International Journal of Computer Science and Applications,

EVALUATING THE PRODUCTIVITY OF RESEARCHERS AND THEIR COMMUNITIES: THE RP-INDEX AND THE CP-INDEX

JÖRN ALTMANN

Technology Management, Economics and Policy Program, School of Industrial and Management Engineering
College of Engineering, Seoul National University
Sillim-Dong, Gwanak-Gu, 151-742 Seoul, South-Korea
jorn.altmann@acm.org

ALIREZA ABBASI

Technology Management, Economics and Policy Program, School of Industrial and Management Engineering
College of Engineering, Seoul National University
Sillim-Dong, Gwanak-Gu, 151-742 Seoul, South-Korea

Research and Development Division, International Affairs, National Iranian Oil Company
Tehran, Iran
aeabbasi@yahoo.com

JUNSEOK HWANG

Technology Management, Economics and Policy Program, School of Industrial and Management Engineering
College of Engineering, Seoul National University
Sillim-Dong, Gwanak-Gu, 151-742 Seoul, South-Korea
junhwang@snu.ac.kr

While the h-Index and the g-Index (as the major indices for quantifying the academic performance of researchers) take into consideration the citation count of publications, some other important indicators of research output (i.e. the number of authors per paper, lead author, year of publication) are omitted. Those indicators have to be considered in order to evaluate the productivity of researchers comprehensively. This paper analyzes the different indicators and proposes two new indices, the RP-Index and the CP-Index. The RP-Index evaluates the productivity of a researcher and the CP-Index evaluates the productivity of a group of researchers. After showing how these new indices can be applied and how they compare to the existing ones, an assessment of the two new indices is given.

Keywords: h-Index, g-Index, research productivity evaluation, performance evaluation of researchers and groups, citation indices, metrics, empirical data analysis, knowledge creation, and knowledge transfer.

1. Introduction

At universities, industry funding, government funding, faculty recruitment, and recruitment of excellent students increasingly depend on the ranking position of the university. In order to rank universities, the performance of researchers, groups of researchers, departments, schools, colleges, and universities has to be quantified, requiring an objective evaluation of researchers' productivity. At the same time, the evaluation of researchers can be used for the internal faculty promotion process, the allocation of academic awards and grants, and even for hiring of new faculty.

Currently, however, there is little data available to evaluate the efficiency of researchers' performance and their communities. In addition to this, there is a lack of tools and measures for the evaluation process. Finally, only little research analysis about knowledge creation and knowledge transfer has been performed so far in this area.

Considering this situation, the overall goal must be to evaluate the efficiency of knowledge creation and knowledge transfer of researchers within an academic community by using tools for collecting and evaluating data about the researchers' output and their collaboration activities. In particular, the following three questions have to be addressed: First, how can we identify the knowledge structure of a research community? Second, how can we evaluate the knowledge structure with respect to the criterion of which community and individuals are more productive and collaborative? Finally, how can communities improve their productivity and collaboration activities?

Within this paper, we focus only on one of these questions, namely on the question of how the performance of researchers and their communities can be measured. Objective measures make the work of researchers comparable. Quantifying the output of researchers and having an index are visible means to achieve this aim. The general idea of quantifying the researchers' productivity is to rank those researchers high that publish a lot, receive a high citation count of their publications, and collaborate with external (with respect to the community that the researcher belongs to) researchers, and acquire research funding. In a second step, which is not in the scope of this paper, we will use the results presented in this paper to analyze the knowledge transfer between researchers, which is based on collaboration with other external researchers, collaboration with industry, giving lectures, and collaboration with researchers of the same department.

However, obtaining the basic data for the productivity analysis is difficult. Citation data is difficult to find and is limited to citation databases [Lehmann (2006)] [Hirsch (2005)]. There are only a few citation databases available. Some of them are commercial, requiring an access charge, and others are publicly accessible [ISI Web of Knowledge (2008)] [Google Scholar (2008)] [Scopus (2008)] [ACM Portal (2008)] [CiteSeer (2008)].

Within this paper, we propose two indices for measuring the productivity of researchers or communities of researchers, the Researcher Productivity Index (RP-Index) and the Community Productivity Index (CP-Index). These new indices combine the advantages of previous indices (i.e. h-Index, g-Index, or its derivates) such as the year of publication, the number of co-authors, and, to reward major contributors to papers, the research-leading author. However, the major difference to existing indices is the idea of calculating the average number of citations of the top cited publications to determine the index value of a researcher. Therefore, the number of papers considered for the RP-Index, depends on the average of citations itself and not on the top-h (according to the h-Index) publications. In addition to this, both indices have as output real numbers, which allows a more detailed evaluation than existing indices (which produce only integer numbers as output).

The remainder of the paper is organized as follows. Section 2 describes the state-of-the-art in performance evaluation of researchers. After briefly describing the data being used in Section 3, the indices are defined in section 4. In section 5, the new indices are applied to the data set and compared with existing indices (e.g. h-Index and g-Index). The paper concludes with a discussion of the results and an outlook in section 6.

2. State of the Art

There have been discussions on whether bibliometric indicators are proper tools for evaluating the research performance of scientists. Peter Weingart stated that the major problem is not the usage of bibliometric data but rather the insufficient development of bibliometric indicators [Weingart (2005)]. However, numerous studies have been conducted that validate the use of citations as an evaluation method. For example, a comparison of the number of citations with peer reviews for communication scholars has been shown in [So (1997)]. The results illustrate that the ranking based on citations and the ranking based on expert judgments are highly correlated.

Hirsch's h-Index is the most known measure for evaluating the performance of a researcher based on bibiometric indicators [Hirsch (2005)]. The input parameters are the number of publications and the number of citations. Because of its simplicity, it has been accepted to be the method for quantifying the research output of scientists. Its definition is easy to understand: A scientist has an h-index of *h* if *h* of his papers have at least *h* citations each and there are no *h+1* papers which have been cited *h+1* times each. That means that the h-Index identifies the core publications of an author by considering the number of citations. These publications are referred to as the Hirsch core (or h-core) [Burrell and Quentin (2007] [Jin (2007)]. This approach of considering only a subset of papers helps addressing the fact that many papers get published without a proper reviewing process. This way, publications with low quality will not be considered.

Although the h-Index represents the most successful concept, it comes with some drawbacks. First, their values only increase over time and, second, they do not consider some basic factors such as the number of co-authors per paper. If the paper has a high number of authors, it is likely that the paper will get a higher reference later on (simply because of the fact that each of the authors cites the paper later on in his/her post works). Therefore, the h-Index has become the source for many variations. Batista et al. modified the h-Index by dividing the standard h-Index by the average number of authors in the articles that contribute to the h-Index [Batista et al. (2006)]. It reduces the effects of co-authorship. Sidiropoulos et al. proposed a contemporary h-Index by giving more weight to recent articles [Sidiropoulos et al. (2007)]. It rewards academics who maintain a steady level of activity. Jin et al. proposed the A-Index, which is the average number of citations of the *h* publications (*h* is defined as for the h-Index) [Jin et al. (2006)]. It has also been proposed to use the tuple (h, AR) as a measure, in which AR defines the age of the publications [Jin (2007)]. Egghe introduced the g-Index to overcome the main shortcomings of the h-Index, namely, ignoring the number of citations in excess of *h*

[Egghe (2006)]. The g-Index is defined as the largest number such that the top *g* articles receive together at least g^2 citations.

The following table gives a summary of the advantages and disadvantages of seven different measures that have been used for evaluating the performance of researchers based on bibliometric data.

Table 1. Summary of the advantages and disadvantages of seven different measures for evaluating the productivity of researchers

	Measures	Advantages	Disadvantages
1.	***Number of papers***	- Measures quantity	- Does not measure impact of papers
2.	***Number of citations***	- Measures impact	- Might be inflated through a small number of highly cited papers with many coauthors. - Gives weight to highly cited articles versus original research contributions
3.	***Average number of citations***	- Allows comparing scientists of different ages	- Rewards researchers with a low number of papers - Penalizes researchers with a high number of papers
4.	***Number of significant papers*** (number of papers with *y* citations]	- Eliminates the disadvantages of measures 1, 2, and 3 - Measures broad and sustained impact	- *y* is arbitrary and will randomly favor or disfavor individuals - *y* needs to be adjusted for different levels of seniority
5.	***Number of citations to the y most-significant papers***	- Eliminates the disadvantages of measures 1, 2, and 3	- Is not a single number, making it more difficult to obtain and compare. - Has the same disadvantages as measure 4
6.	***h-index*** [Hirsch (2005)]	- Measures the quantity and the broad impact of a work - Eliminates the disadvantages of measures 1, 2, and 3 - De-emphasizes single, successful publications	- Is bounded by the number of publications - Has less accuracy than the simpler measure of the average number of citations [Lehmann (2006)] - Depends on the person's scientific age and does not account for the number of authors [Glänzel (2006b)] - Never decreases and does not differentiate between active and inactive researchers [Sidiropoulos (2007)]
7.	***g-index*** [Egghe (2006)]	- - Gives more weight to highly cited articles	- Has the same disadvantages as the h-Index

There are also some studies on measuring the research output of groups of researchers. For example, based on the g-Index, Tol evaluates groups of researchers [Tol (2008)]. A group of researchers has a g1-Index of *g1* if all the researchers have at least an g-Index of *g*. Using the h-index as the foundation, Prathap defined the h1-index and h2-index to quantify the performance of institutes [Prathap (2006)]. Some more variations of the h-indices were defined for evaluating journals, publishers, and nations [Schubert (2007)] [Braun et al. (2005)].

In addition to this, some more criteria could be considered. The leadership factor is one example. The idea of the leadership factor is to give credit to the author who came up with the research idea and conducted most of the work when research is ranked [Oh et al. 2006] [Klavans and Boyak (2008)]. However, the leadership factor is difficult to determine. The order of authors on the paper cannot be considered for this, since it is not always an indication of the amount of work that has been contributed. The reasons can be manifold, ranging from the practice in a research communities to political reasons. Besides, in some communities, the leadership factor is not considered anymore at all. Although we believe that this criterion could be omitted, our index can consider this.

Another example of a criteria that could be considered is the quality of a journal or proceedings. However, the quality of a journal gets less important in some research communities. This is true for example in the area of computer science. Since a simple search on the Web (or in online libraries) will reveal many articles that are freely downloadable from the home pages of researchers, there is no need to search in the most known journal anymore. For our work, we do not consider the importance of journals.

3. Data Collection

The study collected data on five information schools (iSchools)[a], namely the iSchools of University of Pittsburgh (PITT), University of California at Berkeley (UCB), University of Maryland (UMD), University of Michigan (MICH), and Syracuse University (SYR). These schools have been chosen, since they offer similar programs in the area of information management and systems. In addition to this, this kind of schools is new within the university landscape.

The data sources are school reports, which included the list of publications of researchers, the ACM portal[b], Google Scholar[c], and DBLP[d]. The ACM Portal and Google Scholar provided the citation data. This data can be considered sufficiently precise for the purpose of this paper, since Ruane and Tol's analysis showed that the rankings based on Google Scholar highly correlate to the rankings based on the Web of Science and Scopus for Irish researchers. Therefore, it is not relevant that Google Scholar shows higher

[a] Although our program has a similar focus as these iSchools, we excluded our program, TEMEP, from the data collection. Since TEMEP only started in 2003, not sufficient data was available at the time of the data collection.

[b] http://portal.acm.org/portal.cfm

[c] http://scholar.google.com

[d] http://www.informatik.uni-trier.de/~ley/db/

publication and citation numbers and, therefore, higher h-indices and g-indices than Web of Science and Scopus [Ruane and Tol (2007)] [Kousha and Thelwall (2007)].

We did not differentiate between the types of publications (i.e. proceedings of local and international conferences, journals, books, and presentations were ranked equally). The data covered five years, between 2001 and 2005. Since the data for year 2002 was not available for the Maryland iSchool, we substituted it with data from year 2006.

For our study, we used AcaSoNet, a soon-to-be-released Web application for extracting publication information (e.g. author names, title, date of publication, publisher, and number of citations) from the Web. AcaSoNet identifies relationships (e.g. co-authorship) between researchers, and stores the relationships in a database.

After revising the data, 2139 publications of 1815 authors were available. Of those publications, 509 publications have not been cited in Google Scholar. The reason for this is that those 509 publications were published in newspapers, magazines, and on Web sites, which usually give a wider visibility of the author but are not referred to in scientific articles (e.g. Prof. Cox at Pittsburgh). Those articles are also not listed in other article databases than Google Scholar. 5310 co-authorships were detected.

4. Methodology

As a basic performance measure for calculating the productivity of researchers and the productivity of communities of researchers (i.e. productivity of research groups, institutes/departments, schools, colleges, or universities), we considered the number of publications (NP), the number of co-authors, and the number of citations only. We did not consider the researchers' collaboration activities with external researchers (with respect to the community that they belong to) and the researchers' ability to acquire research funding, since data was not available for this[e].

4.1. *Researcher Productivity Index (RP-Index)*

The basis for the Researcher Productivity Index (RP-Index) is the normalized number of paper citations of a researcher *j* (NC_{ji}). The NC_{ji} is calculated as the number of citations of paper *i* of researcher *j* divided by the number of years that the paper is available and multiplied by a factor C_{ji}, which represents the contribution of the researcher *j* to the paper *i* (Eq. 1). The Contribution Factor C_{ji} is in the range between 0 and 1. If all authors of the paper *i* contributed equally, the factor is supposed to be $C_i = 1 / NumberOfAuthorsOfPaper_{ji}$. If there has been a leading author, the factor can be increased. However, the sum of the factors for each author should add up to one

$$NC_{ji} = \frac{NumberOfPaperCitations_i}{AgeOfPaper_{ji}} * C_{ji}, \quad 0 \le C_{ji} \le 1. \tag{1}$$

[e] It should be noted that, although the indices that are introduced in the following sections are used in the research context, these indices could be used in different contexts to evaluate the performance of individuals and communities.

The researcher productivity index (*RP-Index*) of researcher *j* is similarly defined to the h-Index and g-Index. Given that the publications of the researcher *j* are sorted according to the NC_{ji} of each publication *i* in decreasing order, the RP-Index can be defined as the largest natural number *x* such that the top *x* publications have at least in average a value of *x* for their NC_{ji}

$$RP_j^{'} = \arg\max_x \left(\frac{1}{x} \sum_{i=1}^{x} NC_{ji} \;\middle|\; \frac{1}{x} \sum_{i=1}^{x} NC_{ji} \geq x \right), \quad x \in \mathrm{N}. \tag{2}$$

If we allow real numbers for RP_j , we can generate a more fine-grained evaluations of the performance of researchers than the h-Index. Since real numbers are used in many different contexts to evaluate people (e.g. school, college, graduate school), RP_j' is also appropriate to evaluate researchers. Following this approach, we obtain Eq. (3) after slightly modifying equation Eq. (2):

$$RP_j = \max \left(\frac{1}{x} \sum_{i=1}^{x} NC_{ji} \;\middle|\; \frac{1}{x} \sum_{i=1}^{x} NC_{ji} \geq x \right), \quad x \in \mathrm{N}. \tag{3}$$

Within the remainder of the paper, we only refer to equation Eq. (3) when we mention the researcher productivity index.

4.2. *Community Productivity Index (CP-Index)*

A community can be any group of individuals. In the research context, an individual (i.e. researcher) can belong to different communities. For example, at a university, we can distinguish, in hierarchical order, research groups, departments, schools, colleges, and the entire university. Such a hierarchical classification allows comparing the performance of communities at different levels.

Based on the same concept for evaluating the productivity of researchers, a Community Productivity Index (CP-Index) can be introduced. The CP-Index of a research community *k* is defined as the largest natural number *y* such that the top *y* researchers of this research community have at least in average a value of *y* for their RP-Index, given that the researchers are sorted according to their RP-Index in decreasing order

$$CP_k^{'} = \arg\max_y \left(\frac{1}{y} \sum_{j=1}^{y} RP_{kj} \;\middle|\; \frac{1}{y} \sum_{j=1}^{y} RP_{kj} \geq y \right), \quad y \in \mathrm{N}. \tag{4}$$

RP_{kj} denotes that researcher j belongs to community k. In the same way as for the researcher productivity index, we can generate a more fine-grained evaluation of the community productivity by using real numbers. Modifying equation Eq. (2) slightly, we obtain the following equation for CP_k:

$$CP_k = \max\left(\frac{1}{y}\sum_{j=1}^{y} RP_{kj} \;\middle|\; \frac{1}{y}\sum_{j=1}^{y} RP_{kj} \geq y \right), \quad y \in \mathrm{N}. \tag{5}$$

For our analysis of the different research communities (iSchools), we only consider formula Eq. (5).

5. Results

In order to test the newly defined indices, we apply them to the data that we collected about the five iSchools. Table 2 shows the top-10 publications with respect to the highest normalized number of citations, NC (Eq. (1)). The contribution factor has been set to 1. Analyzing Table 2, it is obvious that NC enables newly published papers to get a higher ranking than in a ranking which considers the absolute number of citations.

Table 2. Top-10 publications with respect to NC and their absolute number of citations.

	Title of Paper	Year	Citation	NC
1	Institutional repositories: Essential infrastructure for scholarship in the digital age	2003	285	**57**
2	The political blogosphere and the 2004 US election: Divided they blog	2005	164	**54.66**
3	Adaptive hypermedia	2001	316	**45.14**
4	The library media center: Touchstone for instructional design and technology in the schools	2004	169	**42.25**
5	Friends and neighbors on the Web	2003	179	**35.8**
6	Information systems in organizations and society: Speculating on the next 25 years of research	2004	141	**35.25**
7	Product market regulation in OECD Countries: 1998 to 2003	2005	201	**33.5**
8	The Bangalore boom: From brain drain to brain circulation	2004	119	**29.75**
9	The Silicon Valley-Hsinchu connection: Technical communities and industrial upgrading	2001	208	**29.71**
10	The Role of children in the design of new technology	2002	172	**28.66**

In order to calculate the Researcher Productivity Index, we calculate the NC for each publication of all the iSchool researchers. As an example, Table 3 shows the NC of the top 20 publications of Peter Brusilovsky. The contribution factors *C* of the papers are set reversely proportional to the number of co-authors (Eq. (1)).

Table 3. Top-20 publications of Peter Brusilovsky with respect to NC.

	Title	**Authors**	**Year**	**Citations**	**NC**
1	Adaptive hypermedia	P. Brusilovsky	2001	316	45.14
2	Adaptive navigation support in educational hypermedia: The role of student knowledge level and the case for meta-adaptation	P. Brusilovsky	2003	74	14.8
3	ELM-ART: An adaptive versatile system for Web-based instruction	G. Weber, P. Brusilovsky	2001	193	13.78
4	Course sequencing techniques for large-scale web-based education	P. Brusilovsky, J. Vassileva	2003	96	9.6
5	From adaptive hypermedia to THE adaptive Web	P. Brusilovsky, M. T. Maybury	2002	81	6.75
6	A framework for adaptive e-learning based on distributed re-usable learning activities	P. Brusilovsky, H. Nijhawan	2002	71	5.91
7	Adaptive systems for Web-based education	P. Brusilovsky, N. Henze, E. Millan	2002	79	4.38
8	Course delivery systems for the virtual university	P. Brusilovsky, P. Miller	2001	61	4.35
9	Domain, task, and user models for an adaptive hypermedia performance support system	P. Brusilovsky, D. W. Cooper	2002	49	4.08
10	Map-based horizontal navigation in educational hypertext	P. Brusilovsky, R. Rizzo	2002	47	3.91
11	Social navigation support through annotation based group modeling	R. Farzan, P. Brusilovsky	2005	22	3.66
12	Social adaptive navigation support for open corpus electronic textbooks	P. Brusilovsky, G. Chavan, R. Farzan	2004	44	3.66
13	Preface to special issue on user modeling for Web information retrieval	P. Brusilovsky, C. Tasso	2004	23	2.87
14	Adaptive visualization component of a distributed Web-based adaptive educational system	P. Brusilovsky, H. D. Su	2002	32	2.66
15	Engaging students to work with self-assessment questions: A study of two approaches	P. Brusilovsky, S. Sosnovsky	2005	14	2.33
16	WebEx: Learning from examples in a programming course.	P. Brusilovsky	2001	16	2.28
17	User modeling in a distributed e-learning architecture	P. Brusilovsky, S. Sosnovsky, O. Shcherbinina	2005	17	1.88
18	Assessing student programming knowledge with Web-based dynamic parameterized quizzes	S. Pathak, P. Brusilovsky	2002	21	1.75

19	Layered evaluation of adaptive learning systems	P. Brusilovsky, C. Karagiannidis D. Sampson	2004	21	1.75
20	Comprehensive personalized information access in an educational digital library	P. Brusilovsky, R. Farzan, J. Ahn	2005	12	1.33

Based on the NC values for each publication of each researcher, the Researcher Productivity Index (RP-Index) for each researcher can be calculated. The RP-Index of Peter *Brusilovsky* is 10.57. Table 4 shows the researcher productivity index (RP-index) of those researchers with the highest RP-Index. The table also depicts the number of publications (NP), the average number of citations (CitAvg), the h-Index (h), and the g-Index (g). For a more detailed comparison with the RP-Index, the h-Index and the g-Index are also shown for NC values (h-NC and g-NC). It illustrates the impact of the NC on the indices calculation.

Table 4. The top 20 researchers with respect to the RP-Index[f]

	First Name	Family Name	School	NP	CitAvg	h	g	RP	h-NC	g-NC
1	Lada Ariana	Adamic	MICH	10	100.9	9	10	15.48	7	10
2	AnnaLee	Saxenian	UCB	20	43.8	11	20	11.37	7	13
3	Kevin	Crowston	SYR	42	21.83	17	30	10.58	7	12
4	Peter	Brusilovsky	PITT	52	27.36	17	37	10.57	6	11
5	Martha E.	Pollack	MICH	32	31.81	16	31	9.71	7	10
6	Paul	Resnick	MICH	13	46.07	9	13	9.67	6	10
7	Clifford	Lynch	UCB	20	30.4	9	20	9.51	5	11
8	Danah	Boyd	UCB	25	18.24	9	21	9.41	5	10
9	Hal	Varian	UCB	98	6.69	12	25	8.57	6	9
10	Ping	Zhang	SYR	30	15.86	13	21	8.53	6	10
11	Richard J.	Cox	PITT	89	4.44	9	19	8.18	5	9
12	Allison	Druin	UMD	38	15.18	13	23	7.78	3	8
13	James	Howison	SYR	12	30.33	10	12	7.77	5	8
14	Peter	Honeyman	MICH	12	34.66	6	12	7.62	4	8
15	Marti	Hearst	UCB	33	42.33	15	33	7.49	5	9
16	Mark S.	Ackerman	MICH	24	18.75	11	21	6.77	5	8
17	Dave	Molta	SYR	15	16.93	2	15	6.65	2	7
18	Paul	Conway	MICH	19	14.31	5	16	6.34	3	7
19	Dragomir R.	Radev	MICH	35	20.74	16	26	6.31	5	8
20	Elliot	Soloway	MICH	39	23.74	16	30	6.25	5	8

[f] PITT = Pittsburgh iSchool, UCB = Berkeley iSchool, UMD = Maryland iSchool, MICH = Michigan iSchool, SYR = Syracuse iSchool

From MICH, there are 40% researchers on this list, followed by UCB with 25%, SYR with 20%, PITT with 10%, and UMD with 5%. The highest ranked researcher is from MICH. Apart from two researchers, James Howison and Danah Boyd (Rank 8 and 13), who just finished their PhD, all researchers on this list are faculty members.

The comparison of the RP-Index with the h-Index shows a significantly different ranking of those researchers, although a large fraction can be accounted towards the NC data (the h-Index performed on NC data). The comparison of the RP-Index with the h-Index (applied to NC data) also reveals that the RP-Index differentiates between the researchers in more detail than the h-Index or g-Index. The range of values generated by the RP-Index is much wider (9 units) than for the h-Index (5 units) and the g-Index (6 units). Through the use of real numbers, not many researchers have the same RP-Index. If researchers produce continuously the same quality of research over time

In order to evaluate the productivity of groups of researchers in more detail, we calculate the RP-Index for each of the researchers of each iSchool. The top-10 researchers of each iSchool are listed in decreasing order of their researcher productivity index (RP-index) in Table 5.

Table 5. For each iSchool, the top-10 researchers with respect to the RP-Index are illustrated

	Name	**Family**	**School**	**NP**	**CitAvg**	**h**	**g**	**RP**
1	Peter	Brusilovsky	PITT	**52**	**27.4**	**17**	**37**	**10.6**
2	Richard J.	Cox	PITT	**89**	**4.44**	**9**	**19**	**8.18**
3	James B.D.	Joshi	PITT	**11**	**33.7**	**8**	**11**	**5.1**
4	David	Tipper	PITT	**25**	**10.2**	**6**	**15**	**4.49**
5	Prashant	Krishnamurthy	PITT	**29**	**7.41**	**8**	**14**	**3.72**
6	Paul D.	Robins	PITT	**10**	**16.9**	**3**	**10**	**3.66**
7	Vladimir	Zadorozhny	PITT	**18**	**11.3**	**8**	**14**	**3.34**
8	Stuart W.	Shulman	PITT	**17**	**5**	**6**	**9**	**2.79**
9	Hassan	Karimi	PITT	**16**	**7.87**	**6**	**11**	**2.76**
10	Michael	Lewis	PITT	**42**	**9.28**	**13**	**18**	**2.45**
11	AnnaLee	Saxenian	UCB	**20**	**43.8**	**11**	**20**	**11.4**
12	Clifford	Lynch	UCB	**20**	**30.4**	**9**	**20**	**9.51**
13	Danah	Boyd	UCB	**25**	**18.2**	**9**	**21**	**9.41**
14	Hal	Varian	UCB	**98**	**6.69**	**12**	**25**	**8.57**
15	Marti	Hearst	UCB	**33**	**42.3**	**15**	**33**	**7.49**
16	J. Doug	Tygar	UCB	**16**	**54.1**	**11**	**16**	**5.34**
17	John	Canny	UCB	**17**	**18.3**	**10**	**17**	**5.13**
18	Nathan	Good	UCB	**15**	**31.2**	**7**	**15**	**4.34**
19	Pamela	Samuelson	UCB	**20**	**7.6**	**6**	**12**	**3.98**
20	Nancy	Van House	UCB	**16**	**22.2**	**10**	**16**	**3.94**
21	Allison	Druin	UMD	**38**	**15.2**	**13**	**23**	**7.78**

22	Jennifer J.	Preece	UMD	**26**	**16.4**	**12**	**20**	**5.8**
23	Jimmy	Lin	UMD	**40**	**18.1**	**14**	**26**	**5.31**
24	Douglas W.	Oard	UMD	**71**	**9.35**	**15**	**24**	**5.03**
25	Dina	Demner-Fushman	UMD	**18**	**8.27**	**7**	**12**	**3.78**
26	Dagobert	Soergel	UMD	**30**	**8.1**	**8**	**15**	**2.38**
27	Bo	Xie	UMD	**15**	**4.2**	**4**	**7**	**1.97**
28	Eileen G.	Abels	UMD	**13**	**4.38**	**3**	**7**	**1.48**
29	Benjamin B.	Bederson	UMD	**15**	**11.9**	**7**	**13**	**1.12**
30	Lada Ariana	Adamic	MICH	**10**	**101**	**9**	**10**	**15.5**
31	Martha E.	Pollack	MICH	**32**	**31.8**	**16**	**31**	**9.71**
32	Paul	Resnick	MICH	**13**	**46.1**	**9**	**13**	**9.67**
33	Peter	Honeyman	MICH	**12**	**34.7**	**6**	**12**	**7.62**
34	Mark S.	Ackerman	MICH	**24**	**18.8**	**11**	**21**	**6.77**
35	Paul	Conway	MICH	**19**	**14.3**	**5**	**16**	**6.34**
36	Dragomir R.	Radev	MICH	**35**	**20.7**	**16**	**26**	**6.31**
37	Elliot	Soloway	MICH	**39**	**23.7**	**16**	**30**	**6.25**
38	Thomas	Finholt	MICH	**11**	**50**	**9**	**11**	**6.07**
39	Edmund H.	Durfee	MICH	**35**	**14.2**	**13**	**21**	**5.67**
40	Kevin	Crowston	SYR	**42**	**21.8**	**17**	**30**	**10.6**
41	Ping	Zhang	SYR	**30**	**15.9**	**13**	**21**	**8.53**
42	James	Howison	SYR	**12**	**30.3**	**10**	**12**	**7.77**
43	Dave	Molta	SYR	**15**	**16.9**	**2**	**15**	**6.65**
44	Joon S.	Park	SYR	**24**	**12.2**	**7**	**17**	**5.62**
45	R. David	Lankes	SYR	**17**	**16.4**	**7**	**16**	**5.55**
46	Elizabeth D.	Liddy	SYR	**27**	**14.6**	**9**	**19**	**5.23**
47	Lee W	McKnight	SYR	**19**	**13.7**	**7**	**16**	**5.05**
48	Zixiang	Tan	SYR	**16**	**10.3**	**4**	**12**	**4.84**
49	Martha A	Garcia-Murillo	SYR	**10**	**14.5**	**4**	**10**	**4.83**

Based on Table 5, we calculate the Community Productivity Index (CP-Index) for each of the iSchools, which are illustrated in Table 6. According to the CP-Index, the iSchool at the University of California at Berkeley is ranked the highest, followed by the iSchools of the *University of Michigan*, *Syracuse University*, *University of Pittsburgh*, and *University of Maryland*.

Table 6. Ranking of five research communities (iSchools) in decreasing order of their CP-Index

	School	**CP-Index**
1	UCB	9.3
2	MICH	9.2
3	SYR	8.9
4	PITT	7.8
5	UMD	6.9

In order to show how the different indices correlate, we calculate Spearman's correlation coefficient for each pair of seven performance measures. These measures are the number of publications (P), the number of citations (c), the average citations per paper (m), the h-Index (h), the g-index (g), and the RP-Index (RP). The data used for the calculation contains only data of 91 users (of the 1808 users within our database). All 91 users have more than 10 publications and more than 50 citations. All schools are equally represented. The data itself used for calculating the different indices is based on NC data (Eq. (1)). The results are illustrated in Table 7.

Table 7. Spearman's correlation coefficient between different measures (n=91)

	p	**c**	**m**	**h**	**r**	**g**	**RP**
p	1	0.55	-0.2	0.63	0.49	0.773	0.45
c		1	0.78	0.9	0.99	0.87	0.79
m			1	0.620	0.82	0.52	0.6
h				1	0.88	0.86	0.62
r					1	0.85	0.8
g						1	0.72
RP							1

The results show that RP-Index produces weaker correlation with measures, such as h-Index, r-Index, and g-Index, than the existing measures between themselves. That indicates that the presented RP-Index is significantly different to those indices.

6. Conclusion

The study introduces two new measures (RP-Index and CP-Index) for evaluating the productivity of researchers (knowledge creation) and the productivity of communities of researchers (knowledge creation of communities). The CP-Index is applicable to groups of researchers, departments, schools, colleges, and universities, or countries.

The indices constitute successors of the h-Index and g-Index. However, in addition to the number of publications and the number of citations, our indices consider performance criteria such as the number of co-authors and leadership. These indices decrease over time if individuals or communities are not productive anymore. Besides, our indices overcome the issue of imprecision of the h-Index and the g-Index, which result from the fact that they only deliver integer numbers. Our indices generate real numbers.

To show the workings of these new indices, we used data about publication activities of five information schools (iSchools), working in the area information management and systems at the Syracuse University, University of California at Berkeley, University of

Maryland, University of Michigan, and University of Pittsburgh. The results showed that the RP-Index produces significantly different results than the h-Index and g-Index. If the h-Index and g-Index has been applied to the normalized number of citations of a publication (NC), the difference is still significant as can be seen from the Spearman Correlation Coefficient. The comparison also shows that the RP-Index produces a ranking that differentiates the performance of authors in more detail.

Acknowledgements

The research presented has been funded by Ministry of Information and Communication of the Republic of Korea.

References

ACM Portal (2008), http://portal.acm.org/portal.cfm, October 2008.

Batista, P. D.; Campiteli, M. G.; Kinouchi, O. (2006): Is it possible to compare researchers with different scientific interests? Scientometrics, 68(1), pp. 179-189.

Braun, T.; Glänzel, W.; Schubert, A. (2005): A Hirsch-type index for journals. The Scientist, 19(22), pp. 8.

Burrell, Q.L. (2007): On the h-index, the size of the Hirsch core and Jin's A-index. Journal of Informetrics, 1(2), pp. 170-177.

CiteSeer (2008), http://citeseer.ist.psu.edu/, http://citeseerx.ist.psu.edu/, October 2008.

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

Glänzel, W. (2006a): On the h-index: A mathematical approach to a new measure of publication activity and citation impact. Scientometrics, 67(2), pp. 315–321.

Glänzel, W. (2006b): On the opportunities and limitations of the h-index. Science Focus, 1(1), pp. 10–11.

Google Scholar (2008), http://scholar.google.com/, October 2008.

Hirsch, J. E. (2005): An index to quantify an individual's scientific research output. Proceedings of the National Academy of Sciences, 102(46), pp. 16569-16572.

ISI Web of Knowledge (2008), Journal Citation Report from the ISI website, http://admin-apps.isiknowledge.com/JCR/JCR, October 2008.

Jin, BH. (2006): H-index: An evaluation indicator proposed by scientist. Science Focus (in Chinese), 1(1), pp. 8-9.

Jin, BH. (2007): The AR-index: complementing the h-index. ISSI Newsletter, 2007, 3(1), pp. 6.

Jin, BH.; Liang, LM.; Rousseau, R.; Egghe, L. (2007): The R-and AR-indices: Complementing the h-index. Chinese Science Bulletin, 52(6), pp. 855-863.

Klavans, R.; Boyak, K.W. (2008): Thought leadership: A new indicator for national and institutional comparison. Scientometrics, 75(2), pp. 239-250.

Kousha, K.; Thelwall, M. (2007): Google Scholar citations and Google Web/URL citations: A multi-discipline exploratory analysis. Journal of the American Society for Information Science and Technology, 58(7), pp. 1055–1065.

Lehmann S.; Jackson A. D.; Lautrup B. (2006): Measures for measure. Nature, 444, pp. 1003-1004.

Oh, W.; Choi J.N.; Kim, K. (2006): Coauthorship dynamics and knowledge capital: The patterns of cross-disciplinary collaboration in information systems research. Journal of Management Information Systems, 22(3), pp. 265-292.

Prathap, G. (2006): Hirsch-type indices for ranking institutions' scientific research output. Current Science, 91(11), pp. 1439.

Ruane, F.P.; Tol, R.S.J. (2008): Rational (successive) h-indices: An application to economics in the Republic of Ireland. Scientometrics, 75(2), pp. 395-405.

Scopus (2008), Elsevier, http://www.scopus.com/scopus/home.url/, October 2008.

Sidiropoulos, A.; Katsaros, D.; Manolopoulos, Y. (2007): Generalized h-index for disclosing latent facts in citation networks. Scientometrics, 72 (2), pp. 253-280.

Schubert, A. (2007): Successive h-Indices. Scientometrics, 70(1), pp. 201-205.

So, C.Y.K. (1998): Citation ranking versus expert judgment in evaluating communication scholars: Effects of research specialty size and individual prominence. Scientometrics, 41(3), pp. 325-333.

Tol, R.S.J. (2008): A rational, successive g-index applied to economics departments in Ireland. Journal of Informetrics, 2, pp. 149-155.

Weingart, P. (2005): Impact of bibliometrics upon the science system: Inadvertent consequences? Scientometrics, 62(1), pp. 117-131.