. Rank-citation curve of an assignee's portfolio.

Such a general scale would be highly valuable in patentometric applications. When a large number of assignees[1] are to be evaluated and compared, and each has a sizable patent portfolio, the analysis complexity and effort mounts quickly. A preferable approach is to first rank the assignees along the general scale reflecting their overall performance, and then to skip those having significant distance where their relative performance is well determined and to limit detailed analysis to the remaining, adjacent ones where their relative performance requires further investigation. The analysis complexity and effort as such is significantly reduced.

However, the adequacy of the *h*-index as the general scale is dubious. First of all, the *h*-index has been mostly criticized for being insensitive to some exceptionally highly cited papers, as evident from the large number of so-called *h*-type indices proposed to address this issue and to replace or augment the original *h*-index, such as the *g*-index (Egghe, 2006a, 2006b), the $h^{(2)}$-index (Kosmulski, 2006), the *A*-, *R*-, *AR*-indices (Jin, 2007; Jin, Liang, Rousseau, & Egghe, 2007), the *m*-index (Bornmann, Mutz, & Daniel, 2008), the h_w-index (Egghe & Rousseau, 2008), the *e*-index (Zhang, 2009), the *hg*-index (Alonso, Cabrerizo, Herrera-Viedma, & Herrera, 2010), the q^2-index (Cabrerizo, Alonso, Herrera-Viedma, & Herrera, 2010), and the *w*-index (Wu, 2010). Reviews of these *h*-type indices could be found in Bornmann et al. (2008) and Egghe (2010).

On the other hand, the *h*-index is also regarded as being insensitive to lowly cited papers (Costas & Bordons, 2007; Rousseau, 2008). However, there are only few articles taking the lowly cited papers into consideration (Bornmann, Mutz, & Daniel, 2010; García-Pérez, 2009) and, in these articles, the lowly cited papers play a supplementary role for enhancing the accuracy of the *h*-index. Yet, according to the empirical study by Hall, Jaffe, and Trajtenberg (2000, 2001), a great majority of patents are lowly cited. Therefore, for patent assignees, the lowly cited patents may constitute a significant portion of an assignee's innovation performance, and the *h*-index's ignorance to them may render it inappropriate as our general scale.

Instead of hastily giving up the *h*-index under the suspicion of its lack of information regarding both the exceptionally highly cited and lowly cited patents, we first utilize empirical data to investigate what the ranking by *h*-index really tells us and how the conventional productivity and impact measures are captured and reflected in the ranking by *h*-index. This investigation also provides comprehensive empirical information to the application of *h*-index to patent assignees that should be valuable to patentometric study.

2. Notations and research data

Let $\{P_1, P_2, \ldots, P_{N-1}, P_N\}$ be an assignee's *N* patents sorted in descending order of their respective citation counts $C(P_i)$, $1 \le i \le N$. The assignee is then said to have *h*-index *n* if $C(P_1) \ge \cdots \ge C(P_n) \ge n \ge C(P_{n+1}) \ge \cdots \ge C(P_N)$. The set of highly cited *n* patents $\{P_1, P_2, \ldots, P_{n-1}, P_n\}$ and the set of lowly cited and un-cited $(N-n)$ patents $\{P_{n+1}, P_{n+2}, \ldots, P_{N-1}, P_N\}$ are referred to as the assignee's *h*-core (Rousseau, 2006) and *h*-tail (Ye & Rousseau, 2010), respectively. The *h*-tail can be further divided into two subsets: the lowly cited patents $\{P_{n+1}, P_{n+2}, \ldots, P_{N_c-1}, P_{N_c}\}$ and the un-cited patents $\{P_{N_c+1}, P_{N_c+2}, \ldots, P_{N-1}, P_N\}$, where N_c stands for the number of cited patents (i.e., patents having at least one citation). As un-cited patents are not associated with any impact information, we consider only cited patents and the term "*h*-tail" is referred only to the lowly cited patents $\{P_{n+1}, P_{n+2}, \ldots, P_{N_c-1}, P_{N_c}\}$ hereinafter in this paper.

The assignee's *rank-citation curve* (Ye & Rousseau, 2010) is obtained by arranging $\{P_1, P_2, \ldots, P_{N-1}, P_N\}$ along the horizontal axis and plotting their respective citation counts $C(P_i)$ against the left axis. Without losing generality, an exemplary rank-citation curve is assumed to be smooth and is depicted in Fig. 1. The area beneath the rank-citation curve consists of three parts: the *h*-area (whose size is n^2), the *e*-area (A_e), and the *h*-tail area (A_t) (Ye & Rousseau, 2010). Alternatively, the three parts are referred to as the h^2 center, h^2 upper, and h^2 lower, respectively (Bornmann et al., 2010). We further refer to the area combining the *e*- and *h*-areas as the *h*-core area (A_c).

The geometric representation of Fig. 1 has been adopted by Hirsch (2005), van Eck and Waltman (2008), Zhang (2009), Alonso, Cabrerizo, Herrera-Viedma, and Herrera (2009), Bornmann and Daniel (2009), Bornmann et al. (2010) and Ye and

[1] For simplicity, the term *assignees* is expanded to cover individuals of various levels such as regions, countries, enterprises, corporations, institutions, inventors, etc. whose patent portfolios are to be evaluated.

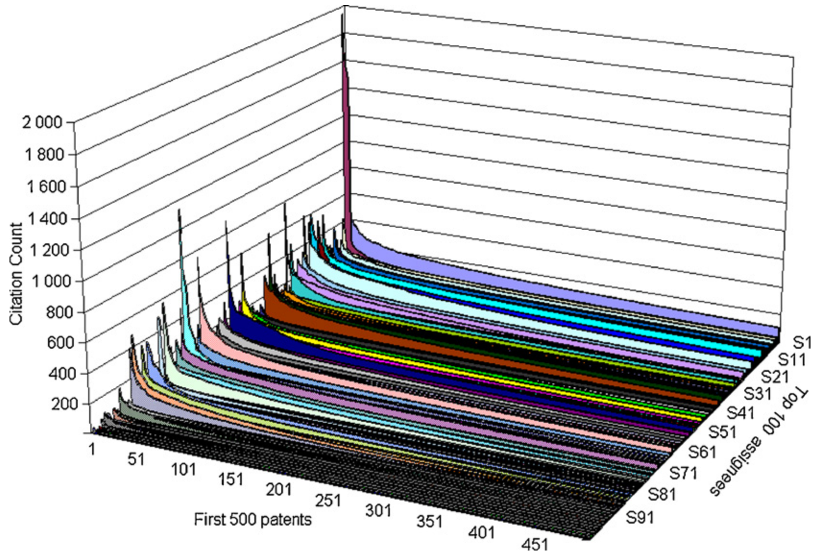

Fig. 2. Rank-citation curves for the 100 assignees.

Rousseau (2010) in providing geometric explanations to the *h*-index and various *h*-type indices. There is also a web site providing the single publication *h*-index as well as its associated rank-citation curve based on the data of Google Scholar (Thor & Bornmann, n.d.).

The empirical data utilized throughout the paper is based on the 100 assignees having the greatest numbers of U.S. patents granted in the year 2009 (U.S. Patent and Trademark Office, 2010). These assignees' U.S. patents issued between 1976 and 2009 are then collected, and the respective *h*-indices are found to range from 161 (IBM with total 58 185 patents) to 3 (LG Display Co., Ltd. with total 872 patents). From the diversity of these *h*-indices, the 100 assignees and their respective patent portfolios seem to constitute a representative set of data for our investigation.

The rank-citation curves for the 100 assignees are drawn in Fig. 2 where the assignees are arranged in descending order[2] of their *h*-indices along the axis towards the viewer with S_i meaning the *i*th assignee. For easier viewing, only the most highly cited 500 patents of these assignees are included.

As illustrated in Fig. 2, pretty much all assignees have significantly long *h*-tail areas while a handful of them have sharp *h*-core areas. We can see that some assignees with smaller *h*-indices have more significant *h*-core areas and/or *h*-tail areas than those with greater *h*-indices.

3. Assignee ranking by *h*-index

If the *h*-index is used as our general scale and an assignee S_i has a smaller *h*-index than that of another assignees S_j, what can we infer about their relative productivity and impact? We first use the patent count (N) and the total citation count (A) as measures for productivity and impact. Then, in Fig. 3, the 100 assignees are sorted in descending order of their *h*-indices along the horizontal axis, and their respective A's and N's are plotted against logarithmic left and right axes, respectively.

Even though, statistically, the ranking by *h*-index is highly correlated to the rankings by N and by A with respective Spearman's rhos at 0.774 and 0.925 (significant at 0.1% level), and both the A and N curves indeed reveal a general downward trend somewhat conforming to the decreasing *h*-indices (the N curve's trend is less obvious), the unpredictable and significant fluctuations along the curves suggest that we cannot confidently claim S_i must receive fewer patents and citations than S_j does.

We then use the cited patent count (N_c) as the measure for productivity in Fig. 4. Now both curves reveal similar downward trends conforming to the decreasing *h*-indices while equally unpredictable and significant fluctuations remain. Again we cannot confidently claim that S_i must receive fewer cited patents than S_j does.

We further try to see whether the total citation count (A) or the citation count received by the *h*-core patents (A_c) is better reflected in the ranking by *h*-index as the measure for impact. As illustrated in Fig. 5, both curves preserve similar downward trends while the A_c curve has fewer fluctuations of significantly reduced magnitudes. In other words, the *h*-core citation counts are better captured by the *h*-indices and we are able to claim with better confidence that S_i's *h*-core patents receive fewer citations than S_j's *h*-core patents do. Yet, the still present fluctuations make such inference unreliable.

[2] Throughout this paper, assignees of the same *h*-index are sorted in descending order of their total citation counts.

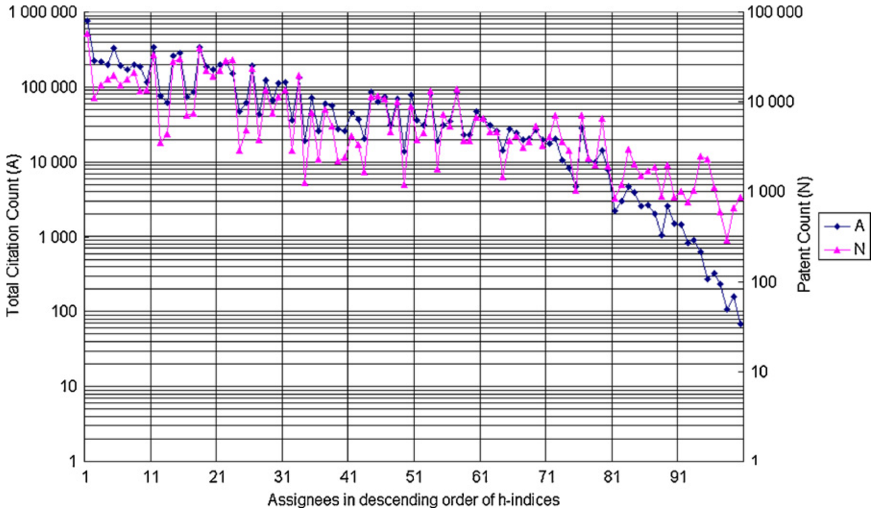

Fig. 3. Total citation counts and patent counts for the 100 assignees.

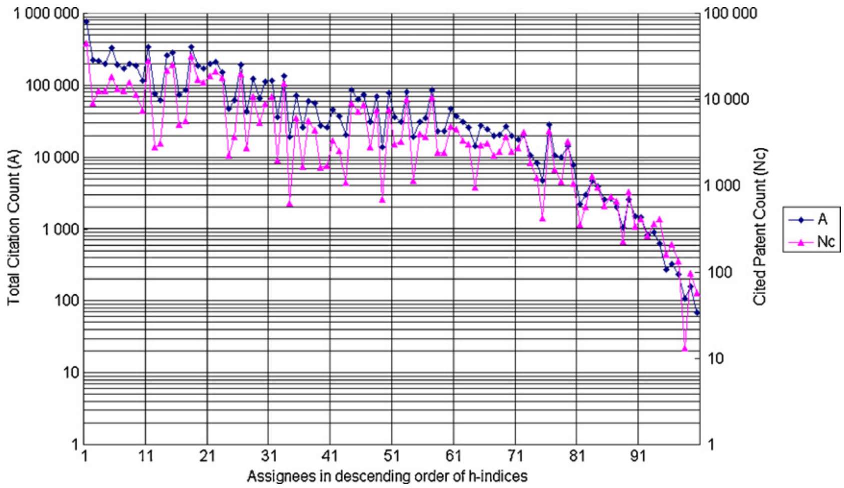

Fig. 4. Total citation counts and cited patent counts for the 100 assignees.

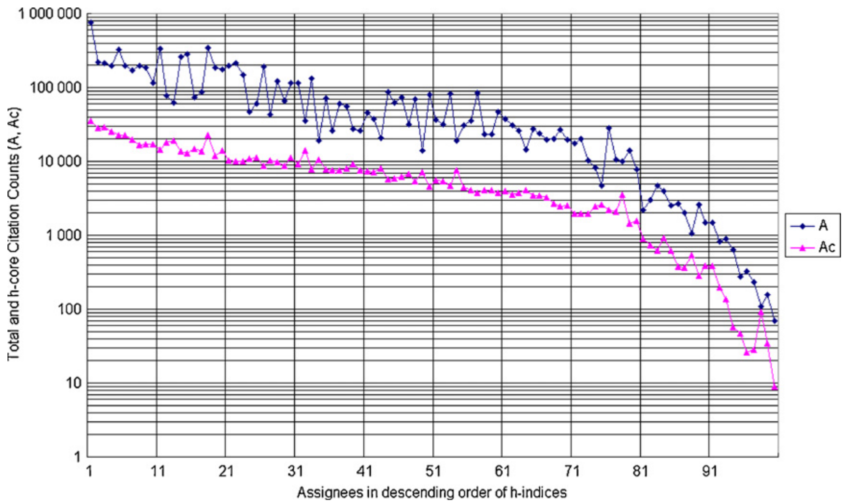

Fig. 5. Total citation counts and *h*-core citation counts for the 100 assignees.

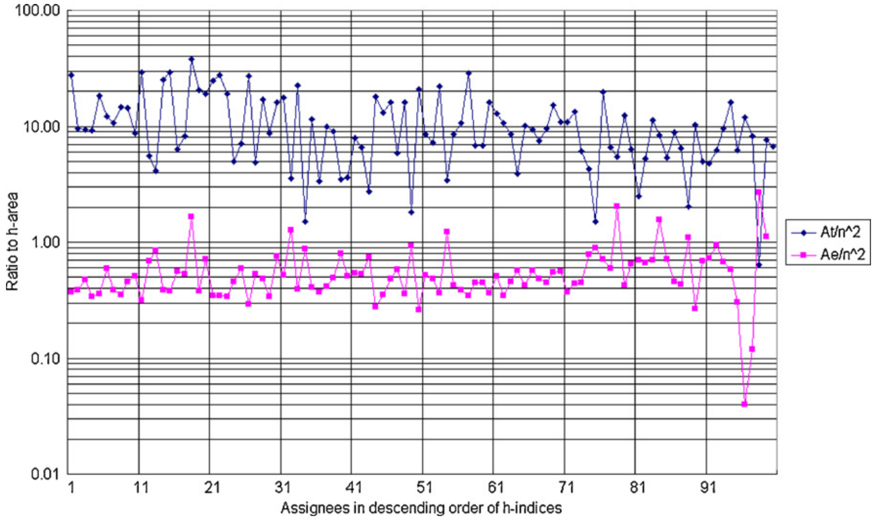

Fig. 6. The ratios A_t/n^2 and A_e/n^2 for the 100 assignees.

As the general trends observed in Figs. 4 and 5 have suggested, the *h*-index indeed has some ranking capability for assignees of distant enough *h*-indices. However, we need to incorporate the missing information contributing to the unpredictable fluctuations so as to develop the *h*-index into our general scale.

We could see from Fig. 5 that the smaller fluctuations along the A_c curve and the more severe fluctuations along the *A* curve must be resulted respectively from the more limited *e*-areas and much greater *h*-tail areas of these assignees. In contrast to the literature that most of the *h*-type indices leave the *h*-tail area barely touched, what is shown in Fig. 6 suggests that the focus has to be reconsidered for patent assignees.

In Fig. 6, the ratios of the assignees' *e*-areas and *h*-tail areas to their respective *h*-areas are plotted against logarithmic left axis. As illustrated, most assignees, or 92 out of the 100 assignees to be exact, have the ratios A_e/n^2 below 1.0, implying that most assignees' *e*-areas are not greater than their respective *h*-areas (i.e., $A_e \leq n^2$, or $A_c \leq 2n^2$). For the 100 assignees, the ratios A_e/n^2 have a mean 0.58 with a standard deviation 0.37.

On the other hand, the ratios A_t/n^2 varies over a great range, with a maximum as large as 38.1 and a minimum as small as 0.6. The mean ratio is 11.2 with a standard deviation 7, indicating that on the average most assignees have *h*-tail areas at least an order of magnitude greater than their respective *h*-areas.

Therefore, to achieve our general scale, an assignee's *h*-tail area constitutes such a huge portion of its productivity and impact to be ignored. As to the assignee's *e*-area, even though most of the times it is bounded, the still present fluctuations observed in Fig. 5 suggest that it should not be ignored either.

Instead of integrating some discordant *h*-type indices dealing separately with the *e*-area and the *h*-tail area to achieve our general scale, we believe that a preferable approach is to treat the *e*-area and the *h*-tail area in a unified manner. We therefore decide to return to the basics and re-examine the rank-citation curve.

4. Geometric interpretation to ranking by *h*-index

The rank-citation curve manifests the complete distribution of citations of an assignee's portfolio from which various information about the assignee's productivity and impact can be derived. For example, the total citation count is the area beneath the rank-citation curve and the cited patent count is where the rank-citation curve intersects the horizontal axis.

The *h*-index *n* corresponds to a point (*n*, *n*) that, for most assignees, is where the rank-citation curve is closest to the origin. The *h*-index therefore can be considered a *descriptor* characterizing the rank-citation curve. The point (*n*, *n*) also partitions the rank-citation curve into two segments manifesting the citation distributions of the *h*-core and *h*-tail patents, respectively. We refer to these segments as the *h*-core and *h*-tail segments.

The ranking by assignees' *h*-indices therefore reflects the positions of their rank-citation curves relative to the origin. Furthermore, if the *h*-index of an assignee S_i is sufficiently smaller than that of another assignee S_j, it is very possible that S_i's rank-citation curve runs completely under that of S_j. We refer to this scenario as S_i's rank-citation curve being *dominated* by S_j's curve, and we could claim that S_i is outperformed by S_j as S_i's *k* th patent always receives a smaller or equal number of citations to S_j's *k* th patent for all valid *k*'s, and S_i's total citation count is less than that of S_j.

To verify our proposition, we pick the assignees at the 11th (having *h*-index 105), . . ., 91 st (having *h*-index 15) places where the differences of their *h*-indices are at least 8 and at most 18. These assignees' relevant data are summarized in Table 1. Then, their rank-citation curves over the most highly cited 105 patents (so as to cover all *h*-core segments) and over the 106th to 1000th patents (as their *h*-tails are too long to include entirely) are plotted in Figs. 7 and 8, respectively. In Fig. 7, the points corresponding to their *h*-indices are connected by a dashed line.

Table 1
Relevant data for assignees at 11th, . . ., 91st places.

Assignee(rank by *h*-index)	N_c	A	A_c	A_t	*h*-Index, *n*
HITACHI(11)	27 756	338 698	14 498	324 200	105
NEC CORPORATION(21)	18 927	197 422	10 194	187 228	87
SHARP(31)	10 778	116 306	9266	107 040	78
NATIONAL SEMICON.(41)	3369	45 272	7327	37 945	69
NORTEL NETWORKS(51)	3030	36 414	5483	30 931	60
SANYO(61)	4518	37 487	3940	33 547	51
DELPHI TECHNOLOGIES(71)	2737	17 615	1979	15 636	38
NVIDIA(81)	354	2211	899	1312	23
FUNAI ELECTRIC(91)	409	1461	388	1073	15

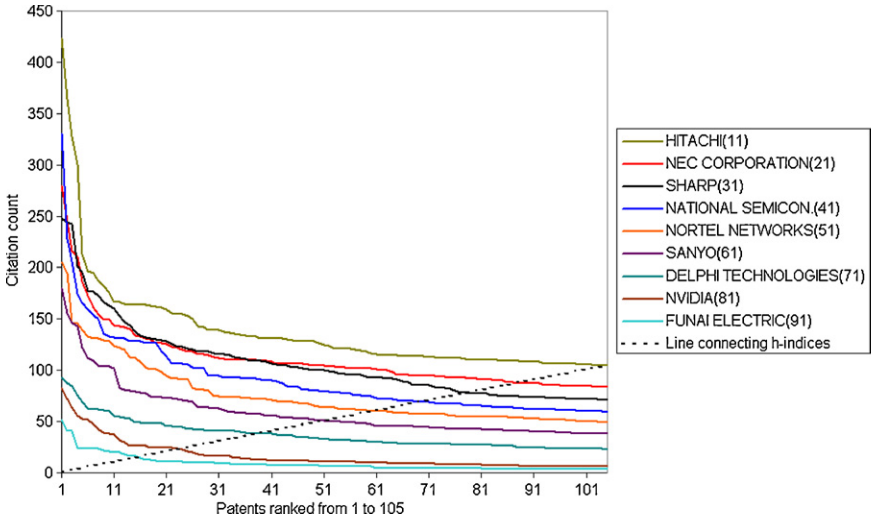

Fig. 7. *h*-Core segments for assignees at 11th, . . ., 91st places.

As illustrated in Figs. 7 and 8, the rank-citation curves for the 9 assignees are sequentially located relative to the origin in consistence with the ordering of their *h*-indices as expected.

The domination relationship among these assignees are indeed present after examining their portfolios, except for the 21st, 31st, and 41st assignees (NEC CORPORATION, SHARP, and NATIONAL SEMICON.) whose *h*-core segments cross each other, and for the 51st, 61st assignees (NORTEL NETWORKS, SANYO) and the 81st, 91st assignees (NVIDIA, FUNAI ELECTRIC) whose *h*-tail segments cross each. Actually, NATIONAL SEMICON. and SANYO also have crossed *h*-tail segments but the crossing happens at their 2183rd patents which are too far to the right to be included in Fig. 8.

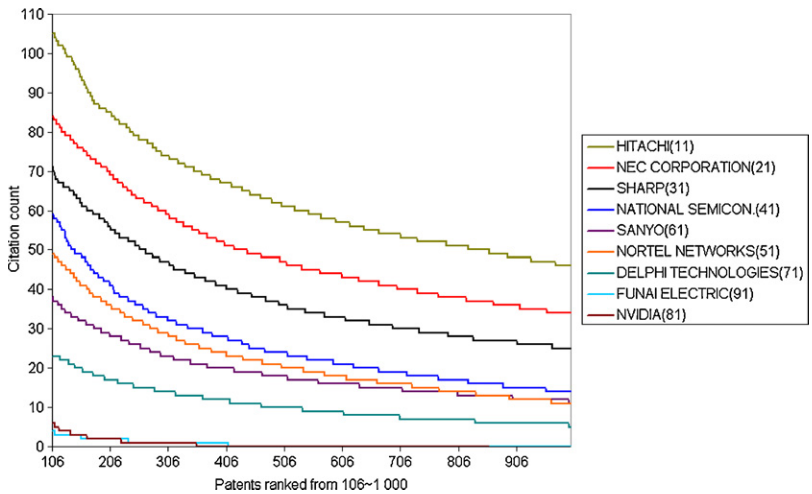

Fig. 8. *h*-Tail segments for assignees at 11th, . . ., 91st places.

Table 2
Relevant data for assignees from 21st to 29th places.

Assignee(rank by h-index)	N_c	A	A_c	A_t	h-Index, n
NEC CORPORATION(21)	18 927	197 422	10 194	187 228	87
SONY(22)	21 333	212 711	9924	202 750	86
SAMSUNG(23)	17 822	149 659	9961	139 735	86
APPLE(24)	2220	47 442	10 771	36 671	86
APPLIED MATERIAL(25)	3732	60 844	11 279	49 565	84
MITSUBISHI(26)	19 894	191 800	8680	183 120	82
TOKYO ELECTRIC(27)	2762	42 913	10 276	32 637	82
NOKIA(28)	10 801	121 515	8776	111 761	81
DUPONT(29)	5341	65 817	9754	57 041	81

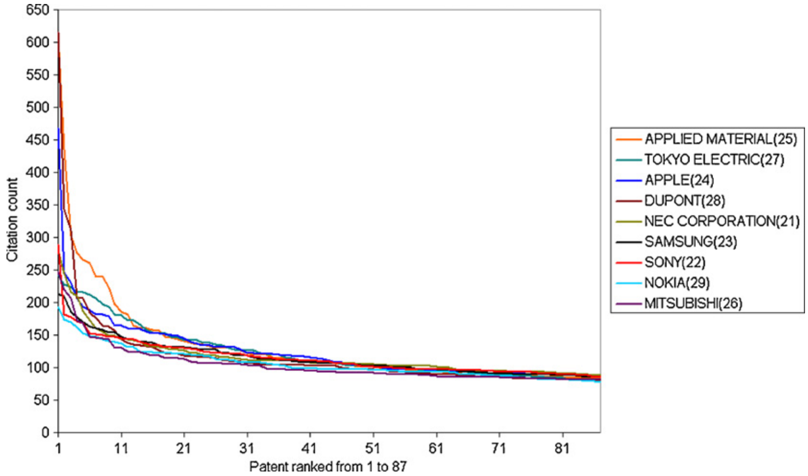

Fig. 9. h-Core segments for assignees from 21st to 29th places.

If distant h-indices suggest dominating rank-citation curves, close or identical h-indices should imply tangled rank-citation curves. To investigate the validity of this proposition, we pick 9 assignees ranked from the 21st place (having h-index 87) to the 29th place (having h-index 81) so that their h-indices are close to each other within a middle range (between 80 and 90) of the 100 assignees. Among them, the same 21st assignee (NEC CORPORATION) from the previous observation is also included as a common reference. The relevant data of these assignees are summarized in Table 2. Again, to facilitate observation, their rank-citation curves over the most highly cited 87 patents and over the 88th to 1000th patents are plotted in Figs. 9 and 10, respectively.

As illustrated in Fig. 9, even though some assignees have some exceptional highly cited patents such as the 25th and 28th assignees (APPLIED MATERIAL and DUPON), all h-core segments entwine as expected. However, only some pairs of

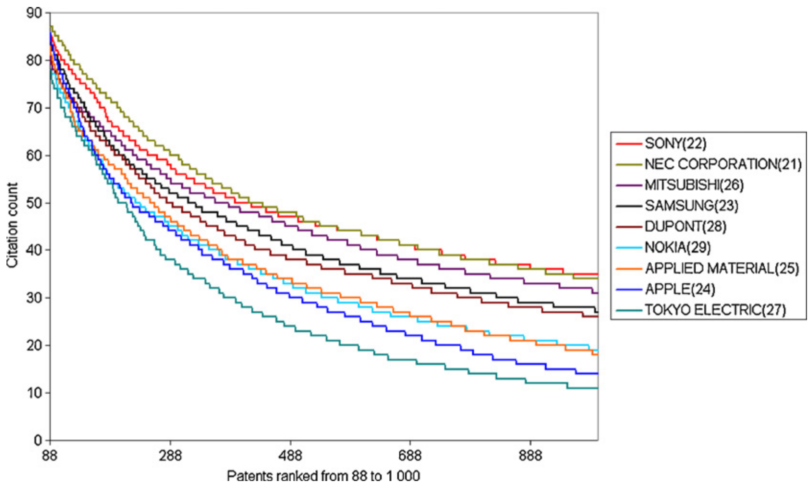

Fig. 10. h-Tail segments for assignees from 21st to 29th places.

Table 3
Descriptors for assignees at 11th, . . ., 91st places.

Assignee(rank by *h*-index)	*h*-Index, *n*	*c*-desc., *c*	*t*-desc., *t*	Order by *n*	Order by *c*	Order by *t*
HITACHI(11)	105	156	6573	1	1	1
NEC CORPORATION(21)	87	127	4451	2	3	2
SHARP(31)	78	130	2626	3	2	3
NATIONAL SEMICON.(41)	69	123	856	4	4	5
NORTEL NETWORKS(51)	60	103	758	5	5	6
SANYO(61)	51	88	1128	6	6	4
DELPHI TECHNOLOGIES(71)	38	56	714	7	7	7
NVIDIA(81)	23	46	105	8	8	9
FUNAI ELECTRIC(91)	15	30	128	9	9	8

h-tail segments entwine with each other in Fig. 10 and some distinct groups of *h*-tail segments are clearly observable. For example, the *h*-tail segments of the 21st and 22nd assignees (NEC CORPORATION and SONY) constitute a group that is spaced apart from the other *h*-tail segments. In other words, assignees with close or identical *h*-indices can actually be further differentiated.

Our geometric interpretation to the ranking of assignees by their *h*-indices so far is proven by empirical data to be rather reliable and accurate, in contrast to the numerous fluctuations observed in Section 3 when the ranking by *h*-index is related to conventional performance measures.

However, as suggested by the exceptions observed, we cannot skip the assignees with distance *h*-indices when conducting patentometric analysis by simply assuming that there is domination relationship among their rank-citation curves and that their relative performance is well determined. Similarly, we also shouldn't give up the further differentiation to the assignees with close or identical *h*-indices by simply assuming that their rank-citation curves are interleaved and that the compatibility of their relative performance is well determined.

5. Descriptors for characterizing *h*-core and *h*-tail segments

To achieve our general scale, we need to further investigate the geometric relationship among the *h*-core and *h*-tail segments of the assignees. Therefore, two descriptors are proposed as supplements to the *h*-index, one characterizing the *h*-core segment and the other one characterizing the *h*-tail segment of the rank-citation curve, and are referred to as the *c*- and *t*-descriptors, respectively. These descriptors are obtained as follows:

$$c\text{-descriptor} = \sum_{i=1}^{n} C(P_i)\left(\frac{C(P_i)}{A_c}\right) = \frac{\sum_{i=1}^{n} C(P_i)^2}{\sum_{i=1}^{n} C(P_i)}, \tag{1}$$

$$t\text{-descriptor} = \sum_{i=n+1}^{N_c} i\left(\frac{C(P_i)}{A_t}\right) = \frac{\sum_{i=n+1}^{N_c} iC(P_i)}{\sum_{i=n+1}^{N_c} C(P_i)}. \tag{2}$$

According to Eqs. (1) and (2), the *c*- and *t*-descriptors are weighted averages of the heights ($C(P_i)$) and the horizontal distances (i) of the points on the *h*-core and the *h*-tails segments, respectively. Considering two assignees having *h*-core areas of identical size, Eq. (1) would achieve a greater *c*-descriptor for the assignee whose *h*-core segment is more skewed to the left. Similarly, for two assignees having equally sized *h*-tail areas, Eq. (2) would achieve a greater *t*-descriptor for the assignee whose *h*-tail segment slops gently to the farther right.

The *c*- and *t*-descriptors can be easily obtained as by-products to the determination of an assignee's *h*-index by iterating through its decreasingly sorted patent portfolio, thereby preserving the *h*-index's simplicity advantage in our approach.

For the assignees of Table 1, their *c*- and *t*-descriptors are summarized in Table 3. The assignees' *h*-indices are also repeated to facilitate the comparison and subsequent discussion.

For two assignees S_i and S_j with *h*-indices h_i and h_j, *c*-descriptors c_i and c_j, and *t*-descriptors t_i and t_j, we would expect that, if h_i is sufficiently greater than h_j, S_i's *h*-core segment should run higher above, and its *h*-tail segment should extend farther to the right. However, if an opposite scenario is indicated by $c_i < c_j$ or $t_i < t_j$, their *h*-core or the *h*-tail segments should cross each other.

To see how *c*- and *t*-descriptors are applied to assignees with distant *h*-indices, the 10 assignees of Table 3 are first sorted decreasingly according to their *h*-indices, and the resulted order is listed in the column "Order by *n*." The 10 assignees are

Table 4
Descriptors for assignees from 21st to 29th places.

Assignee(rank by *h*-index)	*h*-Index, *n*	*c*-desc., *c*	*t*-desc., *t*	Grouping by *t*
NEC CORPORATION(21)	87	127	4451	Group 1
SONY(22)	86	123	4972	Group 1
SAMSUNG(23)	86	122	3877	Group 1
APPLE(24)	86	146	595	Group 2
APPLIED MATERIAL(25)	84	180	884	Group 2
MITSUBISHI(26)	82	115	4829	Group 1
TOKYO ELECTRIC(27)	82	138	639	Group 2
DUPONT(28)	81	162	2595	Group 1
NOKIA(29)	81	113	1189	Group 2

then decreasingly sorted according to their *c*- and *t*-descriptors and the orders obtained are listed in the columns "Order by *c*" and "Order by *t*," respectively.

We can see that the relative positions of the 21st and 31st assignees (NEC CORPORATION and SHARP) are reversed within the two ordered lists by *h*-indices and *c*-descriptors. From Fig. 7, it is verified that NEC CORPORATION and SHARP indeed have crossed *h*-core segments.

Similarly, the reversed positions of the 51st and 61st assignees (NORTEL NETWORKS and SANYO), and the 81st and 91st assignees (NVIDIA and FUNAI ELECTRIC) within the two ordered lists by *h*-indices and *t*-descriptors accurately predict the crossing of their *h*-tail segments, as verified from Fig. 8.

As indicated in Section 4, the *h*-tail segments of NATIONAL SEMICON. (the 41st assignee) and SANYO cross each other at their 2183rd patents. This is also successfully reflected by their revered positions within the two ordered lists by *h*-indices and *t*-descriptors. Please note that our approach looks at the relative positions, not the absolute positions, of assignees. Therefore, for example, even though NATIONAL SEMICON. and NORTEL NETWORKS have different absolute positions in the ordered lists, their relative positions (i.e., $4 < 5$ in the ordered lists by *h*-indices and *c*-descriptors, and $5 < 6$ in the ordered list by *t*-descriptors) are preserved.

As described, our *c*- and *t*-descriptors accurately predict the crossing of the *h*-core and *h*-tail segments for assignees with distant *h*-indices. For these assignees, there is no domination relationship among them and, when conducting patentometric analysis, they are the ones requiring further detailed analysis so as to clarify their relative performance.

In Fig. 7, the *h*-core segment of NATIONAL SEMICON. crosses those of NEC CORPORATION and SHARP, but this scenario is not reflected in Table 3. This is not considered a failure but actually demonstrates the robustness of our approach to misleading scenarios. After examining their portfolios, NATIONAL SEMICON. has only one exceptionally highly cited patent causing its *h*-core segment to cross those of NEC CORPORATION and SHARP. The rest of its *h*-core patents all receive fewer citations than those patents of NEC CORPORATION and SHARP at corresponding ranking positions. Therefore, the *h*-core segment of NATIONAL SEMICON. is substantially dominated by those of NEC CORPORATION and SHARP, as suggested by their *c*-descriptors, and NATIONAL SEMICON. should still be considered outperformed by NEC CORPORATION and SHARP.

To see how *c*- and *t*-descriptors are applied to assignees with close or identical *h*-indices, the *c*- and *t*-descriptors for the assignees of Table 2 are summarized in Table 4.

As we learn from Fig. 9, the *h*-core segments of these assignees are interleaved yet some assignees have exceptionally highly cited patents, which is successfully reflected by their dramatically great *c*-descriptors. As such, when conducting patentometric analysis, we should apply further analysis to these assignees with close or identical *h*-indices yet with significant *c*-descriptor differences so as to clarify their relative performance with respect to their *h*-core patents. On the other hand, for assignees with close or identical *h*-indices and *c*-descriptors, these assignees should have comparable performance at least with respect to their *h*-core patents, and detailed analysis to these assignees' *h*-core patents therefore could be skipped.

From Fig. 10, we also learn that these assignees could be further differentiated with respect to their *h*-tail patents, which is also successfully reflected by their dramatically different *t*-descriptors. In addition, as shown in Table 4's "Grouping by *t*" column, these assignees could actually be separated into two groups based on their *t*-descriptors: those having *t*-descriptors over 2000 and those having *t*-descriptors below 2000. The two groups of assignees are reflected in Fig. 10 as an upper set of *h*-tail segments clearly runs over a lower set of the other *h*-tail segments. These assignees could be further differentiated into more groups. For example, NEC CORPORATION, SONY, and MITSUBISHI (the 26th assignee), all having *t*-descriptors over 4000, are clearly separated from SAMSUNG and DUPONT, both having *t*-descriptors between 2000 and 4000. When conducting patentometric analysis, we could limit detailed analysis to the assignees with close or identical *h*-indices but within the same group (i.e., with close or identical *t*-descriptors) at least with respect to their *h*-tail patents. As to assignees belonging to separate groups, their relative performance at least with respect to the *h*-tail patents should be well determined.

The utilizations of the *c*- and *t*-descriptors to assignees of close of identical *h*-indices are different. This is because the ranking by their *h*-indices successfully predicts the compatibility of their relative performance with respect to their *h*-core patents. Therefore, we only need to focus on the exceptions suggested by dramatically different *c*-descriptors. In contrast, the ranking by their *h*-indices poorly reflects these assignees' relative performance with respect to their *h*-tail patents. Under this circumstance, their *t*-descriptors become more evidential, and dramatically different *t*-descriptors suggest the superiority or inferiority of their relative performance with respect to their *h*-tail patents. We therefore focus on those with close or identical *t*-descriptors.

6. Conclusion

Based on our geometric interpretation to the ranking of assignees by their *h*-indices, the proposed descriptors successfully verify and clarify the geometric relationship among assignees with distant or close *h*-indices. According to our empirical data where the *h*-indices range from 3 to 161, it seems appropriate that, if the difference between two *h*-indices is at least 8, they could be considered distant enough.

In addition to the reliability, accuracy, and robustness observed in Sections 4 and 5, our approach also provides flexibility as the two descriptors could be applied selectively or together. We therefore can evaluate and compare assignees' relative performance only with respect to their *h*-core patents, or only with respect to their *h*-tail patents. This flexibility also allows us to gain more insight into assignees' relative performance and thereby to be more discriminating. For example, our approach is able to tell that, while the relative performance of NEC CORPORATION and SHARP with respect to their *h*-core patents has to be further clarified, NEC CORPORATION actually outperforms SHARP with respect to their *h*-tail patents.

As the *c*-descriptor successfully characterizes the *h*-core segment and therefore the *h*-core patents, it can be individually applied to research performance evaluation for compensating the *h*-index's deficiency to the exceptionally highly cited papers. The application of our approach is therefore not limited to patent assignees only.

References

Alonso, S., Cabrerizo, F. J., Herrera-Viedma, E., & Herrera, F. (2009). *h*-Index: A review focused in its variants, computation and standardization for different scientific fields. *Journal of Informetrics*, *3*, 273–289.

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.

Bornmann, L., & Daniel, H.-D. (2009). The state of *h* index research, is the *h* index the ideal way to measure research performance? *EMBO Reports*, *10*, 2–6.

Bornmann, L., Mutz, R., & Daniel, H.-D. (2008). Are there better indices for evaluation purposes than the h index? A comparison of nine different variants of the *h* index using data from biomedicine. *Journal of the American Society for Information Science and Technology*, *59*, 830–837.

Bornmann, L., Mutz, R., & Daniel, H.-D. (2010). The *h* index research output measurement: Two approaches to enhance its accuracy. *Journal of Informetrics*, *4*, 407–414.

Cabrerizo, F. J., Alonso, S., Herrera-Viedma, E., & Herrera, F. (2010). q^2-Index: Quantitative and qualitative evaluation based on the number and impact of papers in the Hirsch core. *Journal of Informetrics*, *4*, 23–28.

Costas, R., & Bordons, M. (2007). The *h*-index: Advantages, limitations and its relation with other bibliometric indicators at the micro level. *Journal of Informetrics*, *1*, 193–203.

Egghe, L. (2006a). An improvement of the *h*-index: The *g*-index. *ISSI Newsletter*, *2*, 8–9.

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

Egghe, L. (2010). The Hirsch-index and related impact measures. *Annual Review of Information Science and Technology*, *44*, 65–114.

Egghe, L., & Rousseau, R. (2008). An *h*-index weighted by citation impact. *Information Processing and Management*, *44*, 770–780.

García-Pérez, M. A. (2009). A multidimensional extension to Hirschs *h*-index. *Scientometrics*, *81*, 779–785.

Guan, J. C., & Gao, X. (2009). Exploring the *h*-index at patent level. *Journal of the American Society for Information Science and Technology*, *60*, 35–40.

Hall, B. H., Jaffe, A. B., & Trajtenberg, M. (2000). *Market value and patent citations: A first look*. Working paper 7741. National Bureau of Economic Research. Available online at http://www.nber.org/papers/w7741.pdf.

Hall, B. H., Jaffe, A. B., & Trajtenberg, M. (2001). *The NBER patent citations data file: Lessons, insights and methodological tools*. Working paper 8498. National Bureau of Economic Research. Available online at http://www.nber.org/papers/w8498.pdf.

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of United States of America*, *102*, 16569–16572.

Jin, B. (2007). The *AR*-index: Complementing the *h*-index. *ISSI Newsletter*, *3*, 6.

Jin, B., Liang, L., Rousseau, R., & Egghe, L. (2007). The *R*- and *AR*-indices: Complementing the h-index. *Chinese Science Bulletin*, *52*, 855–863.

Kosmulski, M. (2006). A new Hirsch-type index saves time and works equally well as the original *h*-index. *ISSI Newsletter*, *2*, 4–6.

Meyer, M. (2000). What is special about patent citations? Differences between scientific and patent citations. *Scientometrics*, *49*, 93–123.

Meyer, M., & Bhattacharya, S. (2004). Commonalities and differences between scholarly and technical collaboration. *Scientometrics*, *61*, 443–456.

Rousseau, R. (2006). New developments related to the Hirsch index. *Science Focus, 1*, 23–25 (in Chinese). An English translation is available online at http://eprints.rclis.org/6376/.

Rousseau, R. (2008). Reflections on recent developments of the *h*-index and *h*-type indices. *COLLNET Journal of Scientometrics and Information Management*, *2*, 1–8.

Thor, A., & Bornmann, L. (n.d.). Web application to calculate the single publication *h* index (and further metrics) based on Google Scholar. In *Database Group Leipzig*. Retrieved from http://labs.dbs.uni-leipzig.de/gsh/.

U.S. Patent and Trademark Office. (2010). *Patenting by Organizations 2009*. A Patent Technology Monitoring Team Report U.S. Patent and Trademark Office. Available online at http://www.uspto.gov/web/offices/ac/ido/oeip/taf/topo_09.htm.

van Eck, N. J., & Waltman, L. (2008). Generalizing the *h*- and *g*-indices. *Journal of Informetrics*, *2*, 263–271.

Wu, Q. (2010). The *w*-index: A significant improvement of the *h*-index. *Journal of the American Society for Information Science and Technology*, *61*, 609–614.

Ye, F. Y., & Rousseau, R. (2010). Probing the *h*-core: An investigation of the tail-core ratio for rank distributions. *Scientometrics*, *84*, 431–439.

Zhang, C.-T. (2009). The *e*-index, complementing the *h*-index for excess Citations. *PLoS ONE*, *4*, e5429.