Procedia Computer Science
Volume 29, 2014, Pages 465–477
ICCS 2014. 14th International Conference on Computational Science

The p-index: Ranking Scientists using Network Dynamics

U. Senanayake[1], M. Piraveenan[1] and A. Y. Zomaya[2]
[1]Center for Complex Systems Research, The University of Sydney, Australia.
[2]Centre for Distributed and High Performance Computing, The University of Sydney, Australia.
usen8682@uni.sydney.edu.au,
mahendrarajah.piraveenan@sydney.edu.au, albert.zomaya@sydney.edu.au

Abstract

The indices currently used by scholarly databases, such as Google scholar, to rank scientists, do not attach weights to the citations. Neither is the underlying network structure of citations considered in computing these metrics. This results in scientists cited by well-recognized journals not being rewarded, and may lead to potential misuse if documents are created purely to cite others. In this paper we introduce a new ranking metric, the p-index (pagerank-index), which is computed from the underlying citation network of papers, and uses the pagerank algorithm in its computation. The index is a percentile score, and can potentially be implemented in public databases such as Google scholar, and can be applied at many levels of abstraction. We demonstrate that the metric aids in fairer ranking of scientists compared to h-index and its variants. We do this by simulating a realistic model of the evolution of citation and collaboration networks in a particular field, and comparing h-index and p-index of scientists under a number of scenarios. Our results show that the p-index is immune to author behaviors that can result in artificially bloated h-index values.

1 Introduction

The h-index is one of the three metrics used in Google scholar (the others are the number of citations, and the i10-index) to quantify the impact of a scientist's research. The h-index is arguably the most sophisticated among the three, as it accounts for both the quality and the quantity of a scientists' research publications, to some extent. The h-index h was defined by J.E Hirsch (Hirsch 2005) as the number of papers with citation number $\geq h$. This implies that if the scientist has N_T number of papers, (N_T-h) number of papers have no more than h citations. Previous to the h-index, the common measures used to measure a researcher's impact were (i) number of citations (ii) number of papers (iii) and citations per paper. These simplistic measures all had readily evident drawbacks. For example, the number of papers does not take into account the quality or impact of the papers, while the number of citations may

Selection and peer-review under responsibility of the Scientific Programme Committee of ICCS 2014

doi:10.1016/j.procs.2014.05.042

be inflated by a few papers with a great number of citations co-authored by many authors, and the citations per paper measure rewards low productivity. Hirsch argued that his h-index addresses these drawbacks by drawing information from both number of papers and citation count.

The h-index has been widely accepted by the scientific community (Bornmann and Daniel 2007), and there is evidence to suggest that successful scientists comparatively have higher h-index values, at least when 'success' can be measured by employability and grant success (Bornmann, Mutz et al. 2008). However, there has been considerable debate about the merits of h-index, many of its alleged shortcomings have been pointed out (Egghe 2006, Ghoshal and Barabasi 2011), and many variants and alternatives have been suggested (Hirsch 2005, Jin, Liang et al. 2007, Leicht, Clarkson et al. 2007). For instance, h-index depends on the time span of the career of a scientist (Hirsch 2005, Leicht, Clarkson et al. 2007). As such, those scientists who may be brilliant but are in the early stages of their career are reflected unfavorably. Hirsch suggested using a time compensated measure to overcome this restriction. Furthermore, L. Egghe(Egghe 2006) argued that, according to the definition of h-index, once a paper belongs to the top h class (among the papers ranked h or higher), it becomes unimportant whether these papers are further cited or not. Scientists who have high-impact papers are penalized, and this is evidenced by the difference between the lower bound of citations, h^2, proposed by the h-index, and the actual number of citations. This had been observed by Hirsch himself (Hirsch 2005). Egghe therefore suggested the g-index, and defined it as the highest number g of papers that together received g or more citations. Similarly, the R-index and AR index introduced by Jin et al (Jin, Liang et al. 2007) measures the citation intensity of the h-core, also considering the age of the papers. They suggested that these measures could be used together with the h-index to complement each other. There are many other variants and alternative indices introduced, as summarized by (Raan 2006, Waltman, Costas et al. 2012).

A review of the existing variants of h-index by Bornmann et al (Bornmann, Mutz et al. 2008) suggested that all these variants fall into two types: those that describe the size of the h-core, and those that describe the impact of the h-core. The first type redefines the size of the h-core and typically makes it larger (such as g-index), or add additional information about it (such as the r-index) while the second type analyses the citation count of the h-core, negating the disadvantage attributed to scientists who have a few high-impact papers. However, a fundamental issue not even touched-let alone addressed-by all these measures is that they still treat all citations equally. Yet, it is clear that a citation by a paper from a highly regarded journal, such as *Nature*, should be treated differently from a citation by a workshop paper or a technical report. If this does not happen, locally famous authors whose research does not have global impact but gets cited by their colleagues in their country or research circle can get rewarded. Moreover, if all citations are treated equally, then 'massaging' the h-index becomes possible by publishing papers, in whatever available forum, purely with the intention of citing other papers by the same group of authors or collaborators. This process has the danger of encouraging the creation of a huge body of academic papers which nobody reads, let alone utilizes, for further research or application.

In this paper, therefore, we propose to utilize the underlying network of citations to evolve an index which is dynamic and rewards citation by impactful documents. Our goal is to formulate and test a fairer metric which rewards true excellence by giving higher weight to citations from documents with higher academic standing. Therefore, we introduce the p-index, which is designed to address the drawbacks of existing indices by utilizing the underlying citation network dynamics. The index is calculated by running the well-known pagerank algorithm (Page, Brin et al. 1998) on the evolving citation network to give scores to papers. This has been explored before by Walker et al. and Chen et al (Chen, Xie et al. 2007, Walker, Xie et al. 2007) in order to find 'scientific-gems' of publications which are less apparent if number of citations is used as the ranking metric. They have used a modified algorithm called 'citerank' in order to rank scientific publications and this approach is summarized with merits and pitfalls by Maslov et al (Maslov and Redner 2009). Our effort in this research differs such that we are trying to rank scientists as opposed to ranking publications and ranking publications is only an intermediary step. We use the original form of pagerank algorithm to come up with pagerank values for each publication and the score of an author is calculated as a weighted sum of the scores of the papers he/she has written. By employing a realistic simulation system which synthesizes the evolution process of citation and collaboration networks in an emerging field, we demonstrate that the p-index is fairer than h-index in many scenarios in which the authors may be able to massage or manipulate their h-index. In particular, we show that while it is possible to massage h-index by publishing papers in low-impact forums which are used simply to cite other papers of the same group of authors, this is largely impossible with p-index. We also show that while h-index rewards collaboration and authors who focus on quantity, the p-index is much more balanced and equally rewards individual brilliance and quality of papers.

It could be asked, whether we could not have simply weighed citations by using the impact factor of journals in which the citing documents are published. However, careful analysis reveals this method has many pitfalls. There is considerable debate about the dependability of the impact factors of journals (van Nierop 2009), and we cannot therefore consider that impact factors are sacrosanct. In any case, conference papers and relatively new journals do not have impact factors. There is no guarantee either, that a paper which is published in a journal with high impact factor will itself have high impact. Our intention is to reward the quality of the citing document, not the quality of journal in which it is published. Even if we count the number of citations received by the citing document, and weight the citation accordingly in a dynamic manner, it is possible that these weights could be manipulated again by less impactful documents: thus the influence of such documents only becomes indirect, and does not go away. Our intention therefore, is to introduce an index which has 'infinite' levels of feedback, and the impact of citing documents is factored in as much as possible. This is portrayed in Fig. 1. From the figure it can be seen that the number of citations a document receives could be not always proportional to its impact factor, and we must reward actual numbers of citations at all levels, and not the impact factors. This could only be done by fully utilizing the overall dynamic citation network, and by an elegant algorithm such as pagerank, which has had proven success in measuring the impact of such feedback loops in the Google search engine.

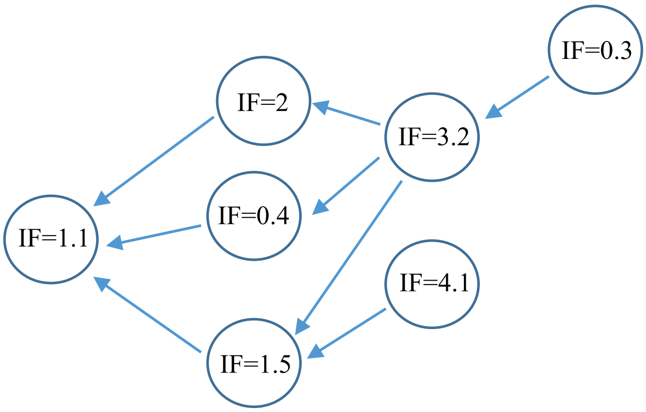

Figure 1: Multiple levels of citations; IF stands for the impact factor of each node (paper) and an arrow represents a citation from the node at the tail end to the node at the head of the arrow.

The rest of this paper is organized as follows. The second section will introduce the methodology while the third section will explain the simulation system we utilized to verify the utility of the p-index. The fourth section will present the results obtained from the experiments. The fifth section will highlight the impact of our findings and present the conclusions. The sixth and final section will discuss possible future work that can be undertaken to further validate the new measure.

2 Methodology

We implement a version of the pagerank algorithm (Page, Brin et al. 1998) on citation networks to compute the new p-index. The premise behind pagerank is that it uses the number of links pointing to a web page, as well as the relative standing of pages from where the links originate, to determine the rank of a web page to be displayed in a Google search. Therefore, in citation network parlance, the 'citations' are 'weighted'. The pagerank of a 'node' in a network of N nodes can be calculated from the following formula:

$$P_t(i) = \frac{1-\alpha}{N} + \alpha \sum_j \frac{A_{i,j} P_{t-1}(i)}{k_{out}(j)} \qquad \text{Eq. (1)}$$

where $A_{i,j}$ is the adjacency matrix of the network, $k_{out(j)}$ is the number of outgoing links from node j and α is a reset parameter. A version of the pagerank algorithm is used to dynamically update the 'pagerank' values of each paper. Asynchronous updating is implemented such that each time a paper is uploaded into a citation database, the pagerank values of all papers are updated. The 'p-index (pagerank index)' of the author is calculated by aggregating all the pagerank values of papers to which the author have contributed. Where there is more than one author per paper, the pagerank value of the paper is shared equally. The pagerank values of new papers are governed by the reset parameter α. Therefore, when an author gets cited by a paper which is just published, it will only change the author's p-index marginally. It is only when that paper gets itself cited, that the author's p-index will begin to increase. If that paper is cited by papers which are themselves becoming famous, the p-index will increase considerably. Therefore, the overall dynamics of the research community determines a particular author's p-index in a non-trivial way.

Since the values generated by the pagerank algorithm are decimal numbers which do not make intuitive sense, we define the p-index as a percentile. Therefore, a very famous scientist may have a p-index of 99.99%, whereas a new research student will have a value close to zero. Let us note here that since it is a percentile, the p-index can be applied to a particular field of research, or the entire scientific community. As such, a computer scientist, for example, may have a p-index of 99.9887% in computer science, and 97.5675% overall. By definition, the overall score will be lower than the field specific score. It must be noted that since the p-index utilizes a citation network, it can only be implemented in a scholarly database. Individual scientists cannot calculate their own p-index based on their own publication record. Therefore, the proposed index needs to be implemented in a scholarly database to be fully tested. For this reason, we utilize a simulated system to verify its utility, as described in the next section.

3 Simulation System

In this section, we describe the simulation system which we use to test the proposed index. The system simulates the process whereby a group of scientists produce papers in collaboration and this spawns a new field of research. Eventually, this gives rise to a citation network, where papers are nodes and citations are links, as well as a collaboration network where the scientists are the nodes and collaborations form the links. The former are directed networks, and the latter are undirected. We opted for this approach to demonstrate the utility of the p-index, since we do not yet have access to a real world citation database to implement and test this index. As described below, the evolution process realistically imitates the growth of a research field and corresponding citation and collaboration networks. We will consider three independent 'case studies' using this simulation system. Each case is meant to illustrate a perceived weakness of h-index, and how the p-index is immune to it.

In our simulation system, there are paper objects and author objects. Each of these objects have a number of variables (attributes). Table 1 and Table 2 list the attributes of these objects respectively where dynamic attributes (attributes that are changed in value every iteration) are shown in black and static attributes are shown in blue. Due to space restrictions, we avoid describing these attributes here, most of which are self-explanatory by their names. The variable *isλ* takes on different names for each simulation scenario, as described below.

At the beginning of the simulation, a paper is 'spawned', and some authors are spawned and assigned to it. The number of authors is randomly assigned between 2-5. Each spawned paper is also randomly assigned an impact factor between 0-20. In this initial study we keep the distribution of impact factors linear, though clearly it is biased towards lower values. We also assume that conference papers and similar documents have a low impact factor. When the next papers are spawned, authors are stochastically assigned from the existing author pool or new authors are spawned. The probability of a new author being spawned decreases as the simulation progresses, so that authors are increasingly assigned from existing author pool. At steady state, the probability of an author being a new author rather than somebody from the existing author pool was fixed at 4%. Each spawned paper is randomly assigned a number between 10 and 50 for the number of references it has.

Paper Object
Paper ID
isλ
number of authors
authors list
Impact Factor
number of citations towards the network
number of references
page rank value

Table 2: Attributes of a Paper Node

Author Object
Author ID
isλ
papers list
number of citations
h-index
p-rank
p-index

Table 1: Attributes of an Author Node

However, not all of these references are necessarily to existing papers in the simulated system. In real world, a paper which is a pioneer in a new field will by necessity refer to papers 'outside' its field, since there are no papers already in the emerging field to cite. Even in a saturated field, a certain proportion of the references will be outside the field anyway. To reflect this, we set a varying proportion σ as the proportion of references which are within the simulated system, and this proportion begins with zero and increases linearly with time elapsed until it settles at 70%. Once the number of internal references is determined, existing paper objects are chosen to be the references in the newly spawned paper. These are chosen by weighted preferential attachment, as described below. Thus, the more a paper is already cited, the higher its chances are to be cited by a newly spawned paper, and the impact factor of the paper also plays a part.

The references for newly spawned papers from existing papers were chosen by the well-known 'preferential attachment' method, introduced by Barabasi et al (Barabasi and Albert 1999). In the context of the citation network, preferential attachment works such that a new paper has a higher probability of referencing an existing paper that is highly cited rather than referencing a paper that is less cited or not cited at all. As such, a directed citation network and an undirected collaboration network begin to evolve. We continue this process for a fixed number of timesteps T. In this paper, typically T=15000 was used. In our experiments, we maintained the number of papers: number of authors ratio at roughly 10:1.

After each timestep, the pagerank algorithm was run on the citation network and the pagerank score, and by extension the p-index, of authors were updated. The h-index of authors was also updated, using the available local information for each author (list of papers and citations for each paper). As such, we were able to compare the evolutionary trends of h-index and p-index for each author.

Below, we describe the three simulation scenarios which highlight perceived weaknesses of h-index and how the p-index overcomes them.

3.1 Manipulating the h-index using low impact publications

The first simulation scenario demonstrates a known weakness in h-index where certain authors can publish low impact papers in order to cite their previous work and thus massage their h-index. Henceforth we refer to the authors of these low impact papers as manipulative authors because their intention is to manipulate their h-index, and the low impact publications will be known as manipulative papers, written and published with the sole intention of manipulating and massaging the h-index. A paper is randomly determined to be a manipulative document with a probability of 0.1 at the time of creation. The *isλ* parameter for paper node will be named as *isManipulativePaper* while the *isλ* parameter for authors will be *isManipulativeAuthor*. The probability of being a manipulative author is set to 0.2 and each new author, when created by the simulation system, will be randomly assigned as either manipulative or not manipulative. Manipulative publications are only authored by manipulative authors while both non-manipulative and manipulative authors can contribute to non-manipulative papers. The sole purpose of a manipulative document is to reference the first author's previous work therein massaging their h-index.

We ran simulation experiments to mimic a citation network of 15,000 documents and these documents had 1317 authors altogether among which 152 authors were manipulative authors and 1500 documents were manipulative documents. Results of this simulation are presented in section 4.1.

3.2 Robustness of p-index towards preferences of the authors

In the second simulation scenario, the intention is to demonstrate that p-index is fairer to authors who are less inclined to collaborate. Here, one sector of authors is assumed to be interested in collaboration and typically publish documents with 1-9 authors per each paper. The second sector of authors is relatively not interested in collaboration and only publishes papers that have three authors at most. The implication of this scenario is that the authors who collaborate will have a better chance of getting a higher h-index because of the fact that their paper count is higher regardless of how much each author contributed. Thus it could be expected that the h-index is biased towards collaborators and we wish to compare this feature with the p-index. As such, we identify two types of papers accordingly as collaborative papers and non-collaborative papers. The *isλ* parameter here is named as *isCollaborativeAuthor* and *isCollaborativePaper* respectively for author and paper nodes. Collaborative authors in this scenario will consist of 90% of the author pool while 10% of authors will be non-collaborative authors. The publications will be assigned as 80% collaborative publications and 20% non-collaborative publications.

The system was simulated for 15000 papers consisting of 3006 non-collaborative papers. The author pool included 1328 authors with 126 non-collaborative authors in total. Results of this simulation are presented in section 4.2.

3.3 Robustness of p-index towards scientists who are concerned about quality over quantity

The third simulation scenario deals with a typical problem scientists face when being ranked by h-index. If a scientist is only concerned about the quality of his work and doesn't mind how many publications he contributes to, his h-index may be deteriorated due to the lack of number of papers. It has been shown that even though the papers published by these selective scientists have relatively high impact factors and large number of citations per paper, their lack of quantity in papers challenges their standing in the scientific community (Raan 2006). As such, we simulated the following scenario to observe how p-index would perform comparatively.

The simulation scenario has two types of authors; those who are quality oriented and those who are quantity oriented. In order to quantitatively justify the impact, the simulations were set up such that the quantity oriented authors publish papers with an impact factor of 0-2 while quality oriented authors publish papers with an impact factor of 0-20. The system will generate roughly 10% of quantity oriented authors with 10% of quantity oriented papers. The simulation system had 15000 papers in total of which 1511 were quantity oriented papers and had an author pool of 1357 authors with 138 quantity oriented authors. Results of this simulation are presented in section 4.3.

3.4 Measures and Visualizations

The results of the respective simulations we described above, and the suitability of the newly introduced p-index was evaluated using a set of heterogeneous measures which are elaborated below.

- Plots of the distribution of the h-index and p-index: These plots are used to visualize the disparity between the two classes of authors in respective scenarios. The h-index and p-index are plotted against author ID.
- Plots of the time variation of average h-index and p-index: Average values of the h-index and p-index for both classes of authors are plotted against time step.
- Plots of the time variation of Maximum h-index and p-index: Highest (best) values of the h-index and p-index for both classes of authors are plotted against time step.
- Ratio of differences: We call this measure the *multiplier* within this paper, which signifies roughly how many times one measure is 'fairer' than other. Equation (2) defines this measure mathematically.

  $$multiplier = \Delta h\text{-}index / \Delta p\text{-}index \qquad \text{Eq. (2)}$$

  where Δh-*index* is the difference between average h-index of the two classes of authors while Δp-*index* is the difference between average p-index of the two classes of authors at each timestep, for a given simulation scenario. Please note that we are not claiming to introduce a measure of mathematical significance, but just using a name convention within this paper for easy description.

4 Results

In this section we describe our results for the above mentioned simulation scenarios.

4.1 Manipulative authors vs non-manipulative authors

As portrayed in Fig. 2, the spread of the h-index for non-manipulative authors and manipulative authors provide evidence to the fact that manipulative authors can indeed massage their h-index by publishing low impact papers with the sole purpose of referencing their previous work. On the contrary, we can detect a mixed distribution for the p-index of manipulative and non-manipulative authors as seen in Fig. 3. Whereas in Fig. 2 there is clear separation between, in Fig. 3 the two categories are overlapping. It should be noted that the IDs of the authors signify their duration as a researcher, thus author ID 0 is bound to get more citations than author ID 1300. This explains the lower values obtained by the authors at the right end of the Fig. 2 and Fig. 3. This effect has been discussed in the context of h-index and can be addressed by using a time dependent version of the same index.

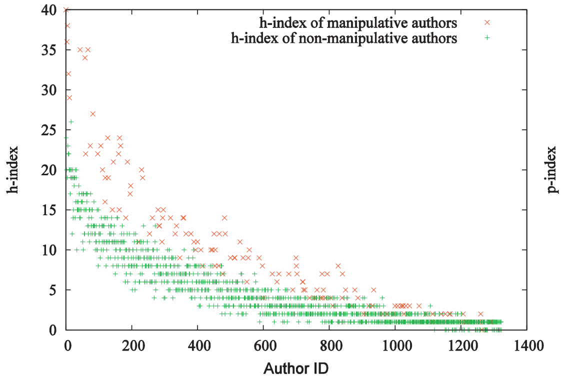

Figure 2: Spread of the h-index for each manipulative and non-manipulative author (as absolute values)

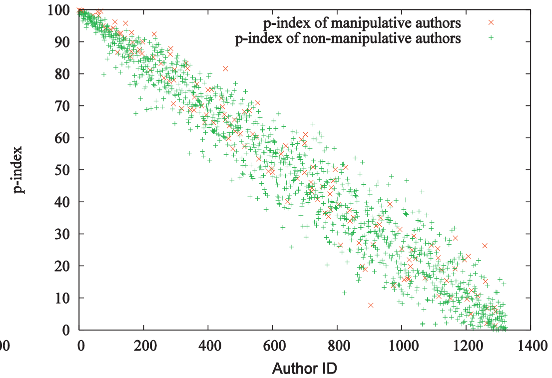

Figure 3: Spread of p-index for each manipulative and non-manipulative author (as percentiles)

Variation of the *average* h-index and p-index for non-manipulative and manipulative authors at each iteration is shown in Fig. 4. It can be inferred from Fig. 4 that p-index is a fairer metric to rank scientists when everyone is not manipulative. The difference between the average h-index of non-manipulative authors and manipulative authors is far greater than the difference between the average p-index of non-manipulative authors and manipulative authors. Quantitatively, the difference of h-index is 110.93% whereas the difference of p-index is only 10.58% which is ten times smaller. The implication is that using h-index, a non-manipulative author will be penalized in average ten times more compared to a manipulative author. We also tested whether the percentile ranking of p-index causes this difference. Thus, we next used the percentile rank of authors with respect to their h-index against the p-index as shown in Fig. 5. Even here, p-index still emerges as the fairer index but the advantage of using p-index diminishes slightly when a percentile h-index is used. The difference of h-index percentile is 53.03% while the difference of p-index is 10.68%. Therefore, the p-index is indeed a fairer measure in comparison to h-index in reducing manipulation by authors.

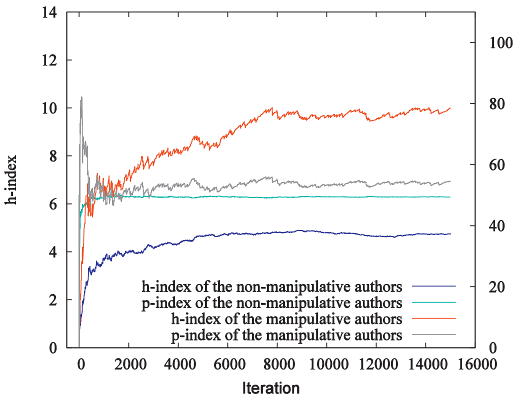

Figure 4: Variation of average h-index and p-index for non-manipulative and manipulative authors at each timestep.

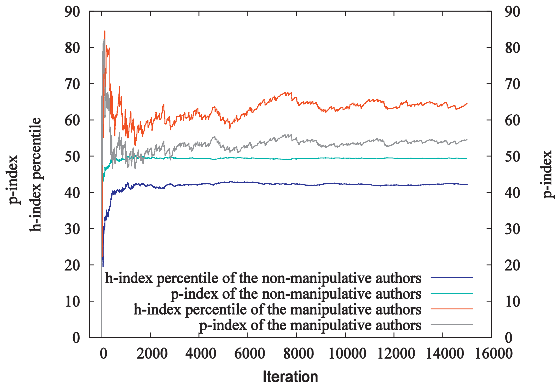

Figure 5: Variation of average h-index percentile and p-index for non-manipulative and manipulative authors at each timestep.

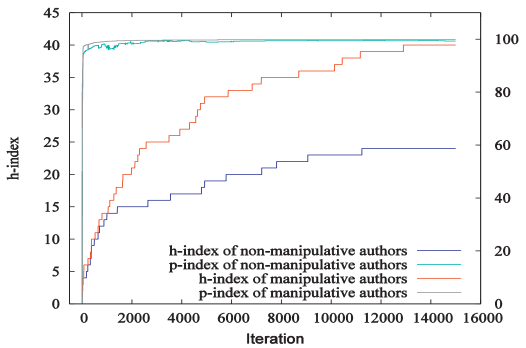

Figure 6: Variation of h-index and p-index for highest ranking non-manipulative and manipulative authors at each timestep.

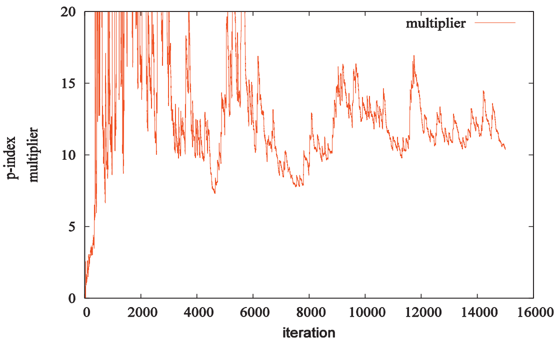

Figure 7: Variation of multiplier at each timestep.

Fig. 6 shows the h-index and p-index respectively for the *highest performing* authors in the simulation. Here the difference is ever starker. The p-index difference between non-manipulative and manipulative authors is insignificant compared to the h-index difference: 0.60% and 66.67% respectively. Fig. 7 characterizes the advantage of using p-index over h-index for an *average* author by using the multiplier defined before. If a non-manipulative author and a manipulative author with adjacent IDs are chosen, what Fig. 7 shows is that using h-index manipulative authors are favored ten times in average more than non-manipulative authors, compared to the p-index. Thus, for large systems of authors, the p-index is more than ten times ‘fairer’.

4.2 Collaborative authors vs non-collaborative authors

Fig. 8 shows the spread of h-index for collaborative and non-collaborative authors, while Fig. 9 shows the spread of p-index for collaborative and non-collaborative authors, over author-ID. We can see that most non-collaborative authors have relatively low h-indices, whereas in terms of p-indices the distribution is more evenly spread. As shown in Fig. 8 and Fig. 9, it is visible that collaborative authors usually have higher h-indices whereas both classes of authors have similar p-indices.

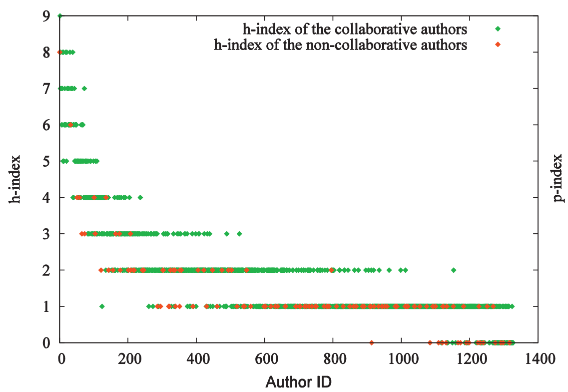

Figure 8: The spread of h-index for collaborative and non-collaborative authors (as absolute values).

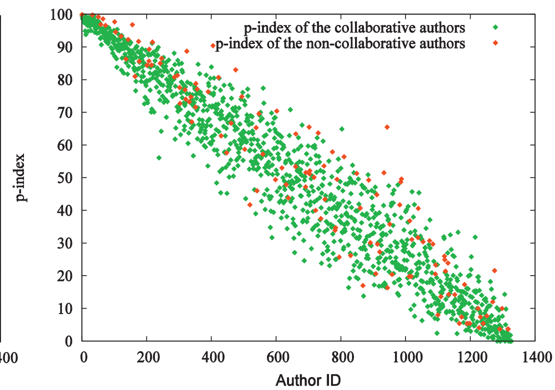

Figure 9: The spread of p-index for collaborative and non-collaborative authors (as percentile).

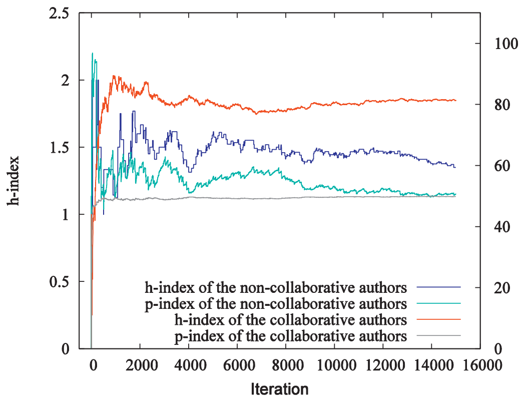

Figure 10: Variation of average h-index and p-index for collaborative and non-collaborative authors.

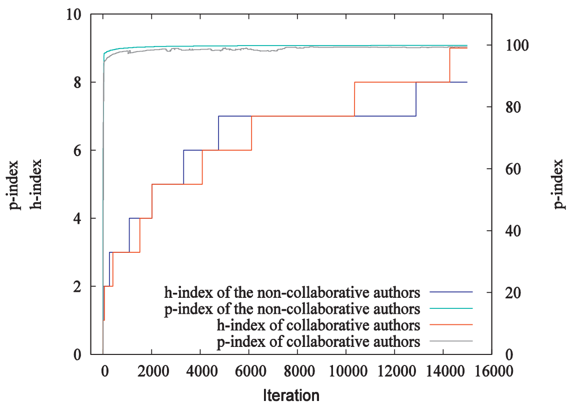

Figure 11: Variation of h-index and p-index for highest ranking collaborative and non-collaborative authors.

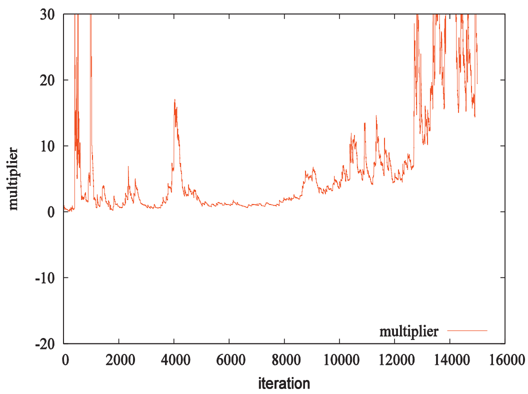

Figure 12: Variation of multiplier over each iteration for collaborative and non-collaborative authors over each timestep.

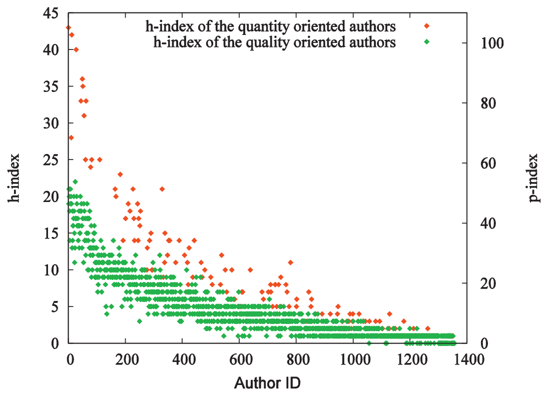

Figure 13: The spread of h-index for quantity oriented authors and quality oriented authors.

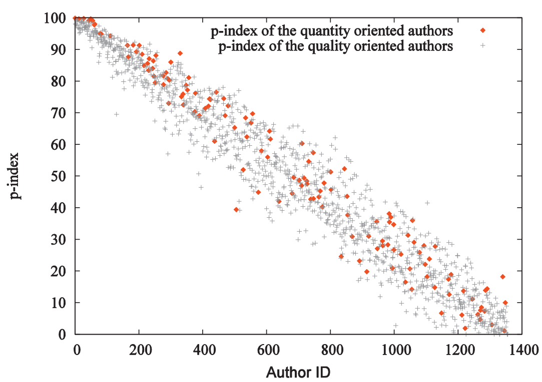

Figure 14: The spread of p-index for quantity oriented authors and quality oriented authors.

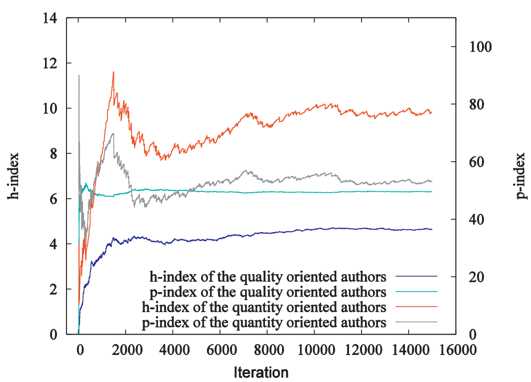

Figure 15: Variation of average h-index and p-index for quantity oriented authors and quality oriented authors.

Analyzing the average fluctuation of h-index and p-index for collaborative and non-collaborative authors reveal interesting characteristics. According to the Fig. 10 which shows the variation of *average* h-index and p-index at each timestep, average h-index of the collaborative authors is 37% higher compared to non-collaborative authors. From Fig. 8, Fig. 9 and Fig. 10 we can infer that p-index is robust against tendencies in collaboration.

The h-index and p-index of the highest ranking authors of both classes are shown in Fig. 11. This re-emphasizes that neither class of authors would gain a quantifiable benefit in position by using p-index whereas it is entirely possible by using h-index. Fig. 12 quantifies the advantage one class of authors can have, on average, over the other class of authors by using h-index instead of p-index, by using the multiplier as defined in the previous section. This figure indicates that the p-index, at steady state, is roughly about twenty times 'fairer' than the h-index when collaborative tendencies are considered.

4.3 Quality oriented authors vs Quantity oriented authors

Fig. 13 shows the spread of h-index for quality oriented and quantity oriented authors, while Fig. 14 shows the spread of p-index for quality oriented and quantity-oriented authors. Again it can be seen that while the h-index distribution favours quantity oriented authors, the two classes nearly overlap in p-index distribution. Fig. 15 shows the variation of the *average* h-index and average p-index of the quality oriented and quantity oriented authors. From this figure, it can be inferred that average h-index of quantity oriented authors is significantly higher whereas average p-index is only marginally affected. The difference of p-index is 7.39% whereas the difference of h-index is well over 110% which puts the robustness of p-index in to perspective in terms of quality vs quantity of publications.

Fig. 16, on the other hand, shows the h-index and p-index variation of the *highest ranked* quality oriented authors and quantity oriented authors where we can infer that the percentage difference of h-index maybe as much as 104% while the difference of p-index is as low as 2% showing that the differences can be amplified more than 50x times when h-index is used and this can have a dire effect on the quality oriented authors. Fig. 17 plots the *multiplier* for the average quality vs quantity oriented authors, in this scenario. It can be observed that the multiplier fluctuates around 15 in steady state, meaning that the advantage a quantity oriented author gains over a quality oriented author is fifteen times less when p-index is used.

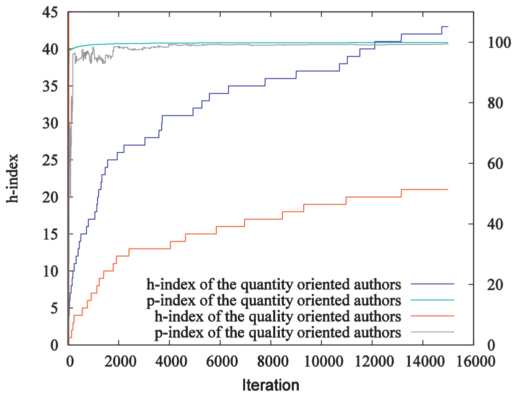

Figure 16: Variation of h-index and p-index for highest ranking quality oriented authors and quantity oriented authors.

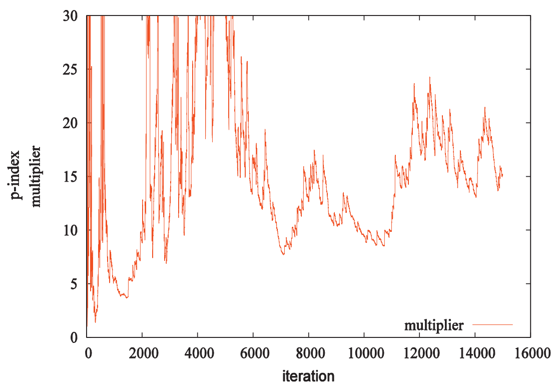

Figure 17: Variation of multiplier of the quantity oriented authors and quality oriented authors over each timestep.

5 Conclusion

We present here a new index, the p-index (pagerank-index) to quantify the scientific output of researchers. The p-index addresses the lack of weighted citations which is a core shortcoming of h-index. Since h-index considers all citations to be equal in nature, it has become achievable to manipulate h-index for personal gain using diverse means. The p-index provides a much fairer means of comparing scientists, and is a lot less prone to manipulation and massaging.

We demonstrated the superiority and fairness of p-index over h-index using three simulated scenarios. The first scenario demonstrated that a manipulative author who is interested in publishing low impact papers in order to reference their own papers can indeed massage their h-index. On the contrary, p-index is much fairer and doesn't allow manipulative authors to gain an edge. The second scenario revealed that the authors who are interested in contributing small parts to a number of papers will certainly gain a higher h-index. However such authors cannot gain the same advantage by using p-index which proves that it is a fairer metric. The third simulation scenario described a situation where authors are interested in maximizing their number of publications so that they may obtain a significant benefit by using h-index. This is not feasible if p-index is used as shown clearly in the results of the simulation. In conclusion, we were able to prove that p-index is robust against manipulations and performs fairer and more effectively in ranking scientists. We quantified our findings, and found that on average, in each scenario the p-index reduces the unfair advantage to one group over the other by ten to twenty times. This is a very significant result.

Even though our simulation system evolves a citation network and the underlying collaboration networks from the beginning of a specific field, we can apply p-index at any point in time for an existing scholarly database. This makes it possible to seamlessly integrate p-index to existing scholarly databases.

6 Future Work

In this work, we used a simulated system to represent the process of a community of authors writing papers. We intend to replicate our results by using a real community of authors, by using authors from a particular field from google scholar. To this end, we will write crawlers which can collect data from a particular emerging field ('networked game theory' for example), and use the real citation network to compute the p-index, and compare it with the temporal h-index values of the authors. Work is progressing in this regard.

References

Barabasi, A. L. and R. Albert (1999). "Emergence of scaling in random networks." Science **286**(5439): 509-512.

Bornmann, L. and H.-D. Daniel (2007). "What do we know about theh index?" Journal of the American Society for Information Science and Technology **58**(9): 1381-1385.

Bornmann, L., R. Mutz and H.-D. Daniel (2008). "Are there better indices for evaluation purposes than theh index? A comparison of nine different variants of theh index using data from biomedicine." Journal of the American Society for Information Science and Technology **59**(5): 830-837.

Chen, P., H. Xie, S. Maslov and S. Redner (2007). "Finding Scientific gems with Google's PageRank algorithm." Journal of Informetrics.

Egghe, L. (2006). "Theory and practise of the g-index." Scientometrics **69**(1): 131-152.

Ghoshal, G. and A. L. Barabasi (2011). "Ranking stability and super-stable nodes in complex networks." Nat Commun **2**: 394.

Hirsch, J. E. (2005). "An index to quantify an individual's scientific research output." Proceedings of the National Academy of Sciences of the United States of America **102**(46): 16569-16572.
Jin, B., L. Liang, R. Rousseau and L. Egghe (2007). "The R- and AR-indices: Complementing the h-index." Chinese Science Bulletin **52**(6): 855-863.
Leicht, E. A., G. Clarkson, K. Shedden and M. E. J. Newman (2007). "Large-scale structure of time evolving citation networks." The European Physical Journal B **59**(1): 75-83.
Maslov, S. and S. Redner (2009). "Promise and Pitfalls of Extending Google's PageRank algorithm to citation networks."
Page, L., S. Brin, R. Motwani and T. Winograd (1998). The PageRank Citation Ranking: Bringing Order to the Web.
Raan, A. J. (2006). "Comparison of the Hirsch-index with standard bibliometric indicators and with peer judgment for 147 chemistry research groups." Scientometrics **67**(3): 491-502.
van Nierop, E. (2009). "Why do statistics journals have low impact factors?" Statistica Neerlandica **63**(1): 52-62.
Walker, D., H. Xie, K. Yan and S. Maslov (2007). "Ranking Scientific Publications using a model of network traffic." Journal of Statistical Mechanics.
Waltman, L., R. Costas and N. Jan van Eck (2012). "Some Limitations of the H Index: A Commentary on Ruscio and Colleagues' Analysis of Bibliometric Indices." Measurement **10**(3): 172-175.