# Capturing and Tracking Performance of Patent Portfolio Using $h$-Complement Area Centroid

Chung-Huei Kuan, Mu-Hsuan Huang, and Dar-Zen Chen

***Abstract*—Assuming that only the bibliometric information contained in patent databases is available and that the citations received by a patent indicate the patent's value or quality, we propose to characterize the performance of a patent assignee's patent portfolio, in terms of both the quantity and quality of its cited and uncited patents, by the centroid of a so-called $h$-complement area from the assignee's citation distribution. This approach is capable of producing a snapshot to a large number of assignees' portfolio performance by simultaneously depicting their $h$-complement area centroids in a 2-D graph. With this 2-D view, these assignees' relative performance, where the performance difference lies, and the degree of such difference can also be quickly determined. In addition, a trajectory manifesting the evolution of an assignee's portfolio performance over a window of time can be obtained by connecting the $h$-complement area centroids at successive epochs within the time window. The pattern revealed by the trajectory provides significant insight into how the assignee's portfolio performance has varied over time. When there is some abrupt pattern change, an analyst should be alarmed to conduct further investigation. A steady pattern, on the other hand, allows forecast to the assignee's future performance.**

***Index Terms*—Centroid, $h$-complement area, $h$-index, portfolio performance, rank-citation curve.**

## I. Introduction

PATENT data, which are structurally organized and substantially objective, are well recognized as a viable source of information for various technology management tasks such as policy making [30], tracing knowledge diffusion [1], [8], [10], strategic planning [3], [16], [29], technology analysis [33], [34], technological forecasting [4], [15], [29], finding relationship among companies and industries [25], and providing assessment to targets of merger and acquisition [7]. There is already a significant amount of related research involving patent data such as [9], [20], [21], [26], and [39].

For an organization's technology manager, competitor analysis often requires the assessment of the organization's technological position among its competitors in a market segment, about a product, or within a technology sector. To this purpose, patents assigned to these competitors are considered as an important source of technological intelligence. Even though some competitors may choose not to file patents for saving cost or to conceal their breakthrough technology, or the value of a patent is not yet manifested as it has not been involved in licensing deals or legal disputes, the freely available patent databases provide the technology manager an economical option in contrast to information such as the R&D spending where costly in-depth research is required. What this paper concerns is that, if given only the bibliometric information contained in free patent databases, can a technology manager conduct an effective competitor analysis work?

Even though we limit ourselves to the patent bibliometric information, there are various ways of applying this information. For example, if a technology is of importance to the competitor, patent applications to multiple countries are filed and a sizeable patent family is established, or continuous development is conducted and manifested in the form of numerous continuing applications. Here, in this paper, we further limit ourselves to use the citations received by a patent as an indication to the patent's value or quality, not only because their ready availability from patent databases but also because they are relatively more objective.[1] And, for simplicity, we do not consider a patent's remaining life span before its expiration.[2]

Under the foregoing assumptions, the $h$-index [24] provides a hint. The $h$-index has been a de facto scientometric indicator for research performance evaluation, as evident from the significant number of related research papers (e.g., the original Hirsch paper is cited more than 2000 times since its introduction in 2005) and its adoption by online databases such as Scopus and Web of Science. To characterize the scientific output or research performance of a researcher, the researcher's $h$-index is $n$ if he or she has published at least $n$ papers, each receiving at least $n$ citations. The $h$-index, therefore, involves only two attributes: *quantity* (a measure related to the number of papers published) and *quality* (a measure related to the citations received by the

Manuscript received September 3, 2011; revised March 5, 2012 and July 26, 2012; accepted August 20, 2012. Date of publication October 5, 2012; date of current version July 13, 2013. Review of this manuscript was arranged by Department Editor S. (Sri) Talluri.

C.-H. Kuan is with the Graduate Institute of Patent, National Taiwan University of Science and Technology, Taipei 106, Taiwan (e-mail: maxkuan@mail.ntust.edu.tw).

M.-H. Huang is with the Department of Library and Information Science, National Taiwan University, Taipei 106, Taiwan (e-mail: mhhuang@ntu.edu.tw).

D.-Z. Chen is with the Department of Mechanical Engineering, National Taiwan University, Taipei 106, Taiwan (e-mail: dzchen@ntu.edu.tw).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TEM.2012.2216882

[1]We believe that the citations to patents are more objective than citations to papers. The reasons are that (1) citations to patents are produced not only by the applicants but also by the patent examiners who can be considered as objective third parties; and (2) self-citation is less a problem to patents than to papers because there is little benefit for an applicant to cite his or her prior applications or patents.

[2]A patent's value should be discounted if it has a shorter remaining life. However, to our best knowledge, no similar research (i.e., using patent citations for performance evaluation) has incorporated patent expiration into consideration. We think that a major reason is that it is very difficult to determine whether a patent is still under maintenance. For U.S. patents, the related data is hidden in USPTO's Public PAIR (Patent Application Information Retrieval) database and one has delved into the database to get the detail one by one for each patent.

published papers), and integrates the two in a single number [11], [24], [38]. This simplicity is why its validity has been questioned even until this date. Even though quite a number of similar measures have been proposed and claimed to be superior [2], [13], the way that the $h$-index integrates the two attributes is the more intuitive one.

The extension of the $h$-index's application to patent assignees seems intuitive due to the many analogous features between published papers and patents [31], [32], yet there are few articles [19], [27] dedicated to this extension. This is probably due to that the adequacy of applying the $h$-index to measure the quantity and quality of a patent assignee's portfolio (jointly referred to as the *portfolio performance* hereinafter) is dubious. The $h$-index, on one hand, has been mostly criticized by the scientific community for being insensitive to some exceptionally highly cited papers, and there are a large number of so-called $h$-type indices proposed to modify or augment the original $h$-index. Some recent reviews to these $h$-type indices can be found in [2] and [13]. On the other hand, the $h$-index is also recognized as being insensitive to the lowly cited and uncited papers [11], [38], even though this is not unanimously regarded as a disadvantage by the scientific community.

Actually a technology manager should concern more about the $h$-index's insensitivity to the lowly cited and uncited patents. As indicated by empirical studies [22], [23], [33], there are few highly cited patents and the great majority of patents are either lowly cited or uncited. Therefore, when applied to patent assignees, the $h$-index may fail to accurately reflect an assignee's portfolio performance as a significant portion of the assignee's effort embodied in its lowly cited and uncited patents is not accounted for.

Therefore, in this paper, we try to extend the $h$-index to the measurement of a patent assignee's performance in terms of its patent portfolios and propose some approaches in adapting the $h$-index for this measurement.

## II. $h$-Index Geometry

Let an assignee have been granted $N$ patents. An operational definition of the assignee's $h$-index is as follows. First, the $N$ patents are sorted in descending order of their respective citation counts into an ordered set $\{P_1, P_2, \ldots, P_{N-1}, P_N\}$, and $C(P_i) \geq C(P_j)$ if $1 \leq i \leq j \leq N$, where $P_i$ is the patent ranked at the $i$th place and $C(P_i) \geq 0$ is the citation count of $P_i$. Then, the assignee has $h$-index $n$ if $C(P_n) \geq n$ and $C(P_{n+1}) \leq n$.

The so-called *rank-citation curve* [40] is a curve manifesting the citation distribution of an assignee's decreasingly sorted patents and has been frequently adopted by $h$-index related researchers to graphically illustrate their various propositions [2], [5], [6], [12], [41]. An assignee's rank-citation curve is obtained by plotting and connecting the points $(i, C(P_i))$, $1 \leq i \leq N$, together in a smooth or stepwise manner in a 2-D coordinate system where the horizontal axis is the ranks of the patents and the vertical axis is the patents' citation counts. A fictitious, stepwise rank-citation curve is depicted in Fig. 1.

Based on the definition above, the $h$-index $n$ corresponds to a point $(n, n)$ on the rank-citation curve. Another operation definition of the $h$-index is, therefore, to draw the line $x = y$, and the line's intersection with the rank-citation curve determines the $h$-index.

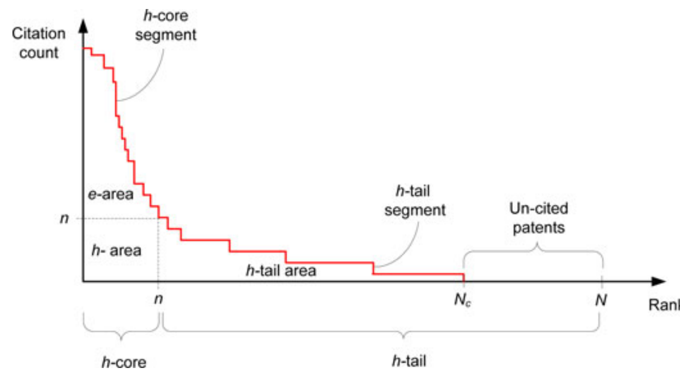


Fig. 1. Fictitious rank-citation curve.

The set of patents $\{P_1, P_2, \ldots, P_{N-1}, P_N\}$ is partitioned by the assignee's $h$-index $n$ into the set of $n$ patents $\{P_1, P_2, \ldots, P_{n-1}, P_n\}$ and the set of $(N-n)$ patents $\{P_{n+1}, P_{n+2}, \ldots, P_{N-1}, P_N\}$, which are referred to as the assignee's $h$-core [37] and $h$-tail [40], respectively. The $h$-tail can be further divided into the set of $(N_c-n)$ cited patents $\{P_{n+1}, P_{n+2}, \ldots, P_{Nc-1}, P_{Nc}\}$ and the set of $(N-N_c)$ uncited patents $\{P_{Nc+1}, P_{Nc+2}, \ldots, P_{N-1}, P_N\}$ where $N_c$ is the number of cited patents (i.e., patents having been cited at least once). Conventionally, the patents in the $h$-core are considered as highly cited ones and those in the $h$-tail are considered as lowly cited.

The point $(n, n)$ also partitions the rank-citation curve into two segments manifesting the citation distributions of the $h$-core and $h$-tail, respectively. We refer to these segments as the *$h$-core* and *$h$-tail segments*, respectively [27]. The areas beneath the $h$-core and $h$-tail segments corresponding to citations received by the $h$-core and $h$-tail are referred to as the *$h$-core area* (whose sizes is denoted as $A_c$) [27] and *$h$-tail area* (whose sizes is denoted as $A_t$) [40]. The $h$-core area is further divided into the $h$-area (whose size is $n^2$) and the $e$-area (whose size is denoted as $A_e = A_c - n^2$) [41]. The area size beneath the entire rank-citation curve is denoted as $A$ ($= A_c + A_t = A_e + n^2 + A_t$).

The uncited patents $\{P_{Nc+1}, P_{Nc+2}, \ldots, P_{N-1}, P_N\}$, as they have zero citation, are not reflected in the $h$-tail area and the $h$-tail segment. For a step-wise rank-citation curve, as illustrated in Fig. 1, the $h$-core and $h$-tail areas can be considered as consisting of $n$ and $(N_c - n)$ rectangles, respectively, where each rectangle has width 1 and height $C(P_i)$ (therefore, area size $C(P_i)$). These observations will be useful in the following sections.

The way the $h$-index characterizes the quantity and quality of an assignee's portfolio is using a single number $n$ to represent the total $N$ patents and the total number of citations $A$, or using a single point $(n, n)$ to represent the entire rank-citation curve. Clearly, the $h$-index only carries limited information content, and a number of researches as such tried to extract more information from the entire area underneath the rank-citation curve or from the entire rank-citation curve. García-Pérez [17] proposed a multidimensional vector quantizing the entire area

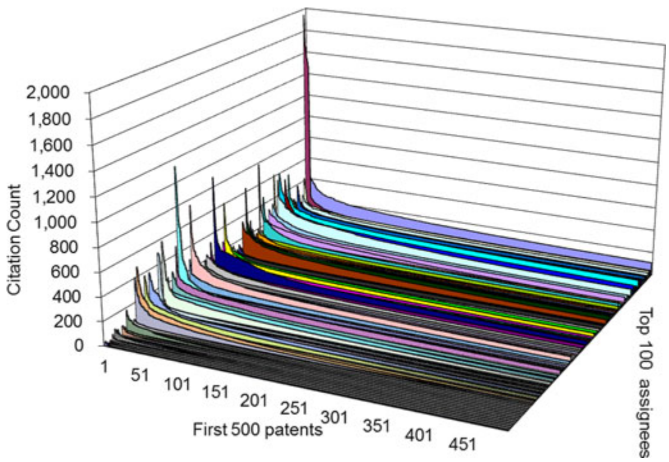


Fig. 2. Rank-citation curves of the top 100 assignees having the most granted U.S. patents in 2009.

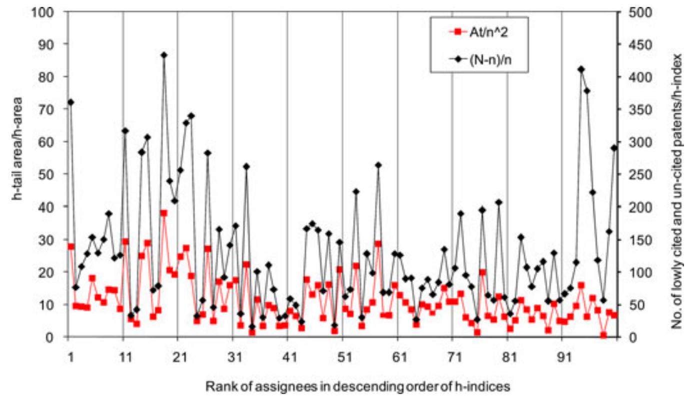


Fig. 3. Ratios $A_t/n^2$ and $(N-n)/n$ for the top 100 assignees having the most granted U.S. patents in 2009.

beneath the rank-citation curve. Bornmann *et al*. [6] proposed to quantify three areas beneath the rank-citation curve: the area of lowly cited patents ranked behind the $h$-index ($h^2$lower), the square area whose width is the $h$-index ($h^2$center), and the area of excessive citations above the $h^2$ center ($h^2$ upper). Ye and Rousseau [40] studied the ratio of the citations received by lowly cited $h$-tail papers to those received by the highly cited $h$-core papers as a so-called tail-core ratio. Egghe [14] adopted a concept referred to as characteristic scores and scales [18] and proposed to summarize the research performance of a researcher by a number of values, similar to García-Pérez's multidimensional vector.

These approaches on one hand rely mainly on the area sizes which are notorious for hiding details. On the other hand, the uncited patents, as having zero area size, are not taken into consideration. In the following sections, we try to incorporate the uncited patents and propose more discriminating approaches.

## III. Research Data

To test the paper's proposition in characterizing an assignee's patent portfolio, the empirical data utilized in this paper are based on the 100 assignees having the greatest numbers of U.S. patents granted in the year 2009 [35]. These assignees' U.S. patents issued between 1976 and 2009 are collected, and the respective $h$-indices are found to range from 161 (IBM with total 58185 patents) to 3 (LG DISPLAY CO., LTD. with total 872 patents). The 100 assignees and their respective patent portfolios constitute a representative set of data for our investigation as suggested by their diverse $h$-indices.

The rank-citation curves of the 100 assignees are drawn in Fig. 2 where the assignees are arranged in descending order of their $h$-indices along the axis toward the viewer. For assignees of the same $h$-index, they are further sorted by their respective total citation counts.

For easier viewing, only the most highly cited 500 patents of each assignee are included. As illustrated in Fig. 2, pretty much all assignees have significantly long $h$-tails whereas a handful of them have sharp $h$-cores. We can also see that some assignees with smaller $h$-indices (i.e., those closer to the viewer) have more significant $h$-cores and/or $h$-tails than those with greater $h$-indices (i.e., those farther away from the viewer).

To confirm our speculation that the $h$-index may fail to accurately reflect an assignee's lowly cited and uncited patents, the top 100 assignees' ratios of the $h$-tail area size $A_t$ to the $h$-area size $n^2$ and ratios of the number of lowly cited and uncited patents $(N-n)$ (i.e., the size of $h$-tail) to the $h$-index $n$ are plotted against the left and right axes, respectively, with the top 100 assignees arranged in descending order of their $h$-indices from left to right in Fig. 3. The division by $n^2$ and $n$ is for normalization.

If $A_t$ and $(N-n)$ are correlated with the $h$-index $n$ in any way, the curves for $A_t/n^2$ and $(N-n)/n$ should at least reflect some pattern. However, the significant and seemingly unpredictable fluctuations along both curves depicted in Fig. 3 suggest that the $h$-index indeed does not carry enough information about an assignee's lowly cited and uncited patents.

## IV. Utilizing Shape Descriptor for Portfolio Performance Characterization

The $h$-index $n$ corresponds to a point $(n, n)$ on the rank-citation curve that, for most assignees, is where the rank-citation curve is closest to the origin. The ranking by the assignees' $h$-indices, therefore, substantially reflects the positions of their rank-citation curves relative to the origin. For example, as shown in Fig. 4, an assignee $S_i$'s rank-citation curve runs completely under that of another assignee $S_j$. We refer to this scenario as $S_i$'s rank-citation curve being *dominated* by $S_j$'s curve, and we can 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. Note that, in this argument, we have not included the uncited patents of $S_i$ and $S_j$ into consideration.

In this scenario, $S_i$'s smaller $h$-index $n_i$ and $S_j$'s greater $h$-index $n_j$ successfully suggest that $S_i$'s rank-citation curve is dominated by that of $S_j$ and that $S_i$ is outperformed by $S_j$.

However, it is not always true that an assignee with a smaller $h$-index is outperformed by another assignee with a greater $h$-index. For example, in the scenario shown in Fig. 5, $S_i$ still has a smaller $h$-index $n_i$ yet its rank-citation curve is not dominated

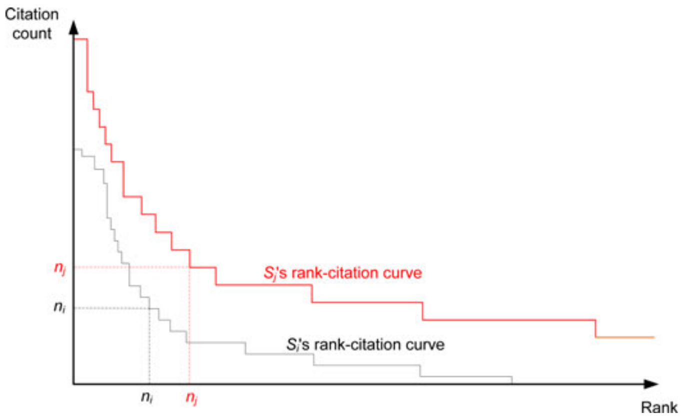


Fig. 4. Two fictitious rank-citation curves having domination relationship.

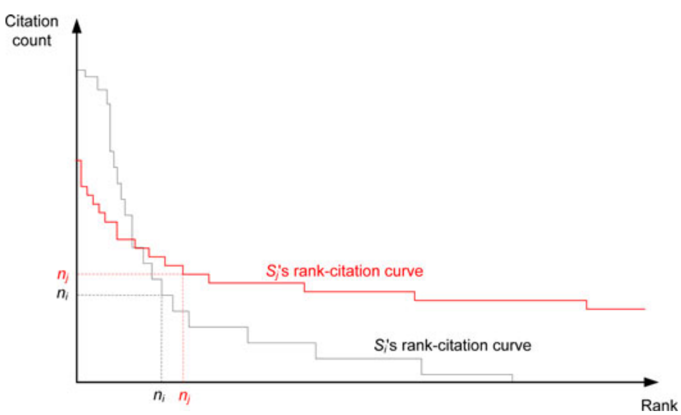


Fig. 5. Two fictitious rank-citation curves without domination relationship.

by that of $S_j$ having a greater $h$-index $n_j$ and we cannot be sure whether $S_i$ is indeed outperformed by $S_j$ without further investigation. The domination relationship fails when one rank-citation curve crosses or intersects with another rank-citation curve. The crossing or intersection can happen between their $h$-core segments (as shown in Fig. 5), their $h$-tail segments, or both, and can happen more than once.

As described above, using assignees' $h$-indices to predict their relative performance is not reliable. When the difference between $h$-indices is small, this prediction is not trustworthy. When the difference between $h$-indices is sufficient large, there still may be exceptions [27].

We, therefore, need additional information to further verify or differentiate 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:

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

$$t\text{−descriptor} = \sum_{i=n+1}^{N} i\left(\frac{C(P_i)}{A_t}\right) = \sum_{i=n+1}^{N_c} i\left(\frac{C(P_i)}{A_t}\right)$$

$$= \frac{\sum_{i=n+1}^{N_c} iC(P_i)}{\sum_{i=n+1}^{N_c} C(P_i)}. \quad (2)$$

According to (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 $h$-tail segments, respectively. Considering two assignees having $h$-core areas of identical size, (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, (2) would achieve a greater $t$-descriptor for the assignee whose $h$-tail segment slops more gently to the farther right. The $c$-descriptor favors the assignees having a sharp spike in their $h$-core so as to account for the excessive citation therein, and the $t$-descriptor favors the assignee having a long $h$-tail, especially when there are a large number of cited patents that cannot make into the $h$-core.

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 can expect that, if $h_i$ is sufficiently greater than $h_j$, $c_i > c_j$ indicates that $S_i$'s $h$-core segment runs higher above, and $t_i > t_j$ indicates that $S_i$'s $h$-tail segment extends farther to the right, thereby verifying the domination relationship among their rank-citation curves. However, $c_i < c_j$ or $t_i < t_j$ indicates an opposite scenario where their $h$-core or the $h$-tail segments cross each other, and further verification is required.

Kuan *et al*. [27] have conducted empirical study and verified that the foregoing shape descriptors are accurate and reliable. However, the application of the shape descriptors to a large number of assignees is quite cumbersome. The assignees have to be sorted first in accordance with their respective $h$-indices. Then, the assignees' relative performance with respect to their $h$-core and $h$-tails is further investigated by comparing their $c$-descriptors and $t$-descriptors, respectively. In the process, a significant number of pair-wise comparisons are required, and, most of all, it is difficult to gain an overall view of the relative positions of the assignees' performance.

## V. Utilizing Area Centroid for Portfolio Performance Characterization

For a large number of assignees, we envision an ideal approach to be a 2-D scheme where an assignee's portfolio performance is represented by a characteristic point in a 2-D coordinate system, preferably with the characteristic point's horizontal and vertical coordinates capturing the quantity and quality sides of the assignee's portfolio, respectively. Through this 2-D scheme, a large number of assignees' portfolio performance can be simultaneously depicted in the 2-D coordinate systems and their relative performance can be immediately determined by observing the relative positions of their characteristic points.

In developing the characteristic point, interestingly, we notice that the calculations of the $c$- and $t$-descriptors specified by (1) and (2) are actually very similar to how the area centroids (i.e., geometric center) of the $h$-core and $h$-tail areas are determined. As a matter of fact, given the $c$- and $t$-descriptors, we can immediately derive the $y$-coordinate of the $h$-core area centroid and the $x$-coordinate of the $h$-tail area centroid.

As mentioned earlier, the $h$-core and $h$-tail areas can be considered as consisting of $n$ and $(N_c - n)$ rectangles, respectively, where each rectangle has width 1, height $C(P_i)$, and area size $C(P_i)$. The area centroid for each rectangle therefore is located at $(i - 0.5, C(P_i)/2)$, $1 \le i \le N_c$. According to geometry, the centroid of a planar shape divisible into a number of smaller constituent shapes can be obtained as the weighted average of the centroids of these smaller constituent shapes. Therefore, the $h$-core area centroid $(c_x, c_y)$ and $h$-tail area centroids $(t_x, t_y)$ can be obtained as follows:

$$c_y = \sum_{i=1}^{n} \frac{C(P_i)}{2} \left(\frac{C(P_i)}{A_c}\right) = \frac{1}{2} \frac{\sum_{i=1}^{n} C(P_i)^2}{\sum_{i=1}^{n} C(P_i)} = \frac{1}{2} c\text{−descriptor} \quad (3)$$

$$c_x = \sum_{i=1}^{n} (i - 0.5) \left(\frac{C(P_i)}{A_c}\right) = \frac{\sum_{i=1}^{n} (i - 0.5) C(P_i)}{\sum_{i=1}^{n} C(P_i)} \quad (4)$$

$$t_y = \sum_{i=n+1}^{N_c} \frac{C(P_i)}{2} \left(\frac{C(P_i)}{A_t}\right) = \frac{1}{2} \frac{\sum_{i=n+1}^{N_c} C(P_i)^2}{\sum_{i=n+1}^{N_c} C(P_i)} \quad (5)$$

$$t_x = \sum_{i=n+1}^{N_c} (i - 0.5) \left(\frac{C(P_i)}{A_t}\right) = \frac{\sum_{i=n+1}^{N_c} (i - 0.5) C(P_i)}{\sum_{i=n+1}^{N_c} C(P_i)} = t\text{−descriptor} - 0.5. \quad (6)$$

As indicated by (3)–(6), the $c$-descriptor is exactly twice as large as the $h$-core area centroid's $y$-coordinate, $c_y$, and the $t$-descriptor on the other hand is basically identical to the $h$-tail area centroid's $x$-coordinate, $t_x$. If the sorted patents are arranged so that the rectangles are shifted to the left for a distance 0.5, the $t$-descriptor is identical to $t_x$.

Similar to the $h$-index as a characterization to the entire rank-citation curve or the entire area beneath the rank-citation curve, the $h$-core and $h$-tail area centroids can be considered as characterizations to the $h$-core and $h$-tail areas. According to (3)–(6), $c_x$ and $t_x$ are weighted average of ranks, and $c_y$ and $t_y$ are weighted averages of citations. The formers are related to the quantity side, whereas the latters are related to the quality side of the $h$-core and $h$-tail. Additionally, if the shapes of the $h$-core and $h$-tail areas are more skewed to the left, $c_y$ and $t_y$ would be higher, and $c_x$ and $t_x$ would be smaller, as more weight is given to the highly cited patents. In contrast, if the shapes of the $h$-core and $h$-tail areas extend smoothly to the right, $c_x$ and $t_x$ would be greater, whereas $c_y$ and $t_y$ reflect a characteristic height for the slow slopes of the $h$-core and $h$-tail segments.

We propose to use the $h$-core and $h$-tail area centroids as the characteristic points, as they are the geometric centers and thereby characterizations to the shapes of the $h$-core and $h$-tail areas, respectively. Even though it should not be hard to design numerous other characteristic points, such as $(n, A_e)$, $(n, A_c)$, $(n, A_t)$, or various combinations of $h$-indices and area-based $h$-type indices, we believe that the shapes of the rank-citation curves or the shapes of the $h$-core and $h$-tail areas are more discriminating than the area sizes.

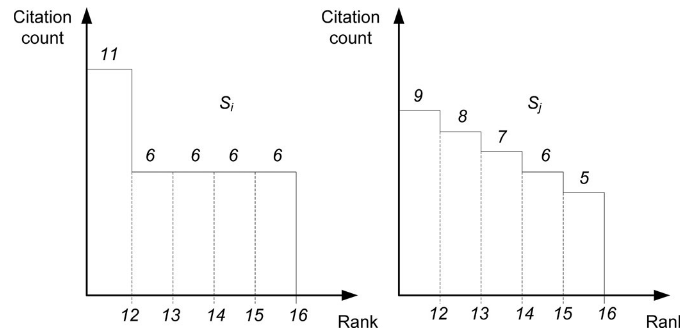


Fig. 6. $h$-tail areas of two fictitious assignees.

Second, despite the effectiveness of the $c$- and $t$-descriptors, there are scenarios where single $c$- or $t$-descriptor cannot achieve differentiation. Fig. 6 gives a simplified example where the $h$-tail areas of two fictitious assignees $S_i$ and $S_j$, both with $h$-index 11, $A_t$ 35, and $N_c$ 16, are depicted. The $t$-descriptors for $S_i$ and $S_j$ are both 13.7 (therefore, $t_x$ is 13.2), yet their different $t_y$'s, 3.8 for $S_i$ and 3.6 for $S_j$, successfully indicate their $h$-tail areas are differently shaped. In other words, when one coordinate of the area centroids fails to provide discrimination, the other coordinate of the area centroids can still be useful.

Kuan *et al.* [28] have empirically demonstrated that the foregoing area centroids are indeed an effective means for positioning the relative performance of a large number of patent assignees. However, an analyst has to deal with area centroids for the assignees' $h$-cores in one analysis and for the assignees' $h$-tails in a separate analysis. In order to make things even more simplified, can these two separate analyses be combined into a single one? Furthermore, the uncited patents are still not reflected in the shape descriptors and area centroids due to their zero citation.

## VI. Accounting for Uncited Patents

To achieve the envisioned single analysis, we may be tempted to use the centroid $(r_x, r_y)$ of the area beneath the rank-citation curve (whose area size is denoted as $A = A_c + A_t$). The point $(r_x, r_y)$ can be obtained similarly as (3) and (4) or (5) and (6) except that the summation over $i$ is from 1 to $N_c$ and $A$ is used instead of $A_c$ or $A_t$. Even though a single analysis is indeed achieved, the $(N - N_c)$ uncited patents are still not taken into consideration.

Therefore, instead of using $(r_x, r_y)$, we propose to use the centroid $(r_x', r_y')$ of an *h-complement area* to capture an assignee's overall performance embodied in its entire set of patents. The $h$-complement area, as illustrated in Fig. 7, is above the entire rank-citation curve, but is bounded by a rectangle whose upper right corner is located at $(N, C(P_1))$. The $h$-complement area size is denoted as $A' = N \cdot C(P_1) - A$.

Since the rectangle consists of the two areas beneath and above the rank-citation curve and has its centroid located at $(N/2, C(P_1)/2)$, the two areas' centroids $(r_x, r_y)$ and $(r_x', r_y')$

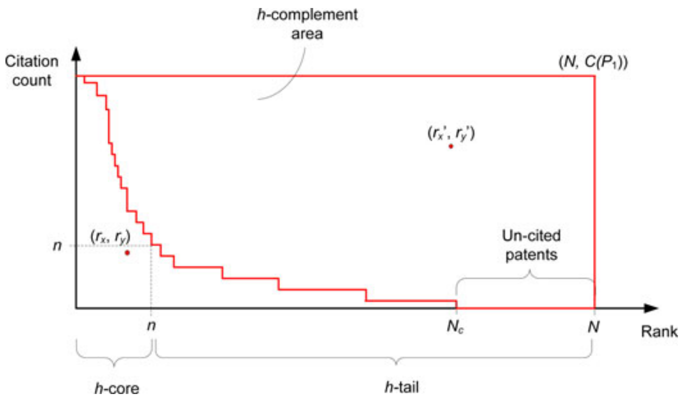


Fig. 7. $h$-complement area and its centroid.

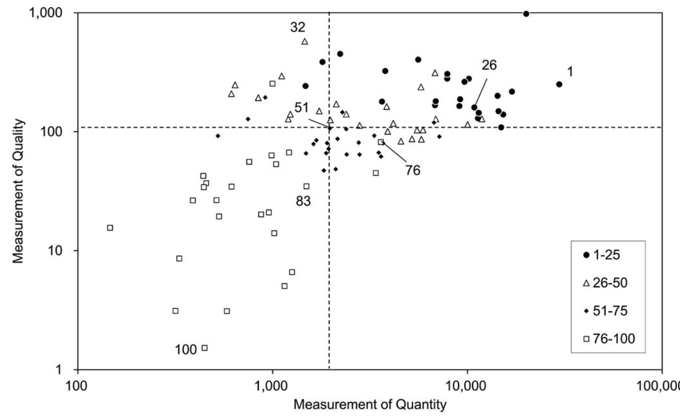


Fig. 8. Distribution of $h$-complement area centroids of the 100 assignees (log-scaled $x$- and $y$-axes).

satisfy the following equations:

$$\frac{C(P_1)}{2} = \frac{r_y A + r_y' A'}{A + A'} = \frac{r_y A + r_y' A'}{N \cdot C(P_1)} \quad (7)$$

$$\frac{N}{2} = \frac{r_x A + r_x' A'}{A + A'} = \frac{r_x A + r_x' A'}{N \cdot C(P_1)}. \quad (8)$$

Then, ($r_x$', $r_y$') can be easily obtained after $(r_x, r_y)$ is determined.

Intuitively, if $(c_x, c_y)$ and $(t_x, t_y)$ faithfully characterize the shapes of the $h$-core and $h$-tail segments or the $h$-core and $h$-tail areas, as suggested in previous sections, we can expect that ($r_x$', $r_y$') also faithfully characterizes the shape of the entire rank-citation curve (including the segments corresponding to the $(N-N_c)$ uncited patents) or the entire $h$-complement area and, therefore, the overall performance of the assignee's patent portfolio.

Additionally, the previous section points out that the $x$-coordinates, $c_x$ and $t_x$, represent the respective quantity sides of an assignee's $h$-core patents and $h$-tail patents (that is, how good the assignee is at producing these patents), and the $y$-coordinates, $c_y$ and $t_y$, represent the respective quality sides (that is, how good these patents are at attracting citations). We, therefore, expect that $r_x$' and $r_y$' should have similar functionality.

Empirical evidence to the roles of $r_x$' and $r_y$' is provided later. For the moment, we can imagine that, if an assignee continuously produces new patents and these new patents are not cited yet, $r_x$' would increase and ($r_x$', $r_y$') would move toward the right as $N$ increases and the $h$-complement area extends laterally. Similarly, if an assignee continuously receives more citations whereas its $N$ and $C(P_1)$ remain the same, $r_y$' would increase and ($r_x$', $r_y$') would move upward as $A$ increases and the $h$-complement area shrinks vertically. With the reasoning above, we can generally and quite comfortably consider that $r_x$' and $r_y$' have reflected the quantity (including both cited and uncited patents) and quality of an assignee's portfolio, respectively.

## VII. Positioning an Assignee's Portfolio Performance

### A. Two-dimensional Portfolio Performance Snapshot

The centroids of the 100 assignees' $h$-complement areas are obtained according to (7) and (8) and plotted in Fig. 8 with log-scaled $x$- and $y$-axes for better readability. For the same set of assignees, their $h$-core and $h$-tail area centroids according to (3)–(6) were obtained and figures similar to Fig. 8 were developed in [28].

The centroids in Fig. 8 are plotted with different markers depending on the corresponding assignees' ranks by $h$-index. For assignees ranked from the 1st to the 25th places, from the 26th to the 50th places, from the 51st to the 75th places, and from the 76th to the 100th places, four different markers, solid circles, hollow triangles, solid diamonds, and hollow squares, are used, respectively. The centroids of the assignees ranked at the 1st (IBM with $h$-index 161), 26th (Mitsubishi with $h$-index 82), 51st (Nortel Networks with $h$-index 60), 76th (Hon Hai Precision with $h$-index 36), and 100th places (LG Display Co. Ltd. with $h$-index 3) are specifically labeled with their respective ranks.

As illustrated, Fig. 8 provides a 2-D snapshot to the overall portfolio performance of the 100 assignees at the end of 2009, with highly cited $h$-core patents, lowly cited $h$-tail patents, and uncited $h$-tail patents all taken into consideration.

With the help of the different markers, we can see that the $h$-index is not entirely useless to the overall performance evaluation for patent assignees. For example, despite the entanglement of the centroids, assignees ranked in the top 25% generally have their solid-circle centroids positioned to the upper right of the hollow-square centroids of those in the last 25%. More specifically, the assignee ranked at the first place is always to the upper right of the assignees at the 51st and 100th places, the assignee at the 100th place is always to the lower left of the assignees at the 51st and 1st places, and the assignee at the 51st place is always in the middle between the assignees at the 1st and 100th places.

To utilize the 2-D view of Fig. 8 for positioning competitors, imagine that a technology manager is working for Nortel Networks, the assignee ranked at the 51st place with $h$-index 60.

TABLE I
RELEVANT DATA FOR THE FIVE SAMPLE ASSIGNEES

| | MITSUBISHI | QUALCOMM | NORTEL | HYNIX | HON HAI |
|---|---|---|---|---|---|
| Quadrant | 1 | 2 | reference | 3 | 4 |
| $h$-index | 82 | 78 | 60 | 19 | 36 |
| Rank | 26 | 32 | 51 | 83 | 76 |
| $C(P_1)$ | 248 | 1,139 | 205 | 68 | 160 |
| $N$ | 23,243 | 2,898 | 3,810 | 2,930 | 7,061 |
| $N_c$ | 19,894 | 1,966 | 3,030 | 1,304 | 4,328 |
| $A$ | 191,800 | 35,610 | 36,414 | 4,674 | 27,968 |
| $A_c$ | 8,680 | 13,874 | 5,483 | 615 | 2,218 |
| $A'$ | 5,572,464 | 3,265,212 | 744,636 | 194,566 | 1,101,792 |
| $r_x$ | 4,611.33 | 293.63 | 647.50 | 325.89 | 1,040.27 |
| $r_y$ | 12.88 | 82.46 | 17.35 | 5.08 | 8.53 |
| $r_x$' | 11,862.78 | 1,461.60 | 1,966.49 | 1,492.36 | 3,593.71 |
| $r_y$' | 127.82 | 574.81 | 106.66 | 34.69 | 81.81 |

Then, by treating Nortel Networks' $h$-complement area centroid as a reference point, the centroids of the rest of the assignees are partitioned relative to the reference point into four quadrants as outlined by the crosshair.

For the assignees whose $h$-complement area centroids are located in Nortel Networks' first quadrant, they can be considered to have both superior quantity and quality and, for those whose $h$-complement area centroids are located in the third quadrant, they can be considered to have both inferior quantity and quality. As to those whose $h$-complement area centroids are located in the second (or fourth quadrant), they can be considered to have inferior quantity but superior quality (or superior quantity but inferior quality), relative to Nortel Networks, as verified in the next section.

## B. *Empirical Verification*

Kuan *et al.* [28] tested empirically that similar inferences based on the distribution $h$-core and $h$-tail area centroids are valid. To see that the inference is also valid for $h$-complement area centroids, we pick only one assignee from each quadrant of Nortel Networks since [28] has already provided basic confidence. Among the four assignees, Mitsubish (ranked at the 26th place with $h$-index 82) and Hon Hai Precision (ranked at the 76th place with $h$-index 36) are in the first and fourth quadrants. As to the second and third quadrants, we randomly choose Qualcomm Inc. (ranked at the 32nd place with $h$-index 78) and Hynix Semiconductor (ranked at the 83rd place with $h$-index 19).

The relevant data of the four assignees, as well as Nortel Networks, are summarized in Table I, and their rank-citation curves are plotted altogether in Fig. 9 where, for better readability, only the first 10 000 patents in their decreasingly sorted portfolios are included and log-scaled $x$-axis is used. As shown in Fig. 9, Mitsushi of the first quadrant outperforms Nortel Networks both in terms of quantity and quality, as its rank-citation curve is entirely above that of Nortel Networks. Similarly, Nortel Networks obviously outperforms Hynix Semiconductor of the third quadrant, as its rank-citation curve is entirely beneath that of

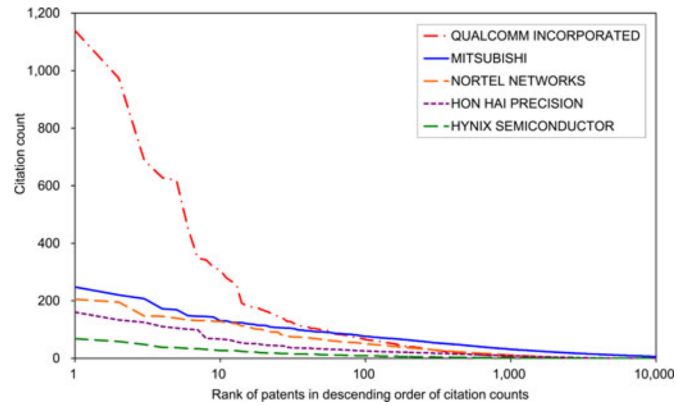


Fig. 9. Rank-citation curves for the five sample assignees of Table I (log-scaled $x$-axis).

Nortel Networks. These observations are verified by comparing the portfolios of Mitsushi and Hynix Semiconductor against that of Nortel Networks.

As to Hon Hai Precision in the fourth quadrant, even though it appears that its rank-citation curve runs completely beneath that of Nortel Networks in Fig. 9, actually its curve runs below Nortel Networks first, coincides with Nortel Networks after the 1848th patent, and then runs above Nortel Networks after the 2063rd patent. Since the two curves cross each other, we cannot immediately tell which one has better performance, but Hon Hai Precision's falling in the fourth quadrant indicates that it should have superior quantity yet inferior quality relative to Nortel Networks. We can see from Table I that this is indeed the case. Despite its greater number of cited and uncited patents, Hon Hai Precision's portfolio receives much fewer citations than Nortel Networks' smaller portfolio does.

We can also see from Table I that Qualcomm Inc. has a smaller number of patents and receives fewer total citations than Nortel Networks. However, its being positioned in the second quadrant suggests that it should have inferior quantity but superior quality relative to Nortel Networks. The quantity side is indeed true, but why should Qualcomm Inc. be considered to have superior quality? We can see from Fig. 9 that it has some patents so highly cited that, due to the way the $h$-complement area centroid is calculated, Qualcomm Inc.'s centroid is raised above that of Nortel Networks and landed in the second quadrant. This scenario reflects a property of our approach that a smaller number of highly cited patents may be considered to have achieved better quality than a larger number of mediocre patents. This property seems to be a reasonable one, especially considering that patents are rarely highly cited [36].

# VIII. TRACKING PORTFOLIO PERFORMANCE EVOLUTION

## A. *Trajectory of h-Complement Area Centroid*

With the 2-D performance snapshot, not only that we can tell who outperforms or is outperformed by a specific patent assignee in terms of quantity and quality, but also that we can see where the performance difference lies (i.e., in quantity or in quality or in both). Additionally, by the distance between two

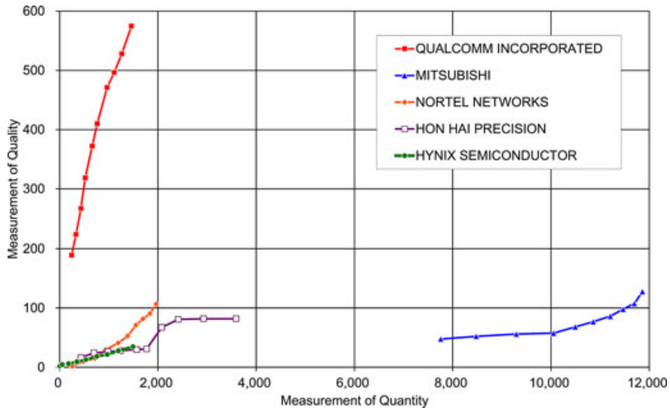


Fig. 10. Trajectories of $h$-complement area centroids from 2000 to 2009 for the five sample assignees.

assignees' $h$-complement area centroids, we can infer whether they have comparable or different or significantly disparate performance, instead of simply ordering them in a linear list as many patentometric indicators do.

Another major advantage of the proposed 2-D approach is that it can be extended to track the performance evolution for a number of competitors over a period of time, in addition to providing a static view at a specific epoch.

To illustrate this application, the five assignees of Table I are again used as example and their patents issued between 1976 and 2000, between 1976 and 2001, and so on up to between 1976 and 2009, are collected respectively. In other words, their patent portfolios since 1976 are obtained for each successive year between 2000 and 2009. Then, these portfolios' $h$-complement area centroids are obtained according to (7) and (8), plotted in a 2-D coordinate system, and connected chronologically into separate trajectories, as shown in Fig. 10. Similar trajectories for these portfolios' $h$-core and $h$-tail area centroids can be found in [27], [28].

As mentioned in Section VI, if an assignee's portfolio continuously increases with newly granted patents and the portfolio continuously receives new citations, it is very possible that its $h$-complement area centroid ($r_x$', $r_y$') would move toward the upper right. This speculation is indeed reflected in Fig. 10.

The historical view manifested in Fig. 10 provides a great deal of insight into how the portfolios of these assignees change over time. The trajectories shown in Fig. 10 generally develop gradually over time and reveal rather smooth trends. Even though Hon Hai Precision's trajectory undergoes some abrupt change between 2005 and 2007, it resumes a smooth trend after 2007. These smooth trends allow us to forecast the assignees' future performance by extrapolation.

On the other hand, when there is some abrupt change in the trend, such as Hon Hai Precision's jump between 2005 and 2007, such a scenario entails further investigation. After examining Hon Hai Precision's data, we find that some of its patents received unusually more citations in 2006 (e.g., D431825's citations increased from 51 in 2005 to 131 in 2006). For technology managers, special attention should be paid to these patents.

## *B. Empirical Verification*

We find that, starting from 2006, Hon Hai Precision's number of uncited patents increased rapidly to 1408, 1626, 2064, and 2733 (all over 34% of its total number of patents) in subsequent years. Therefore, even though its lowly cited $h$-tail patents grew steadily both in terms of quantity and quality over these years [28], its trajectory in Fig. 10 runs flatly since its increased $h$-core and $h$-tail quality was significantly discounted by its large number of uncited patents.

Compared to Hon Hai Precision, Nortel Networks was inferior in terms of quantity and quality before 2003 as shown in Fig. 10. Yet, after that year, Nortel Networks caught up pretty fast in terms of quality and quantity. After examining its data, we find that Nortel Networks' uncited patents played a diminishing role as their percentage to the total number of patents dropped from 801 (44%) in 2003 to 780 (20%) in 2009. This is the major factor causing Nortel Networks' trajectory to shoot upward in the last few years, as shown in Fig. 10.

Our observation to Nortel Networks is applicable to Mitsubish as well, which is the most productive assignee, as their trajectories in Fig. 10 reveal similar trends. Mitsubish's data indeed confirm our speculation. Its uncited patents play an even diminishing role as their percentage to the total number of patents dropped from 23% (3516) in 2000 to 14% (3349) in 2009. This explains Mitsubishi's sharper growing quality in the last few years.

As to Hynix Semiconductor, it has achieved the least growth both in terms of quantity and quality, even though it was not the worst one in the first few years. Qualcomm Inc., on the other hand, has produced very highly cited patents and has achieved exceptional quality with very limited quantity. Actually, in 2009, its quantity was even worse than that of the Hynix Semiconductor. However, Qualcomm Inc.'s growth in quality seemed to saturate a bit after 2006 as the trajectory of $h$-complement area centroids starts to traverse a path of smaller slope.

Additionally, we can derive a number of inferences of these assignees' performance evolution. For example, we can see that Qualcomm Inc. has been consistently showing a trend that its quality grows at a much faster rate than the other assignees. On the other hand, both Mitsubishi and Nortel Networks present similar trends of increased quality. Then, as already seen, Hon Hai Precision is very productive in producing patents, yet its fast-growing portfolio does not have better quality, and Hynix Semiconductor has such a steady trend that requires no special attention.

We can also see from Fig. 10 how these assignees evolve against each other. For example, we can see that Hynix Semiconductor and Mitsubishi have always been located at Nortel Networks' third and first quadrants, respectively, in each successive year. However, Qualcomm Inc. was in Nortel Networks' first quadrant initially but shifted to the second quadrant in later years. This is because Nortel Networks caught up on the quantity side over the years. Similarly, we can see that, as Nortel Networks caught up on the quality side, Hon Hai Precision moved from the first quadrant finally into the fourth quadrant of Nortel Networks.

## IX. Conclusion

Based on the citation distribution of a competitor's decreasingly sorted patent portfolio, we define an $h$-complement area and then propose to use the centroid of the $h$-complement area, which characterizes the shape of the $h$-complement area or the shape of the entire rank-citation curve, to capture the competitor's portfolio performance.

The intelligence we gain from analyzing the static and historical views of the $h$-complement area centroids is valuable in making various business decisions. Assuming that we are considering one of the five assignees for acquisition and that the value of intellectual property is a major concern, clearly Qualcomm Inc. should be our first target. Alternatively, if we are the owner of Qualcomm Inc. and we are putting our patent portfolio on the market for sale, Fig. 10 then can be used as evidence to the valuation we put on the patent portfolio.

In this paper, we have only demonstrated the effectiveness of the $h$-complement area centroid and how it can be utilized. An interesting future topic would be to use some firm-level innovation or R&D metric together with the $h$-complement area centroid so as to probe into the technology strategy, business activity, or R&D decision of the firm.

The approaches proposed by this paper are rather simplified by considering only the number of patents and their citations, and a technology manager has to utilize these approaches with their limitations in mind. For example, patents do expire and valid patents have different remaining lives but both factors are not considered in this paper. A more elaborate approach would be to remove the expired patents from the portfolio and apply some discounting scheme based on a patent's remaining life. However, these involve significant effort and can be interesting topics for further investigation. Additionally, competitors usually have several lines of products and patents in different technological areas should be treated separately, not altogether, as different technological areas involve different patterns of patenting activities. Also, different phases of a technology's life cycle may also involve different patterns of patenting activities. All these factors should be carefully weighted so as not to make misleading conclusions.

## References


[1] P. Almeida, "Knowledge sourcing by foreign multinationals: Patent citation analysis in the U.S. semiconductor industry," *Strategic Manage. J.*, vol. 17, Winter Special Issue, pp. 155–165, 1996.

[2] S. Alonso, F. J. Cabrerizo, E. Herrera-Viedma, and F. Herrera, "*h*-Index: A review focused in its variants, computation and standardization for different scientific fields," *J. Informetr.*, vol. 3, no. 4, pp. 273–289, 2009.

[3] W. B. Ashton and R. K. Sen, "Using patent information in technology business planning-I," *Res. Technol. Manage.*, vol. 31, no. 6, pp. 42–46, 1988.

[4] B. L. Basberg, "Patents and the measurement of technological change: A survey of the literature," *Res. Policy*, vol. 16, no. 2–4, pp. 131–141, 1987.

[5] L. Bornmann and H.-D. Daniel, "The state of h index research, is the h index the ideal way to measure research performance?" *EMBO Rep.*, vol. 10, no. 1, pp. 2–6, 2009.

[6] L. Bornmann, R. Mutz, and H.-D. Daniel, "The *h* index research output measurement: Two approaches to enhance its accuracy," *J. Informetr.*, vol. 4, no. 3, pp. 407–414, 2010.

[7] A. Breitzman and P. Thomas, "Using patent citation to target/value M&A candidates," *Res. Technol. Manage.*, vol. 45, no. 5, pp. 28–36, 2002.

[8] A. K. Chakrabarti, I. Dror, and N. Eakabuse, "Interorganizational transfer of knowledge: An analysis of patent citations of a defense firm," *IEEE Trans. Eng. Manage.*, vol. 40, no. 1, pp. 91–94, Feb. 1993.

[9] R. Chao and E. Loutskina. (2011). "How complexity impacts R&D portfolio decisions and technological search distance," Darden School of Business Working Paper No. 1917607 [Online]. Available at http://ssrn.com/abstract=1917607

[10] C. Chen and D. Hicks, "Tracing knowledge diffusion," *Scientometrics*, vol. 59, no. 2, pp. 199–211, 2004.

[11] R. Costas and M. Bordons, "The *h*-index: Advantages, limitations and its relation with other bibliometric indicators at the micro level," *J. Informetr.*, vol. 1, no. 3, pp. 193–203, 2007.

[12] N. J. van Eck and L. Waltman, "Generalizing the *h*- and g-indices," *J. Informetr.*, vol. 2, no. 4, pp. 263–271, 2008.

[13] L. Egghe, "The Hirsch-index and related impact measures," *Annu. Rev. Inf. Sci.*, vol. 44, pp. 65–114, 2010.

[14] L. Egghe, "Characteristic scores and scales based on *h*-type indices," *J. Informetr.*, vol. 4, no. 1, pp. 14–22, 2010.

[15] H. Ernst, "The use of patent data for technological forecasting: The diffusion of CNC-technology in the machine tool industry," *Small Bus. Econ.*, vol. 9, no. 4, pp. 361–381, 1987.

[16] H. Ernst, "Patent information for strategic technology management," *World Patent Inf.*, vol. 25, no. 3, pp. 233–242, 2003.

[17] M. A. García-Pérez, "A multidimensional extension to Hirsch's *h*-index," *Scientometrics*, vol. 81, no. 3, pp. 779–785, 2009.

[18] W. Glänzel and A. Schubert, "Characteristic scores and scales in assessing citation impact," *J. Inf. Sci.*, vol. 14, no. 2, pp. 123–127, 1988.

[19] J. C. Guan and X. Gao, "Exploring the *h*-index at patent level," *J. Amer. Soc. Inf. Sci. Technol.*, vol. 60, no. 1, pp. 35–40, 2009.

[20] L. Fleming, "Recombinant uncertainty in technological search," *Manage. Sci.*, vol. 47, no. 1, pp. 117–132, 2001.

[21] L. Fleming and O. Sorenson, "Technology as a complex adaptive system: Evidence from patent data," *Res. Policy*, vol. 30, no. 7, pp. 1019–1039, 2001.

[22] B. H. Hall, A. B. Jaffe, and M. Trajtenberg. (2000). "Market value and patent citations: A first look," NBER, Cambridge, MA, Working paper 7741. [Online]. Available at http://www.nber.org/papers/w7741.pdf

[23] B. H. Hall, A. B. Jaffe, M. Trajtenberg. (2001). "The NBER patent citations data file: Lessons, insights and methodological tools," NBER, Cambridge, MA, Working Paper 8498. [Online].Available: http://www.nber.org/papers/w8498.pdf

[24] J. E. Hirsch, "An index to quantify an individual's scientific research output," *Proc. Natl. Acad. Sci. USA.*, vol. 102, no. 46, pp. 16569–16572, 2005.

[25] M.-H. Huang, 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, no. 3, pp. 489–506, 2003.

[26] A. Jaffe, "Technological opportunity and spillovers of R & D: Evidence from firms' patents, profits, and market value," *Am. Econ. Rev.*, vol. 76, no. 5, pp. 984–1001, 1986.

[27] C.-H. Kuan, M.-H. Huang, and D.-Z. Chen, "Ranking patent assignee performance by *h*-index and shape descriptors," *J. Informetr.*, vol. 5, no. 2, pp. 303–312, 2011.

[28] C.-H. Kuan, M.-H. Huang, and D.-Z. Chen, "Positioning research and innovation performance using shape centroids of *h*-core and *h*-tail," *J. Informetr.*, vol. 5, no. 4, pp. 515–528, 2011.

[29] S.-J. Liu and J. Shyu, "Strategic planning for technology development with patent analysis," *Int. J. Technol. Manage.*, vol. 13, no. 5/6, pp. 661–680, 1997.

[30] E. Mansfield, "Patents and innovation: An empirical study," *Manage. Sci.*, vol. 32, no. 2, pp. 173–181, 1986.

[31] M. Meyer, "What is special about patent citations? Differences between scientific and patent citations," *Scientometrics*, vol. 49, no. 1, pp. 93–123, 2000.

[32] M. Meyer and S. Bhattacharya, "Commonalities and differences between scholarly and technical collaboration," *Scientometrics*, vol. 61, no. 3, pp. 443–456, 2004.

[33] M. E. Mogee, "Using patent data for technology analysis and planning," *Res. Technol. Manage.*, vol. 34, no. 4, pp. 43–49, 1991.

[34] M. E. Mogee and R. G. Kolar, "International patent analysis as a tool for corporate technology analysis and planning," *Technol. Anal. Strateg.*, vol. 6, no. 4, pp. 485–503, 1994.

[35] Patent Technology Monitoring Team Report. (2010). "Patenting by Organizations 2009," USPTO, Alexandria, VA. [Online]. Available at http://www.uspto.gov/web/offices/ac/ido/oeip/taf/topo_09.htm

[36] F. Narin and D. Olivastro, "Technology indicators based on patents and patent citations," in *Handbook of Quantitative Studies of Science and Technology*, A. F. J. van Raan, Ed. Amsterdam, The Netherlands: Elsevier, 1988, pp. 465–507.
[37] R. Rousseau, "New developments related to the Hirsch index" (in Chinese)," *Sci. Focus*, vol. 1, no. 4, pp. 23–25, 2006, ([Online]. An English translation is available at http://eprints.rclis.org/6376/.)
[38] R. Rousseau, "Reflections on recent developments of the *h*-index and *h*-type indices," *COLLNET J. Scientometrics Inf. Manage.*, vol. 2, no. 1, pp. 1–8, 2008.
[39] T. Stuart and J. Podolny, "Local search and the evolution of technological capabilities," *Strategic Manage. J.*, vol. 17, no. S1, pp. 21–38, 1996.
[40] F. Y. Ye and R. Rousseau, "Probing the *h*-core: An investigation of the tail-core ratio for rank distributions," *Scientometrics*, vol. 84, no. 2, pp. 431–439, 2010.
[41] C.-T. Zhang, "The e-index, complementing the *h*-index for excess citations," *PLoS ONE*, vol. 4, no. 5, p. e5429, 2009.

**Chung-Huei Kuan** received the B.S. degree in electrical engineering from National Taiwan University in 1983, the M.S. degree in electrical and computer engineering from the University of Massachusetts, Amherst, in 1989, and the Ph.D. degree in industrial engineering and management from the mechanical engineering Department, National Taiwan University, in 2012.

He is currently an Assistant Professor at the Graduate Institute of Patent, National Taiwan University of Science and Technology, Taipei, Taiwan. He is also the Director of the University's Technology Transfer Center and Intellectual Property Business Center. His research interest includes patentometrics, patent information mining, and bibliometrics. Most of his research is inspired by his work in managing and promoting the intellectual properties and technologies for the university.

**Mu-Hsuan Huang** received the B.A. and M.A. degrees in library science form National Taiwan University, Taiwan and the Ph.D. degree in library and information science from the University of Maryland, MD.

She is currently a Professor in the Department of Library and Information Science, National Taiwan University, Taipei, Taiwan. Her recent research interests include bibliometrics, science and technology policy, intellectual property, and patent information. She is also the Project Investigator for several national-scale research projects, including the Performance Ranking of Scientific Papers for World Universities (Taiwan Ranking) funded by Higher Education Evaluation & Accreditation Council of Taiwan, and National Science and Technology Program-Energy funded by National Science Council, Taiwan.

**Dar-Zen Chen** received the M.A. and Ph.D. degrees in mechanical engineering from the University of Maryland, MD.

He is currently a Professor in the Department of Mechanical Engineering and Institute of Industrial Engineering, National Taiwan University, Taipei, Taiwan. His research interest includes robotics, automation, kinematics, and mechanism design. He has also worked intensively on topics in intellectual property management, patentometrics, and competitive analysis. He also participates in several national-scale research projects such as National Science and Technology Program-Energy funded by National Science Council, Taiwan. He is the corresponding author of this paper.