International Journal of Hybrid Information Technology
Vol.7, No.3 (2014), pp.285-302
http://dx.doi.org/10.14257/ijhit.2014.7.3.27

Improving the Sensitivity of Research Publication Indices Using Exponential Function and the Area under Curve

Majdi A. Maabreh [1], Izzat M. Alsmadi [2] and Mohammed N. Al-Kabi [3]

[1]*Computer Center, Hashemite University, P.O. Box 150459, 13115 Zarqa, Jordan*
[2] *Department of Information Systems, Prince Sultan University, Riyadh, Saudi Arabia*
[3] *Faculty of Sciences and Information Technology, Zarqa University, 13132 Zarqa, Jordan*
majdi@hu.edu.jo, ialsmadi@cis.psu.edu.sa, malkabi@zu.edu.jo

Abstract

In collecting and analyzing information about recent research papers, authors, and journals with their citation and indices, we noticed that in many cases such indices fail to reflect reality and are incapable to distinguish between different authors in some cases while they clearly have different research profiles. There are many research papers that can be found in literature that discussed the behavior and limitations of the current citation or publication indices. Most of them referred to the sensitivity problem and the fact that most indices neglect some publications during index calculation. These two problems are the most important limitations that may judge index fairness. In this paper, we proposed new indices; MIExp-Index using exponential function and MIArea-Index using the concept of the area under curve (AUC) in order to enhance the sensitivity and thus the fairness of index assessment. To demonstrate our two novel indices, data of real and hypothetical authors were used to demonstrate these two approaches. The results showed that both of the proposed indices are very sensitive for each single citation that the paper receives after being published. In addition, the two indices consider all authors' publications, even the new ones, into account.

Keywords: *H-Index, MIEXP-Index, MIArea-Index, Research assessment, Impact Factor, Citations-based indices.*

1. Introduction

The need to a criterion to assess the quality of scientific outputs of a scholar (scientist), conference, journal, university, etc. leads to the emergence of a number of performance indicators such as *H-index* , *G-index*, etc. These performance indicators are classified into three classes according to (Franceschet, 2009); Productivity metrics, impact metrics and hybrid metrics where the number of citations is mainly used to assess the performance indicators of those three classes. These indicators are essential in our present time to make objective judgment about activities such as: Scientific promotion, research grants, recruitment, selection of scientific committees, and awards.

The assessment of research quality was based on one of the bibliometric indicators (*e.g.* total number of papers, total number of citations) before inventing performance indicators in 2005 by Jorge Hirsch. This means that no formula exists to combine more than one bibliometric indicators. The proposal of the first impact metric (citation index) *H-index* by Argentine American physicist Jorge Hirsch (Hirsch 2005) invokes scientists to propose

ISSN: 1738-9968 IJHIT

several other performance metrics in general and impact metrics in particular. These metrics are used to evaluate the research impact of a particular: author, paper or journal. Jorge Hirsch metric (*H-index*) is used to quantify the scientific output of a single researcher by one indicator, where this indicator combines the number of publications and number of citations in one formula. Hirsch indicator (*H-index*) is an objective metric which is easy to compute and verify. *H-index* is considered better than one criterion indicators such as: Total number of citations, rate of citation per paper, or total number of publications. *H-index* receives a positive acceptance by a number of studies which shows its usefulness such as (Bornmann, & Daniel, 2005). Glänzel (2006b) shows also that one of the *H-index* strengths is its ability to assess a small set of articles. (Cronin, & Meho, 2006) used the H-index to rank 31 well-known information science scientists, where the self citation is included and excluded to conclude that self citation has no influential effect on the rank of those scientists. Many authors discussed the pros and cons of *H-index*. *Further*, many variants to Hirsch indicator are casted. For example (Bornmann and Daniel, 2007) found that around 30 papers published during the first year after publishing *H-index*.

Several researchers realized shortly after publishing Hirsch index that it can be used to evaluate also scientific journals, universities, members of national academies, etc. (Braun, Glänzel, and Schubert, 2006). Therefore the scope of using *H-index* is not restricted to quantifying the research outputs of different scientists around the world.

Hirsch himself in (2007) argues that his index can be used to predict any scientist's scientific research future productivity, and proves it is better for prediction than other bibliometric indicators. He claims also that *H-index* is a better predicator than other indicators even those variants which assign more weight on highly cited articles.

Several researchers criticize the *H-index* early since it was released. (Glänzel, 2006a) lists 8 main disadvantages of *H-index*. (Jin, Liang, Rousseau and Egghe, 2007) also presents 14 pros and cons of *H-index* in their study. (Hirsch 2005) noted that the scientific research field may affect the *H-index* value, therefore the highest *H-index* value of a physicist (namely, Witten E) was 110, and the highest *H-index* value of a scientist (namely, Snyder SH) was 192 where difference between both is significant. Egghe (2006a) noticed early that once a paper belongs to the *h*-defined inner class (known as *h*-core), it becomes ineffective whether it has 0 or 1000 extra citations, and H-index should be then restricted to compare scientist's with equal years of experience. Egghe (2006a) also noticed that *H-index* value is a non-decreasing one. As a result, an inactive scientist (later on his/her life) will maintain the same *H-index* value. Accordingly, the study (Egghe 2006b) proposed one of the most important variants of the *H-index* which is called *G-index*. The justification of proposing this new variant (*G-index*) was due to insensitiveness of the *H-index* to a set of non-cited (or modest cited) papers and to a set of highly cited papers. (Egghe, 2006c) shows that *G-index* is sensitive to both cases of low or high number of citations for publications.

Examples of other citation indices that were proposed in relevant research papers in addition to the *H-index* include: *g*-, *h*(2)-, *w*-, and *hw*-Index where those depend on *H-index* behavior in calculating the impact value, in addition to: *hg*-, q^2 -, *AR*-, *m*-quotient, *m*-,*e*-, *A*-, *r*- , *w*-, *j*-index, etc. where those directly depend on *H-index* value in their calculations. This dependency relation is considered as a limitation of the proposed indices as they should calculate the *h* value before their assessment.

Each index represents a different formula to calculate the impact value based largely on only two factors: Author's number of publications and number of citations for all his/her publications. Some of the proposed indices added new attributes such as: Publications dates or periods, average number of publications, citations, etc. Examples of such indices include: Contemporary *H-index*, *m*-quotient, and *AR*- index.

The number of citations is still used as the main value for research assessment in most current indices. In the web, each webpage takes its popularity from those WebPages that point to it regardless of the webpage contents. If the webpage has for example 100 In-links, then it can be considered as more popular than another one which has just 5 In-Links. In the research community, the matter is not different. The paper which has 100 citations is more important or influential than another one which has just 5 citations, regardless of the actual research value in both.

All indices, including our proposed one in this paper, do not take into account the contents but just the In-Links or citations. Since the number of citations is the primary value for research indices' calculations then this value should be thoroughly investigated and calculated. However, even if we assume that the citation values are accurately collected, the citation indices do not calculate these values in fairly manners. For example, if we have a software program that is developed for collecting authors' publications and citations, and the software returns information about an author who has just 5 publications (P_1, P_2, P_3, P_4, and P_5). Further, these publications have citation values of: {100, 50, 20, 10, and 5} respectively. According to the *H-index citation* point of view, the author has an *H-index* value of 5. However, the *H-index* value will not change, if we knew that the software while collecting papers' citations missed many citations for the authors' publications. Let's assume that the correct citation values are: {500, 250, 200, 10, and 9}. H-index in this case will also stay the same as earlier. This means that the *H-index* did not consider how many citations the paper has as far as the citation value exceeds its order value (*i.e.* 1, 2, 3, 4, etc.). If we assume that the software also misses some publications for the same author, for example {P_6, P_7, and P_8}, and these papers have citation values of: {2, 2, 1} respectively, these possibly recent publications will not change the *H-index* value. This insensitivity of the index to variations of a single author publications and citations or variations between different authors is what triggered our trial to propose more sensitive indices.

In this paper, new citation indices are proposed and new mathematical formulas are applied in order to tackle some of the issues mentioned above. We will benefit from the sensitivity of the exponential function as well as the concept of the area under curve in developing more sensitive indices which take all citations into consideration.

2. Related Work

H-index with its simplicity built a baseline for most of later publications indices. All *H-index* variants tried to solve one or more limitation of the *H-index*. For example, (Schreiber, Malesios, Psarakis, 2012) study presented 17 *H-index* variants and examined the performance and properties of each one of them. (Bornmann, Mutz, Hug, & Daniel 2011) conducted a larger study than the previous one and presented in their study the computation of correlations between the *H-index* and 37 of its variants. Authors found that the *H-index* is highly correlated with most of its variants. The high correlation between *H-index* and most of its variants means that these variants are somewhat redundant to *H-index*. (SCI2S, 2013) presented and classified a large number of indices based on *H-index*.

One of the disadvantages of *H-index* is in its neglecting of the number of authors in research papers and also their order in the paper. Therefore a complementary version of *H-index* called h_I is presented by (Batista, Campiteli, & Kinouchi, 2006) to solve this problem based on conclusions of previous studies showing that the number of authors generally influenced the number of citations. If all studies of an author are solo for the author as a single author then the *H-index* value is equal to h_I-index value. In addition, h_I-index values are generally less sensitive to the scientific research field relative to *H-index* values. Schreiber (2008) presented a solution to the number of authors issue by introducing a modified *H-index*

(h_m*-index*). Author compares the performance of *H-index*, h_I*-index*, and h_m*-index* using the scientific outputs of eight famous physicists, where author claims that h_m*-index* was a better representative for author's research production.

Egghe (2006b) proposed a new index called *G-index*. *G-index* represents an improvement to *H-index*, since it takes into its account highly cited papers, which are ignored by *H-index*. Nonetheless, *G-index* could not solve all the problems observed in *H-index*. For example both *H-index* and *G-index* are insensitive to the number of coauthors of a given article. To solve this problem h_I*-index* and h_m*-index* were introduced by (Batista, Campiteli and Kinouchi, 2006) and Schreiber (2008) as mentioned before. (Schreiber, 2009), Schreiber (2010a) and Schreiber (2010b) present a modified *G-index* (g_m*-index*) which takes into account the number of coauthors in each article. Therefore the values of g_m*-index* of different scientists who have multi-author articles will be lower than the values of *G-index* for the same scientists. This means that g_m*-index* will be in favor of scientists who have fewer numbers of authors.

One of the *H-index* issues is the precision problem. This occurs in some cases when an author publishes papers under different names as noticed by (Jin et al., 2007). Different names are due to misspellings, transliterated names, marriage, and changing names. Kosmulski (2006) *h(2)-index* is easier to compute than *H-index* and is defined as the highest number $h(2)$ of articles that have at least $((h(2))^2$ or more citations. *h(2)-index* aims to solve the problem of insensitivity of the *H-index* to highly cited articles. Kosmulski (2006) defines the scientist's *h(2)-index* as "the highest natural number such that $h(2)$ most-cited papers receive each at least $[h(2)]^2$ citations." *h(2)-index* helps to reduce the precision problem according to (Jin *et al.,* 2007). Wu (2010) proposes a novel performance index called *W-index*. It is also more sensitive than *H-index* for highly cited articles.

Jin (2006) and Rousseau (2006) discussed an *A-index* to measure citation intensity. The *A-index* is defined as the average number of citations received by the articles given in the *h-core*. Rousseau (2006) called it *A-index* since it is related to average values. It is argued that the *A-index* can solve the problem of highly cited articles ignored by *H-index*. This makes *A-index* close to the *G-index,* since it also solves the same problem. One of the problems facing the *A-index* is summarized by the possibility that a decrease in the value of *A-index* may occur when the number of citations is increased since formula involves a division by *H-index* value (Maabreh and Alsmadi, 2012).

Three different indices are presented by (Sidiropoulos, Katsaros, Manolopoulos, 2007), which are variants of *H-index* and called: normalized-, contemporary *H-index* and trend *H-indices*. The comparison of different scientists according to their publications faces the challenge that those authors may have significantly different profiles in terms of number of publications. As such, Normalized *H-index* is proposed to solve this problem. Trend *H-index* takes into its account citations' age, so that articles cited for few years get lower values than those cited for a long period of time. Contemporary *H-index* takes into its account the age of the article.

(Bornmann, Mutz and Daniel, 2008) paper proposes the *m*-index as a variation of the *A-index*. *M-index* is the median number of citations received by papers in the *h-core*. Median is one of the statistical central tendency measures. The number of citations received by a paper in the *h-core* is sorted in descending order to select the middle value easily. *M-index* is a simple computed metric and needs precedent determination of *H-index*. Citation distribution counts is usually skewed and is not included in *M-index* (Bornmann, Mutz, & Daniel, 2008). An increment in the *h-core* citations may not affect the value of *M-index* helps to reduce the impact of heavily cited papers in the *h-core*. Values of *M-index* are higher than their *H-index* counterparts.

As mentioned before the *H-index* value is a non-decreasing one and it is insensitive to the age of the paper and actual number of citations. Therefore another complementary version of *H-index* called *AR-index* was introduced by Jin (2007) to solve these limitations. The *AR-index* formula is characterized by its simplicity, since it is equal to "the square root of the sum of the average numbers of citations per year of articles included in the *h-core*". So the values of *AR-index* can decrease according to its formula.

(Jin, Liang, Rousseau and Egghe, 2007) proposed a novel metric to measure the citation intensity in the *h-core* called *R-index*. *R-index* is equal to the square root of the sum of all citations received by articles included in the *h-core* at time *T*. The *R-index* must be used in conjunction with the *H-index*.

(Egghe, and Rousseau, 2008) paper proposes a citation weighted *H-index* which is called h_w*-index*. This performance indicator aims to solve the problem of the *H-index* insensitivity to performance changes similar to the *AR-index*. h_w*-index* is equal to the square root of the total number of citation(*s*) received by the highest number of articles that each received *s/h* or more citations. However, similar to *A*-index issue, an increase in the number of citations could lead to a decrease in the value of h_w*-index*.

(Alonso, Cabrerizo, Herrera-Viedma and Herrera 2010) paper presents another index called *hg-index*. It is called so since it is based on both *H-index* and *G-index as it tried to combine both indices into one*. The main goal of *hg-index* is to maintain the pros of these two indices (*H-index* and *G-index*) and minimize their cons. Empirical tests of *hg-index by authors show its advantages on both original citations*.

(Cabrerizo, Alonso, Herrera-Viedma and Herrera, 2010) paper presents a metric called q^2*-index*. q^2*-index* includes two dimensions of scientist's scientific research output; First dimension represents the number of articles, and the second dimension represents the number of citations (impact of papers). Authors stated that: "q^2-index is based on the geometric mean of *H-index* and the median number of citations received by papers in the *h-core*, *i.e.*, the *M-index*, which allows us to combine the advantages of both indices."

3. Citations Indices

3.1. Sensitivity and Comprehensiveness Issues

By sensitivity problem we mean that most citations focus on the *h-core* or the publications that get the largest number of citations. For most citation indices, activities outside the *h-core* are ignored. For example, an author may get 30 new publications in the current year without impacting or changing his/her citation index value. Furthermore, activities in the *h-core* may not also impact the citation index. Citation indices are largely interested in the publications that are about to pass the edge from the *h-tail* to the *h-core* zone. The *H-index* value would not increase in response to citation increments gained by *h-core* publications. For example if an author has 5 publications that were cited as: (10, 6, 3, 3, 1) then author will have an *H-index* value of 3. If the publications in *h-core* (*i.e.* 10, 6, 3) gain more citations in next years, say (20, 15, 9), and no citations occur in the rest of publications then the *H-index* will not be changed.

To overcome the *H-index* limitations (Egghe, 2006) proposed the *G-index*, which is now a major index in addition to the *H-index*. *G-index* was proposed based on the behavior of *H-index* with some enhancements to improve its sensitivity. However, the "*h-core*" or "big-hits" problem still affect or decide the *G-index* calculations. The *G-index* partially overcomes the sensitivity problem in comparison with the *H-index*. However, the sensitivity problem can still be seen in the *G-index* with many scenarios. For example, if an author has 5 publications with citation numbers as: (7, 5, 4, 2, and 0), then he/she has a *G-index* value of 4. There will be no change on *G-index* value if the author papers' citations became as: (10, 7, 4, 2, and 0). The sensitivity problem was also raised with many of *h* type indices such as: *W*-, *h*(2)-, *M*-, *hg*-, *m*-quotient, q^2, *hw*- and *j*-indices. Another problem was seen through experiments with *M-index*, where the increase in the number of citations has no effect on the *M-index* value, if this increment would not lead to a median change. This problem is also clear in those indices which depend on square root function such as: *A*-, *R*-, *AR*-, *E*-, *hg*-, q^2-indices as examples.

In addition to the sensitivity problem, for what is called *h-tail* (Ye and Rousseau 2009) (*i.e.* the rest of the publications with the exclusion of *h-core* publications), there is no logical base to exclude those publications from *H-index* calculations. While, it makes sense to give more weight to publications with the high number of citations, however, this should not mean to, almost, totally ignore the rest of publications. This problem was also raised in all indices which depend on *h-core* contents such as: *A*-, *R*-, *AR*-, *E*-, *M*-, and *hg*. In many scenarios, the rest of publications may have more than 10 citations and may be neglected.

In terms of comprehensiveness, we discussed some of the earlier issues related to the limitations of citation indices and ignoring some publications, citations. Some other important information such as the number or sequence of authors, age of paper, etc. can be also categorized under incomprehensiveness issues.

3.2. MIExponential (*MIExp-Index*) Using the Exponential Function

Most of the indices presented in this study use the number of citations received by the articles to quantify the value of scientist's publications. Typically, high citation to an article should indicate useful information in the papers for others to cite from or refer to.

The number of citations received by the article varies from zero to thousands. We have to distinguish between the real number of citations (zero or positive integer) and the order of these values as in *H-index*. Using the number of citations could be impractical without preprocessing or normalizations.

Most of research studies can be cited from the date of its publication, and some studies are cited before its date of publication (*e.g.* in pre-processing or publication cases). The number of citations for each paper is accumulated through years. Suppose a paper *X* was published in 1990 and from its date of publication and till now each year the total number of citations of paper *X* is increased. If the present number of citations of paper *X* is 1000. This number was the lowest in the date of publication and currently is the maximum. In other ways, total number of citations should always increase, or stay fixed with time and not decrease. That means when the age of any paper is increased, so the number of citations for this paper should not be decreased.

A common logarithm with respect to the base $b = 10$ is used in our proposed formula to deal with the cumulative nature of the number of citations. Therefore, the number of citations is transformed into a smaller number. From a mathematical point of view logarithms can be used to find the growing rate of the individual authors' research paper during their academic age. Logarithms can also help to answer the following question: How can we model the citation growth in a given time interval from X level to $X + n$ level? We will add number 1 to the citation values in order to deal with those authors who have publications with zero citations since ($\text{Log}_{10} 0$ is undefined).

To get a sensitive formula to any change in the number of citations, a natural exponential function is used in our proposed formula (1), where e is used as a base for our proposed function. A common logarithm to base 10 of the number of citation plus 1 represents the exponent of our new proposed formula. The Exponential function is considered as a sensitive function in response to any change in the input values.

Then,

$$MIExp - Index = e^{\log_{10}(No.of\ citations+1)} \quad (1)$$

3.3. MIArea (*MIArea-Index*) Using the Concept of the Area under Curve

The area under curve is a mathematical concept that can be used in many applications. In this study, we used this concept to represent authors' publications and their citation values. The *y*-axis corresponds to the number of citations received by a paper, whereas the *x*-axis represents the number of authors' publications. For example, Table 1 represents such authors' research papers.

Table 1. A sample of author research papers

No. of Publications	1	2	3	4	5	6	7	8	9	10
No. of Citations	12	7	5	4	4	3	3	1	1	1

The citation curve exhibited in Figure 1 was based on the citations presented in Table 1. The citation curve shows clearly the drop out in the number of citations.

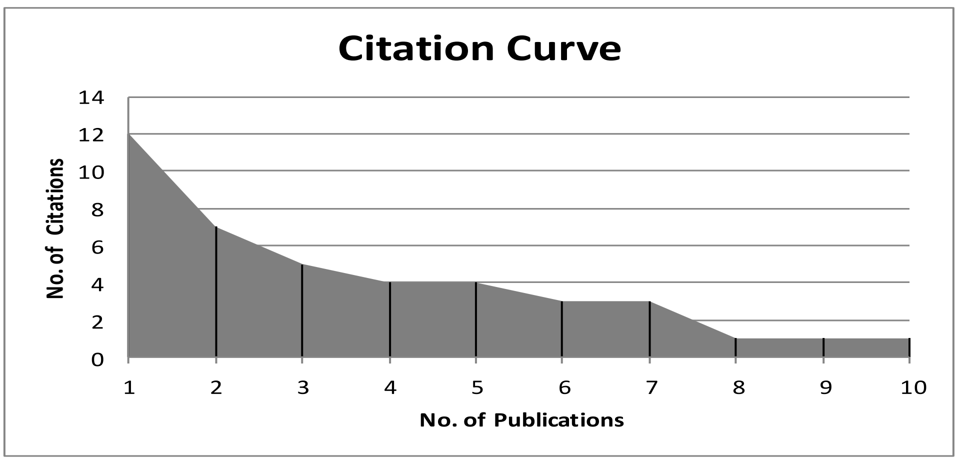

Figure 1. Citations vs. Publications

We may use two methods to find the area under curve or the area between the curve and the *x*-axis. First we can use a definite integral to compute area under a curve. However this method needs the presence of a function which generates *y*-values which can be computationally complex. In fact, there is no clear, formal or general relationship between number of publications and number of citations. Here, we assume that papers for each author with their citations can form an area chart like those areas presented in Figure 1. The second method needs not any knowledge in calculus as MS Excel trend line option can be a simple solution.

In Figure 1, each publication and its citations form a sub-area of the whole chart. Furthermore, each sub-area is a trapezoid or a special type of trapezoid. As such, to estimate the area of a trapezoid, the following formula is used:

$$\frac{b-a}{2N}\left(f(x_1)+2f(x_2)+2f(x_3)+\ldots 2f(x_N)+2f(x_{N+1})\right)$$

Where, the value of *b* represents the order of the author latest publication(s) (1, 2, 3, *n*). The value of *a* in this approach is always equal to 1. *N* refers to the number of generated trapezoids, and it equals to $(b - a)$. $f(x_i)$ refers to the number of citations that the paper (i) received, where $i = 1, 2, 3, \ldots, n$.

Conditions:

Using the Area under Curve (AUC) discussed above is regulated by the following conditions:

1- It is applicable only for those authors who have more than one publication. The author who has one publication will get ZERO area units even if the single publication gained more than 100 citations. The value Zero means then that the author has no or single publication.
2- The citation values should follow the *H*-index order (i.e. descending order).
3- The publications with zero citation will be ignored as there is no specific impact yet.

For the above example shown in Figure 1, the values are ordered as follows: (12, 7, 5, 4, 4, 3, 3, 1, 1, and 1).

$$AUC = \frac{10 - 1}{2 \times 9}(12 + (2 \times 7) + (2 \times 5) + (2 \times 4) + (2 \times 4) + (2 \times 3) + (2 \times 3) + (2 \times 1) + (2 \times 1) + 1)$$
$$= 34.5 \text{ Area Unit}$$

If we assume that any paper of this author gains 1 more citation value, then the AUC will be 39 Area Unit. Therefore, the *MIArea-index* (Which equals to AUC) is very sensitive for citation values. Moreover, the publication will not be neglected even if it has at least one citation. For normalization issue, we can use a factor such as α, where α value may be = {0.1, 0.2, 0.25….0.5…1}. Identifying the value of α needs further studies and extensive experiments. In this study we assume that $\alpha = 1$.

3.4. Example

An example showing the calculations of the two proposed novel indices (*MIExp*- and *MIArea - Index*) is presented in this section. Two authors were used in this example, where each one of these two authors has 10 research papers, which differently cited as shown in Table 2.

Table 2. Two authors scientific outputs with their citations

Order		1	2	3	4	5	6	7	8	9	10	Total
$Author_1$	Publications	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	10
	Citations	8	8	7	7	5	2	1	1	1	0	40
$Author_2$	Publications	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	10
	Citations	22	7	6	5	5	2	1	0	0	0	48

The calculations of *MIExp-Index* are presented first in this section using the data presented in Table 2.

MIExp-Index for the first author ($Author_1$) equals to:

$$MIExp - Index = e^{\log_{10}(40+1)} = 5.017$$

MIExp-Index for the second author ($Author_2$) equals to:

$$MIExp - Index = e^{\log_{10}(48+1)} = 5.421$$

Next the calculations of *MIArea-Index* are presented using the data presented in Table 2.

MIArea-Index for the first author ($Author_1$) equals to:

$$MIArea - Index = \left(\frac{10 - 1}{2 \times 9}\right)\begin{pmatrix} 8 + (2 \times 8) + (2 \times 7) + (2 \times 7) + (2 \times 5) \\ + (2 \times 2) + (2 \times 1) + (2 \times 1) + 1 \end{pmatrix}$$

$$MIArea - Index = 35.5 \text{ Area unit.}$$

MIArea-Index for the second author ($Author_2$) equals to:

$$MIArea - Index = \left(\frac{10 - 1}{2 \times 9}\right)\begin{pmatrix} 22 + (2 \times 7) + (2 \times 6) + (2 \times 5) + (2 \times 5) \\ + (2 \times 2) + 1 \end{pmatrix}$$

$$MIArea - Index = 36.5 \text{ Area unit.}$$

4. Evaluation

In order to assess the proposed two indices, and show how they may overcome some of *H*- and *G-index* weaknesses; mainly, the "Sensitivity" and "Comprehensiveness" problems, we will conduct two simple but significant experiments. In the first experiment, we will use artificial examples of authors, publications and citations. In the second one, we will show how the proposed indices behave with real data in comparison with *H*- and *G*-indices.

4.1. First Experiment

The assessment process is conducted using four different scenarios listed in Table 3. We believe that these scenarios simulate realistic scenarios of researchers.

The first scenario (FF) includes those authors who can be classified as newcomers or beginners, where the publication volume and its influence are somewhat small. The second scenario (FM) represents those authors who are usually interested in producing few core papers, large influence, regardless of the production volume. The authors in scenario three (MF) focus on volume or number of publications with possibly few number of citations, at least at the beginning. For the last one (MM) represents senior authors with large number of publications and citations. We can say that in any research field researchers may fall largely in one of those four classes or categories.

Table 3. First Experiment of Four Simulated Cases of Researchers

Scenario ID	# of Publications	# of Citations
1	Few	Few
2	Few	Many
3	Many	Few
4	Many	Many

For each scenario, we will simulate the increase in citations value during the publications life time by applying three of common citation increments. Table 4 shows the increment scenarios. We use these scenarios in order to examine the sensitivity of the two indices under study in dealing with citation increments with different scenarios of publication and citation values.

Table 4.Citation Increments

Scenario-ID	Description
1	Add citations to the first publication where they are in a descending order.
2	Add citations to others except the first one, where they are in a descending order.
3	Add a new publication with only one citation value.

4.1.1. The First Suggested Scenario

This scenario simulates new researchers who have few numbers of publications and their publications have a few numbers of citations. Table 5 shows four authors who are classified as new researchers.

Table 5. New Researchers

$Author_1$	**Pub.**	**1**	**2**	**3**	**4**	**$Author_2$**	**Pub.**	**1**	**2**	**3**	**4**	**5**
	Cit.	2	1	0	0		Cit.	3	2	1	1	0
$Author_3$	Pub.	1	2	3	X	**$Author_4$**	Pub.	1	2	3	X	X
	Cit.	3	1	0	X		Cit.	8	4	0	X	X

The index values of these four cases are presented in table 6.

Table 6. *MIExp* and *MIArea* values for the four new researchers

	Total Publications	Total Citations	*H-index*	*G-index*	*MIExp*-Index	*MIArea*-Index
$Author_1$	4	3	1	1	1.826	1.5
$Author_2$	5	7	2	2	2.467	5
$Author_3$	3	4	1	2	2	2
$Author_4$	3	12	2	3	3.046	6

Table 6 shows a considerable difference in *MIExp and MIArea- index* values for these four authors.

MIExp and *MIArea* measure the impact of authors ($Author_1$ and $Author_3$), and ($Author_2$ and $Author_4$) in different aspects in comparison with *h* values that were almost the same for the four cases. Variations can be also noticed between proposed indices and *G-index* values. Even though the assessments are somewhat close to each other, but the comparison between $Author_2$ and $Author_3$ with respect to values of *G-index* and the proposed indices may indicate that they are not consistent.

4.1.2. The Second Suggested Scenario

In this scenario, we concentrate on those authors who have few publications and many citations. These publications can be classified as core papers in their field. Table 7 shows artificial examples of this scenario.

There are no differences between the *H-index* values shown in table 7. They have the same *h* –level (*i.e.* 10). In table 7 the total number of citations of author A23 is 8,444 and the total number of citations of author A24 is 34,033. The difference between these two values is large (25,589), and the *H-index* value determines the most influential author. In this case the number of papers is ignored. The *G-index* and our two proposed indices (*MIExp* and *MIArea*) show different citation values for the same scenario.

After applying the first increment scenario, the *H-index* and the *G-index* did not consider the new values of citations into account. On the other hand, the *MIExp-index* and *MIArea-index* values are affected by citation changes. Both of *MIExp-index* and *MIArea-index* are sensitive to citation changes.

There are no criteria to increase the values of citations in the second scenario. The results show that both of *H-* and *G*-indices' values have not changed, while the *MIExp* and *MIArea* values were impacted by the new citations.

Table 7. Highly Cited Authors with Few Publications

Original Values of Citations

P.ID	1	2	3	4	5	6	7	8	9	10	TOT	H	G	MIExp	MIArea
A21	7500	5530	2450	1855	1103	900	295	211	101	55	20000	10	141	73.777	16222.5
A22	4300	1555	1200	950	790	750	455	X	X	X	10000	7	100	54.601	7622.5
A23	2800	2440	610	601	500	455	300	288	250	200	8444	10	91	50.734	6944
A24	9000	8850	7000	6500	1000	705	595	233	100	50	34033	10	184	92.936	29508

After applying first citation increament scenario(Add 10 citations to the first publication)

P.ID	1	2	3	4	5	6	7	8	9	10	TOT	H	G	MIExp	MIArea
A21	7510	5530	2450	1855	1103	900	295	211	101	55	20010	10	141	73.793	16227.5
A22	4310	1555	1200	950	790	750	455	X	X	X	10010	7	100	54.624	7627.5
A23	2810	2440	610	601	500	455	300	288	250	200	8454	10	91	50.760	6949
A24	9010	8850	7000	6500	1000	705	595	233	100	50	34043	10	184	92.948	29513

After applying second citation increament scenario(Add 10 citations to any publication except the first one)

P.ID	1	2	3	4	5	6	7	8	9	10	TOT	H	G	MIExp	MIArea
A21	7500	5530	2450	1855	1103	*910*	295	211	101	55	20010	10	141	73.793	16227.5
A22	4300	1555	*1210*	950	790	750	455	X	X	X	10010	7	100	54.624	7627.5
A23	2800	2440	610	601	*510*	455	300	288	250	200	8454	10	91	50.760	6949
A24	9000	8850	7000	6500	1000	705	595	233	100	*60*	34043	10	184	92.948	29513

After applying third citation increament scenario (Add new publication with one citation)

P.ID	1	2	3	4	5	6	7	8	9	10	11	TOT	H	G	MIExp	MIArea
A21	7500	5530	2450	1855	1103	900	295	211	101	55	1	20001	10	141	73.779	16251
A22	4300	1555	1200	950	790	750	455	1	X	X	X	10001	7	100	54.603	7851
A23	2800	2440	610	601	500	455	300	288	250	200	1	8445	10	91	50.737	7045
A24	9000	8850	7000	6500	1000	705	595	233	100	50	1	34034	10	184	92.937	29534

The third increment case shown in table 8 shows how the new two proposed indices (*MIExp* and *MIArea*) behave with the extension of the publication list. Values of *MIExp* and *MIArea* indices are affected and changed when one publication with one citation is added. On the other hand, the index values of *H*- and *G*-indices have not changes. In Table 8 the values of *MIArea-index* of scenario 3 are completely different from the values of *MIArea-index* of scenario 1, where the values of *MIArea-index* of scenario 3 are higher than the values of *MIArea-index* of scenario 1 using the trapezoidal rule to compute those values which represent area under the curve. In addition, these values can be used to predict the number of citations for different articles published by the same author. Table 8 shows the second scenario with the three increment cases. The results show that the proposed indices (*MIExp* and *MIArea*) are more sensitive in response to the changes in the total number of citations if those indices were compared with traditional *H*- and *G*- indices.

4.1.3. The Third Suggested Scenario

Table 8 shows the three cases and the results of the third suggested scenario of publications and citation values, and the effect of the three increment cases on the evaluated indices. As shown in Table 8, the *G-index* was slightly affected in response to the citation changes in comparison with *H-index* where the citation increment values made no impact on the *H-index*. However, the sensitivity degree of our proposed indices (*MIExp* and *MIArea*) is noticeable in response to the increment in the number of citations and publications presented in Table 8.

4.1.4. The Fourth Suggested Scenario

Table 9 shows artificial examples of publications, citations and the behavior of the indices understudy. The 100 new citations added to the number of citations in the first and second increment cases have no effect on the values of *H*- and *G*- indices. Table 9 shows also clearly the sensitivity of the new proposed indices (*MIExp* and *MIArea*) to those new 100 citations added to the original citations. Adding one new publication with one citation to the current publication list for each author affects also the values of *MIExp* and *MIArea* indices. Therefore adding one citation to thousands of citations for a long list of publications can positively affect the values of proposed indices whereas it has no effect on the values of *H*- and *G*-indices.

Table 8. Cases of a large number of publications and low citation

Original Citations Values

P.ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	TOT	H	G	MIExp	MIArea
A31	8	8	8	8	8	7	7	7	7	7	7	7	7	6	6	6	6	6	6	5	5	5	5	4	4	4	3	3	3	3	2	2	2	2	1	1	1	1	1	1	190	7	7	9.787	185.5
A32	10	9	9	9	9	9	9	9	8	8	7	7	7	4	4	4	3	3	3	3	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	158	8	9	9.038	152.5
A33	9	9	9	8	8	8	8	8	8	8	8	7	7	7	7	7	7	7	5	5	5	5	5	5	5	5	5	5	5	4	4	4	2	2	1	1	1	1	1	1	217	8	8	10.365	212
A34	11	6	6	6	5	5	5	4	4	4	4	4	4	4	4	3	3	3	3	3	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	118	5	6	7.969	112

The results after first citations increment scenario (Add 5 citations to the first publication)

P.ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	TOT	H	G	MIExp	MIArea
A31	13	8	8	8	8	7	7	7	7	7	7	7	7	6	6	6	6	6	6	5	5	5	5	4	4	4	3	3	3	3	2	2	2	2	1	1	1	1	1	1	195	7	8	9.897	188
A32	15	9	9	9	9	9	9	9	8	8	7	7	7	4	4	4	3	3	3	3	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	163	8	9	9.160	155
A33	14	9	9	8	8	8	8	8	8	8	8	7	7	7	7	7	7	7	5	5	5	5	5	5	5	5	5	5	5	4	4	4	2	2	1	1	1	1	1	1	222	8	8	10.468	214.5
A34	16	6	6	6	5	5	5	4	4	4	4	4	4	4	4	3	3	3	3	3	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	123	5	7	8.113	114.5

The results after second citations increment scenario (Add 5 citations to any publication except the first one

P.ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	TOT	H	G	MIExp	MIArea
A31	13	8	8	8	8	7	7	7	7	7	7	7	7	6	6	6	6	6	6	5	5	5	5	4	4	4	3	3	3	3	2	2	2	2	1	1	1	1	1	1	195	7	8	9.897	188
A32	14	10	9	9	9	9	9	9	8	8	7	7	7	4	4	4	3	3	3	3	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	163	8	9	9.160	155.5
A33	14	9	9	8	8	8	8	8	8	8	8	7	7	7	7	7	7	7	5	5	5	5	5	5	5	5	5	5	5	4	4	4	2	2	1	1	1	1	1	1	222	8	8	10.468	214.5
A34	11	11	6	6	5	5	5	4	4	4	4	4	4	4	4	3	3	3	3	3	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	123	5	7	8.113	117

The results after Third citations increment scenario (Add new publication with one citation)

P.ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	TOT	H	G	MIExp	MIArea
A31	8	8	8	8	8	7	7	7	7	7	7	7	7	6	6	6	6	6	6	5	5	5	5	4	4	4	3	3	3	3	2	2	2	2	1	1	1	1	1	1	1	191	7	7	9.809	186.5
A32	10	9	9	9	9	9	9	9	8	8	7	7	7	4	4	4	3	3	3	3	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	159	8	9	9.062	153.5
A33	9	9	9	8	8	8	8	8	8	8	8	7	7	7	7	7	7	7	5	5	5	5	5	5	5	5	5	5	5	4	4	4	2	2	1	1	1	1	1	1	1	218	8	8	10.386	213
A34	11	6	6	6	5	5	5	4	4	4	4	4	4	4	4	3	3	3	3	3	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	119	5	6	7.998	113

Table 9. Cases of many publications and citations

Original Citations Values

P.ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	TOT	H	G	MIExp	MIArea
A41	6800	6201	5500	4003	3810	3211	3002	2850	2550	2500	1660	1601	1500	1500	1000	905	900	850	850	850	850	770	760	750	750	740	700	688	670	620	600	401	305	300	210	150	100	70	60	50	61587	40	248	120.241	58162
A42	4000	4000	3550	3500	3250	3150	3100	3002	3000	1500	1440	1300	1200	1150	1100	1050	1000	950	900	550	500	500	420	400	390	350	200	200	190	150	140	110	90	88	70	68	60	60	55	55	46788	40	216	106.713	44760.5
A43	7000	6690	3500	3200	3100	1250	1200	1200	1180	1100	1090	1077	1005	1001	999	980	970	960	922	911	902	901	887	850	840	822	810	805	800	705	122	105	100	98	91	88	88	75	70	63	48557	40	220	108.447	45025.5
A44	11000	9950	8500	7660	7040	6441	5000	2400	2060	2005	1991	1905	1807	1000	1000	1000	1000	991	888	777	770	701	701	688	646	638	600	460	305	300	201	192	155	106	101	80	70	60	60	59	81308	40	285	135.658	75778.5

The results after first citations increment scenario (Add 100 citations to the first publication

P.ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	TOT	H	G	MIExp	MIArea
A41	6900	6201	5500	4003	3810	3211	3002	2850	2550	2500	1660	1601	1500	1500	1000	905	900	850	850	850	850	770	760	750	750	740	700	688	670	620	600	401	305	300	210	150	100	70	60	50	61687	40	248	120.326	58212
A42	4100	4000	3550	3500	3250	3150	3100	3002	3000	1500	1440	1300	1200	1150	1100	1050	1000	950	900	550	500	500	420	400	390	350	200	200	190	150	140	110	90	88	70	68	60	60	55	55	46888	40	216	106.812	44810.5
A43	7100	6690	3500	3200	3100	1250	1200	1200	1180	1100	1090	1077	1005	1001	999	980	970	960	922	911	902	901	887	850	840	822	810	805	800	705	122	105	100	98	91	88	88	75	70	63	48657	40	220	108.544	45075.5
A44	11100	9950	8500	7660	7040	6441	5000	2400	2060	2005	1991	1905	1807	1000	1000	1000	1000	991	888	777	770	701	701	688	646	638	600	460	305	300	201	192	155	106	101	80	70	60	60	59	81408	40	285	135.731	75828.5

The results after second citations increment scenario (Add 100 citations to any publication except the first one

P.ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	TOT	H	G	MIExp	MIArea
A41	6800	6201	5500	4003	3810	3211	3002	2850	2550	2500	1760	1601	1500	1500	1000	905	900	850	850	850	850	770	760	750	750	740	700	688	670	620	600	401	305	300	210	150	100	70	60	50	61687	40	248	120.326	58262
A42	4000	4000	3550	3500	3250	3150	3100	3002	3000	1500	1440	1300	1200	1150	1100	1050	1000	950	900	650	500	500	420	400	390	350	200	200	190	150	140	110	90	88	70	68	60	60	55	55	46888	40	216	106.812	44860.5
A43	7000	6690	3500	3300	3100	1250	1200	1200	1180	1100	1090	1077	1005	1001	999	980	970	960	922	911	902	901	887	850	840	822	810	805	800	705	122	105	100	98	91	88	88	75	70	63	48657	40	220	108.544	45125.5
A44	11000	9950	8500	7660	7040	6441	5000	2400	2060	2005	1991	1905	1807	1000	1000	1000	1000	991	888	777	770	701	701	688	646	638	600	560	305	300	201	192	155	106	101	80	70	60	60	59	81408	40	285	135.731	75878.5

The results after Third citations increment scenario (Add new publication with one citation)

P.ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	TOT	H	G	MIExp	MIArea
A41	6800	6201	5500	4003	3810	3211	3002	2850	2550	2500	1660	1601	1500	1500	1000	905	900	850	850	850	850	770	760	750	750	740	700	688	670	620	600	401	305	300	210	150	100	70	60	50	1	61588	40	248	120.242	58187.5
A42	4000	4000	3550	3500	3250	3150	3100	3002	3000	1500	1440	1300	1200	1150	1100	1050	1000	950	900	550	500	500	420	400	390	350	200	200	190	150	140	110	90	88	70	68	60	60	55	55	1	46789	40	216	106.714	44788.5
A43	7000	6690	3500	3200	3100	1250	1200	1200	1180	1100	1090	1077	1005	1001	999	980	970	960	922	911	902	901	887	850	840	822	810	805	800	705	122	105	100	98	91	88	88	75	70	63	1	48558	40	220	108.448	45057.5
A44	11000	9950	8500	7660	7040	6441	5000	2400	2060	2005	1991	1905	1807	1000	1000	1000	1000	991	888	777	770	701	701	688	646	638	600	460	305	300	201	192	155	106	101	80	70	60	60	59	1	81309	40	285	135.659	75908

4.2. Applying MIExp and MIArea indices on real data

Table 10 shows the first 15 authors in the field of software engineering ordered as they appear in the website of Microsoft academic research (http://academic.research.microsoft.com). Microsoft academic uses special criteria to rank the authors. Their ranking policy depends mainly on *H-index* and other factors such as: distribution of publications and citations of author works. We applied and examined the four indices under study on these top ranked authors. The *G-index* behaves in a more fairly manner in comparison with *H-index* to assess the authors' publications. For example, the use of *G-index* shows that the studies of Edmund Clarke are more influential than the studies of Victor Basili. It is clear that the use of *H-index* in such situations leads to unfair results, since it is strange for the index to rank the author with: 453 publications and 45,375 total citations of Edmund Clarke in a rank after an author with: 603 publications and 20,367 total citations of Victor Basili.

Table 10. Microsoft Academic - Top 15 Researchers in Software Engineering

Author ID	Author	Total Publication	Total Citations	*H*-Index	*G*-Index	*MIExp*	*MIArea*
1	Victor Basili	603	20367	68	136	74.362	19727
2	Edmund M. Clarke	453	45375	66	212	105.301	41060
3	Mary Jean Harrold	238	12144	59	108	59.406	11883
4	Barry Boehm	761	35466	67	184	94.616	31837
5	Amir Pnueli	446	34084	85	181	92.997	32159
6	David Garlan	472	24698	65	154	80.857	22461.5
7	Jeff Kramer	489	18526	59	132	71.365	17875
8	Lionel C. Briand	217	10965	50	103	56.829	10368
9	Thomas A. Henzinger	511	31236	80	171	89.539	30435
10	David L. Dill	392	26711	63	162	83.655	24126.5
11	Jeff Offutt	199	4597	30	66	38.961	4396
12	Gregg Rothermel	258	9509	52	95	53.420	9248
13	Thomas W. Reps	380	15413	62	121	65.885	14750.5
14	Gail Murphy	583	14488	62	109	64.138	13837
15	David Lorge Parnas	203	14796	38	121	64.727	12421.5

Table 11 shows the rank of each author ordered based on the indices under study. The *G-index*, *MIexp* and *MIArea* have the same point of view on authors ranking, and it is clear that they are strongly correlated. However, the sensitivity property of the proposed indices could make them more effective indicators on authors' impacts.

Table 11. Fifteen Authors Ordered according to their Index Values

	Rank														
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
	Author ID														
Based on *H*-Index	5	9	1	4	2	6	10	13	14	3	7	12	8	15	11
Based on *G*-Index	2	4	5	9	10	6	1	7	13	15	14	3	8	12	11
Based on *MIExp*-Index	2	4	5	9	10	6	1	7	13	15	14	3	8	12	11
Based on *MIArea*-Index	2	5	4	9	10	6	1	7	13	14	15	3	8	12	11

There are different points of views about the production volume and its influence. It is not easy to say which scenario is the best? Those authors who have few publications and many citations, or those who have many publication and few citations? We can say that production volume is an important factor, since productive researchers tend to share their thoughts with others in the research community, even though their publications may not get a large number of citations. On the other hand, productivity can be seen as a rival for publication quality (if we consider citation number is the quality metric).

Regardless of these points of views, we showed in this work through the use of a number of artificial examples and real data that *MIExp* and *MIArea* indices can be more balanced, comprehensive and sensitive to evaluation publications if compared with other citation indices. However, the values produced by *MIArea* are too high and it is not easy to deal with *MIArea* in its current form. We proposed the idea of using Area under curve, but we should conduct more experiments in order to normalize the *MIArea* index values while preserving its sensitivity.

5. Conclusions

All citation indices, including proposed ones in this paper, depend mainly and radically on the number of citations that each research paper received. Currently, the *H*- and *G*-indices are the most common indices used to assess the authors and their publications. However, there are many limitations related to indices' calculations. Most of those limitations are related to the sensitivity of the indices and the coverage or comprehensives of the index to cover different publication or author aspects in addition to the pure numbers. In this paper, we proposed and evaluated new indices: *MIExp* and *MIArea* which provide simple enhancements to the behaviors of: *H*- and *G*-indices using the exponential function and the concept of area under curve. The evaluation process showed that the new indices can solve the problems and behave in a more effective manner than traditional indices. More experiments and in depth studies are needed to normalize the *MIArea* values as they are too large to deal with. On the other hand, it is critical to keep the sensitivity of those indices that we showed, where each index can response to any addition in publications or citations.

References

[1] M. Franceschet, "A Cluster Analysis of Scholar and Journal Bibliometric Indicators", Journal of the American Society for Information Science and Technology, vol. 60, no. 10, **(2009),** pp. 1950–1964.
[2] J. E. Hirsch, "An index to quantify an individual's scientific research output", Proceedings of the National Academy of Science of the USA, vol. 102, no. 46, **(2005)**, pp. 16569–16572.
[3] L. Bornmann and H.-D. Daniel, "Does the H-index for ranking of scientists really work? Scientometrics, vol. 65, no. 3, **(2005)**, pp. 391–392.
[4] W. Glänzel, "On the H-index – a mathematical approach to a new measure of publication activity and citation impact", Scientometrics, vol. 67, no. 2, **(2006b)**, pp. 315–321.
[5] B. Cronin and L. Meho, "Using the H-index to rank influential information scientists", Journal of the American Society for Information Science and Technology", vol. 57, no. 9, **(2006)**, pp. 1275–1278.
[6] L. Bornmann, and H.-D. Daniel, "What do we know about the h index?" Journal of the American Society for Information Science and Technology, vol. 58, **(2007)**, pp. 1381–1385.
[7] T. Braun, W. Glänzel, and A. Schubert, "A Hirsch-type index for journals", Scientometrics, vol. 69, no. 1, **(2006)**, pp. 169–173.
[8] J. E. Hirsch, "Does the H-index have predictive power?", Proceedings of the National Academy of Sciences of the United States of America, vol. 104, no. 49, **(2007)**, pp. 19193–19198.
[9] W. Glänzel, "On the opportunities and limitations of the *H*-index", Science Focus, vol. 1, **(2006a)**, pp. 10–11.
[10] B. Jin, L. Liang, R. Rousseau and L. Egghe, "The R- and ARindices: Complementing the H-index", Chinese Science Bulletin, vol. 52, no. 6, **(2007),** pp. 855–863.
[11] L. Egghe, "How to improve the H-index?" The Scientist, vol. 20, no. 3, **(2006a),** p. 14.

[12] L. Egghe, "Theory and practise of the *G-index*", Scientometrics, vol. 69, no. 1, **(2006b)**, pp. 131–152.
[13] L. Egghe, "An improvement of the H-index : The G-index. ISSINewsletter, vol. 2, **(2006c)**, pp. 8–9.
[14] M. Schreiber, C. C. Malesios, and S. Psarakis, "Exploratory factor analysis for the Hirsch index 17 h-type variants, and some traditional bibliometric indicators", Journal of Informetrics, vol. 6, no. 3, **(2012)**, pp. 347–358.
[15] L. Bornmann, R. Mutz, S. E. Hug, and H.-D. Daniel, "A multilevelmeta-analysis of studies reporting correlations between the h index and 37 different h index variants", Journal of Informetrics, vol. 5, no. 3, **(2011)**, pp. 346–359.
[16] SCI2S (Soft Computing and Intelligent Information Systems) - SCI2S Thematic Public Websites - H-index and Variants, **(2013)**, Retrieved July 4, 2013,http://sci2s.ugr.es/hindex/.
[17] P. D. Batista, M. G. Campiteli, and O. Kinouchi, "Is it possible to compare researchers with different scientific interests?", Scientometrics, vol. 68, no. 1, **(2006)**, pp. 179–189.
[18] M. Schreiber, "To share the fame in a fair way, h_m modifies h formulti-authored manuscripts", New Journal of Physics, vol. 10, **(2008)**, pp. 040201.
[19] M. Schreiber, "Fractionalized Counting of Publications for the G-index", Journal of the American Society for Information Science and Technology, vol. 60, no. 10, **(2009)**, pp. 2145–2150.
[20] M. Schreiber, "A case study of the modified g index: Counting multi-author publications fractionally", Journal of Informetrics, vol. 4, **(2010a)**, pp. 636–643.
[21] M. Schreiber, "How to modify the G-index for multi-authored manuscripts", Journal of Informetrics, 4, **(2010b)**, pp. 42–54.
[22] B. Jin, "The AR-index: complementing the *H-index*", ISSI Newsletter, vol. 3, no. 1, **(2007)**, p. 6.
[23] M. Kosmulski, "A new Hirsch-type index saves time and works equally well as the original H-index", ISSI Newsletter, vol. 2, no. 3, **(2006)**, pp. 4–6.
[24] Q. Wu, "The w-index: A measure to assess scientific impact by focusing on widely cited papers", Journal of the American Society for Information Science and Technology, vol. 61, no. 3, **(2010)**, pp. 609–614.
[25] B. H. Jin, "H-index : an evaluation indicator proposed by scientist", Science Focus, vol. 1, no. 1, **(2006)**, pp. 8-9. (In Chinese)
[26] R. Rousseau, "New developments related to the Hirsch index", Science Focus, vol. 1, no. 4, **(2006)**, pp. 23–25.
[27] M. Maabreh, and I. Alsmadi, "A Survey of Impact and Citation Indices: Limitations and Issues", International Journal of Advanced Science and Technology, vol. 40, **(2012)**, pp. 35-54.
[28] A. Sidiropoulos, D. Katsaros, and Y. Manolopoulos, "Generalized Hirsch H-index for disclosing latent facts in citation networks", Scientometrics, vol. 72, no. 2, **(2007)**, pp. 253–280.
[29] L. Bornmann, R. Mutz, and H.-D. Daniel, "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, vol. 59, no. 5, **(2008)**, pp. 830–837.
[30] L. Egghe, and R. Rousseau, "An H-index weighted by citation impact", Information Processing and Management, vol. 44, **(2008)**, pp. 770–780.
[31] S. Alonso, F. J. Cabrerizo, E. Herrera-Viedma, and F. Herrera, "hG-index: a new index to characterize the scientific output of researchers based on the h- and g-indices", Scientometrics, vol. 82, no. 2, **(2010),** pp. 391–400.
[32] F. J. Cabrerizo, S. Alonso, E. Herrera-Viedma, and F. Herrera, "q2-Index: Quantitative and qualitative evaluation based on the number and impact of papers in the Hirsch core", Journal of Informetrics, vol. 4, no. 1, **(2010)**, pp. 23–28.
[33] 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, **(2009)**, pp. 431-439.

Authors

Majdi Ahmad Maabreh, he is an Oracle applications developer and DBA at Hashemite University, Jordan. He received his bachelor's degree in computer science from Yarmouk University in Jordan in July 2004. Since 2004, he joined the team of programming and development of database systems at the Hashemite University. He received his master degree in Computer Information Systems from Yarmouk University in 2012. The most topics interest includes

Data Mining, Information Retrieval as well as Databases Technology.

Izzat Mahmoud Alsmadi, he is an assistant professor in the department of computer information systems at Yarmouk University in Jordan. He obtained his Ph.D degree in software engineering from NDSU (USA), his second master in software engineering from NDSU (USA) and his first master in CIS from University of Phoenix (USA). He had B.sc degree in telecommunication engineering from Mutah University in Jordan. He has several published books, journals and conference articles largely in software engineering and information retrieval fields.

Mohammed Naji Alkabi, he obtained his Ph.D. degree in Mathematics from the University of Lodz/Poland (2001), his master's degree in Computer Science from the University of Baghdad/Iraq (1989), and his bachelor degree in statistics from the University of Baghdad/Iraq (1981). Mohammed Naji AL-Kabi is an assistant Professor in the Faculty of Sciences and IT, at Zarqa University. Prior to joining Zarqa University, he worked many years at Yarmouk University in Jordan, Nahrain University and Mustanserya University in Iraq. He also worked as a part time lecturer at Jordan University of Science and Technology, and Sunderland university. AL-Kabi's research interests include Information Retrieval, Web search engines, Data Mining, Sentiment analysis and opinion mining, Natural Language Processing and Software Engineering. He is the author of more than 68 peer reviewed articles in these topics. His teaching interests focus on information retrieval, Web programming, data mining, DBMS (ORACLE, MYSQL & MS Access).