A comprehensive author ranking evaluation of network and bibliographic indices

Sahar Maqsood[1], Muhammad Arshad Islam[1,2]*, Muhammad Tanvir Afzal[1] and Nayyer Masood[1]
[1]Capital University of Science and Technology,
Islamabad, PAKISTAN
[2]National University of Computer and Emerging Sciences Science,
Islamabad, PAKISTAN
e-mail: Saharmaqsood2016@gmail.com; arshad.islam@nu.edu.pk*(corresponding author);
mafzal@custedu.pk; nayyer@cust.edu.pk

ABSTRACT

With the immense growth of scientific literature over the Web, the authors of research papers are being ranked for various purposes such as for being shortlisted for different scientific awards, for prestigious position, for tenured appointments, for keynote speaker invitation or for allocation of research grants. The traditional paradigm to attain these aspects is based on bibliometric indices, such as publication count, citation count and h-index. The quintessential indicator among all these indices are based on a number of citations received by the publication of an author. Generally, when a research paper is published, it receives citations after some time and may take more than a couple of years to attain a reasonable number of received citations which could make a difference in researcher ranking based on bibliometric parameters. Therefore, dependability over these bibliometric indices for authors ranking is not beneficial for researchers at the start of their academic career. Such authors are less likely to have a fair chance to compete with their senior counterparts. This study aims at overcoming the above-mentioned deficiency with the assistance of network centrality measures using the co-author network. The experiments are conducted to rank authors in order to identify the awardees of various scientific societies. The obtained ranking is further evaluated by comparing the results of bibliographic indices and network-based indices. The results revealed that network indices have equal potential to identify influential authors as compared with bibliometric indices such as h-index, particularly in the case of networks having a high value of cluster coefficient. Based on the obtained results, it is observed, awardees that are exclusively identified by network centrality measures are more than 5 years younger as compared to their counterparts who are exclusively identified by bibliometric indices.

Keywords: Co-authorship network; Citation analysis; Centrality measures; Citation-based indices; Network-based indices

INTRODUCTION

The idea of collaboration among authors is not unique and researchers have been collaborating since the early 19th century (Beaver and Rosen 1978). The primary objectives of joint research production are to complete the scientific study within a minimum span of time and produce a comprehensive output by a mutual effort (Erfanmanesh, Rohani and Abrizah 2017) . Usually, authors having a similar research interest, collaborate with each other to produce a novel piece of work (McCarty et al. 2013). Collaboration among authors from different backgrounds could be beneficial to make significant improvements in the corresponding fields (Bridle et al. 2013).

https://doi.org/10.22452/mjlis.vol25no1.2

Behind this, there is an immense effort of researchers for which they must be acknowledged. At the same time, this vast plethora of scientific studies has made it crucial to track the contributions of an individual author (Asif and Islam 2016). In the current state-of-the-art, efforts of researchers are being acknowledged by adopting different ranking mechanisms wherein different parameters of authors are considered to determine their academic influence over other researchers. Researchers are ranked to identify the potential authors eligible for research funding or grants, to supervise and coordinate industrial megaprojects, to be invited as keynote speakers, to serve as editors for journals, or to offer tenured positions in any organization (Petersen, Wang and Stanley 2010) . The traditional paradigm of author ranking mechanisms relies on certain bibliographic indices such as publication count (Ghani et al. 2019), citation count (Moreira, Calado and Martins 2015) and *h*-index (Hirsch 2005).

The authors having the highest number of publications or citations are deemed as potential research contributors (Dunaiski, Visser and Geldenhuys 2016). Jorge Hirsch proposed *h*-index in 2005 to measure the research contribution of an individual author (Hirsch 2005). *h*-index a ranking measure that couples both publication and citation count. Some researchers have considered *h*-index as one of the best measures among existing ones, while others presented different extensions of *h*-index to overcome its existing deficiencies. These extensions include *g*-index(Glänzel 2006), A-index, (Alonso et al. 2009) l-index and Ar-index (Jin et al. 2007).

However, most of these indices employ citation count-based measures. Citations take a span of two to three years to be received by the publication (Ottaviani 2016). It can be inferred that older research articles may have more citations as compared to the recent ones. Therefore, authors who are at the start of their academic career suffer from the inherent disadvantage of citation dependent metrics (Dorta-Gonzalez and Dorta-Gonzalez 2013) . These issues hinder to conduct equitable comparison among the researchers irrespective of their experience.

Centrality measures are considered a very vital tool in the field of network analysis to identify a central node that can play a key role in disseminating information in the network. A network or a graph G=*(V, E)* is represented as a set of vertices *V* and edges *E*. A co-author network is formed considering authors as nodes and two author nodes that have collaborated with each other are considered adjacent. Figure 1 shows co-author network that contains 7 authors represented by vertices $V = \{A,B,C,D,E,F,G\}$ containing 10 undirected edges specifying a co-authorship relation among them.

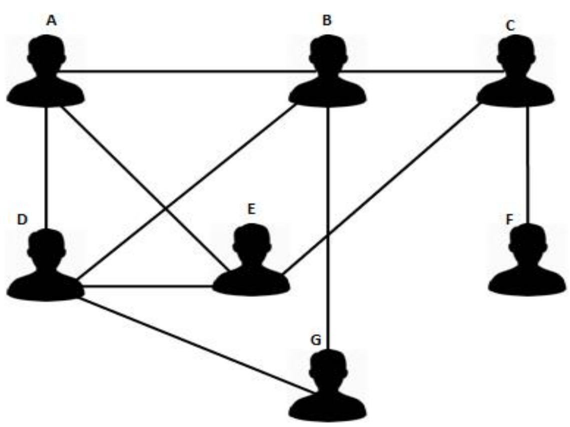

Figure 1: Co-author Network

The authors who have just commenced their academic career should also be acknowledged so that they can avail appraising awards to further excel in the career. Based on this idea, this paper evaluates the ranking mechanism obtained using network-based centrality indices and citation-based (bibliographic) indices. The primary concern of this work is to investigate how accurately centrality indices can rank the authors ranked by bibliometric indices. The motivation for using network-based indices is to provide a fair chance to such novice authors, publications of whom have yet to receive the due number of citations; however, the quality of their research work makes them quite eligible to be considered for prominent positions. The study argues that an author is likely to have an important position in a co-author network if he has published quality research articles. In other words, two authors have the same publication and citation count can be ranked differently if one of them has co-authored with influential researchers. Network centrality measures are commonly harnessed to identify influential vertices from a given network (Freeman 1978). These measures identify important vertices using various, aspects such as the number of adjacent edges, using the number of shortest paths passing through them, and average distance of the network.

In this study, a co-author network of researchers belonging to the Mathematics domain has been constructed. To find the influential author, four centrality measures are employed i.e.; degree, closeness, betweenness, and PageRank. To evaluate the proposed scheme, the obtained author ranking is correlated with the Awardees from four prestigious societies. The results revealed that network indices have the potential to identify the same influential authors as identified by bibliographic indices. It is observed that awarding societies implicitly depend upon the betweenness centrality measure, which indicates that awarding societies tend to award those authors who have multidisciplinary scope for their research. Moreover, obtained results agree with the scheme presented in Adali, Lu and Magdon-Ismail (2013) , that centrality indices can provide competitive ranking when authors within one cluster are considered.

RELATED WORK

The rapid growth of scientific literature has created a hustle for the scientific community to find the expertise of researchers while using different citation indices. Various researchers have proposed their own qualitative and quantitative parameters to find the highly ranked authors in the different fields of study such as mathematics, medical sciences, social sciences, management science, engineering technology and computer science.

Individual authors, scholars and researchers can be ranked by measuring the impact of their publications. Indices such as the author's publication count, citation count, mean and median citation numbers are included in bibliometric indices (Moreira, Calado and Martins 2015). Mean citation distribution becomes highly skewed which is not satisfactory and the median citation produces a very long tail. To overcome these limitations, *h*-index has been introduced (Hirsch 2005). Apart from *h*-index, publication count and citation count, researchers harness different variants of *h*-index such as *g*-index and m-quotient to measure the scientific contributions of authors (Bornmann et al. 2011) . Author level eigenvector and author impact factor are also used to rank the authors (Alarfaj et al. 2012).

Co-author networks have widely been used to identify the key researchers and to extract collaboration patterns. Sarigöl et al. (2014) have acknowledged the influence of social structure in scientific literature, arguing that the centrality of highly cited authors significantly varies from authors with lower citation count. They have considered 2-year sliding windows in order to compute four centrality metrics. Their results show that authors at the start of their career, tend to gain high citation count swiftly when they collaborate with influential co-authors in a co-author network. Additionally, they have elaborated the predictive power of the centrality metrics for the citation count of their articles. Co-author networks are usually extracted from bibliographic records (Biryukov 2008), or large scientific databases such as MEDLINE, the Los Alamos National Library and Networked Computer Science Technical Report Library (NCSTRL). One of the earlier examples of the co-author network consists of 511 mathematicians who have collaborated with a renowned mathematician, Paul Erdös (Castro and Grossman 1999). Gang et al. (2015) have modified the PageRank algorithm as the LeaderRank algorithm that considers Paul Erdös' co-author network and helps a researcher to find his influence in the scientific community. Li et al. (2014) have modified Katz Bonacich centrality to define the authors' prestige, which uses the idea of PageRank algorithm. To find the influential authors, each author's influence score is calculated using a co-author network of Paul Erdös. Another study also evaluated the two parameters used for computing Bonacich centrality; α provides a priori information related to authors and β can provide information related to the robustness of the results in case of the incomplete network (Hicks et al. 2019).

Collaboration behaviour of authors using three metrics, i.e., the number of co-authors, the structure of collaboration, characteristics of co-authors has been investigated (Jinsong et al. 2019). Contrary to the hypothesis, network structure failed to provide predictions related to *h*-index, however, it was concluded that the highest *h*-index could be achieved by working with many high valued *h*-index authors (McCarty et al. 2013). Apart from measuring the author's impact by using his number of citations, Ausloos (2013) has proposed a scheme to rank the number of co-authors according to the number of joint publications. Results showed that a strong association exists between joint publications count of the co-authors and their citations. Franceschet (2011) has analyzed the evolution of collaboration and affiliation network to conclude that papers with authors belonging to multiple affiliations attract more citations than papers with authors to a single affiliation.

In summary, various studies have proposed around 40 variants of *h*-index and all of them are citations dependent (Bornmann et al. 2011). Such indices require a period of 2-3 years until a paper reaches its prime time to gain citations (Ottaviani 2016). From the literature, few examples of co-author network-based ranking are also delineated, which have effectively correlated the network-based metrics with citation-based metrics. However, there should be a ranking mechanism that does not depend on citations so that authors who are at the early phase of their career can be ranked fairly along with their senior counterparts. This study aims to investigate network-based indices for authors ranking that have not been given due attention in the literature.

RESEARCH QUESTIONS

The aforementioned gaps in the current state-of-the-art have led us to scrutinize these research questions:

(a) Which of the bibliographic indices (citation count, *h*-index, *g*-index) contributes most to bring the international awardees in the list of top-ranked authors?

(b) Which of the co-author network indices (degree, closeness, betweenness, and PageRank) contributes most to bring the international awardees in the list of top-ranked authors?

In order to formulate the next research question, mathematical illustration has been employed to convey the concept and reasoning effectively. Suppose, most contributing index obtained from bibliographic indices on top is labeled as C_I and the most contributing index obtained from network indices is labeled as G_I. Set of awardees on top rank using C_I are known as C_A and the set of top-ranked awardees obtained using G_I are known as G_A. To evaluate the obtained ranking lists, the top 50, top 100, top 200, top 500, top 1000, and top 5000 authors have been considered from the dataset. The symbols are described in Table 2.

Table 2: Description of Mathematical Symbols

Symbol	Description
C_I	*Bibliometric Index that identifies the maximum awardees*
C_A	Set of awardees identified by C_I
G_I	Network Index that identifies the maximum number of awardees
G_A	Set of awardees identified by G_I

(c) What percent of G_A is closer to the C_A?
(d) Which mathematical awarding society depends more on the co-author network indices?
(e) How much difference is there between the work experience of awardees identified by bibliographic and graph-based metrics?

METHOD

In the previous studies, several types of schemes have employed network-based indices to find the most influential authors from the co-author network and ranked them within that network (Zhang, Liu and Lu 2014). The power of centrality metrics has been acknowledged by evaluating their capability to identify the awardees of prestigious scientific societies. Fiala and Tutkoy (2017) have shown that PageRank based ranking has outperformed citation-based ranking to identify the awardees of Turing Awards. Similarly, Dunaiski and colleagues have examined multiple datasets and ranked the awardees and their research articles using citation-based and network-based metrics. The results evaluated with the help of the median rank of 249 awardees show that PageRank related ranking algorithms have achieved the best results (Dunaiski, Visser and Geldenhuys 2016).

This study scrutinizes the network-based indices; degree, closeness, betweenness and PageRank, to analyze to what extent these indices can behave similar or closer to the citation-based indices. The awards given by the leading scientific societies are considered for evaluation due to the absence of a benchmark (Ayaz and Afzal 2016). In this study, data of 24 prestigious awards and 671 awardees in mathematics field were collected. A brief description related to each society and number of awardees is presented in Table 1. Details of each award and names of the awardees are available at http://cdsc-cust.org/research/scientometrics/. The awarding societies considered for evaluation in this paper are:

(a) American Mathematics Society (AMS): An association of professional mathematicians established according to the interest of mathematical research and scholarship. It serves the national and international community through its publications, meetings and other programmes.

(b) International Mathematics Union (IMU): An international scientific organization that promotes international cooperation in mathematics. The objectives of this society are to promote international cooperation in mathematics, to support the scientific meetings or conferences and to contribute to all sub-branches of mathematics.
(c) London Mathematics Society (LMS): The UK learned mathematics society. It publishes journals, books and provides grants for the promotion of mathematics field, and scientific meetings and lectures.
(d) Norwegian Academy of Science and Letters (NASL): NASL is not a domain-specific society. It collaborates with all fields of study. The most prestigious awards of this society are small in number as compared to other prestigious awarding societies.

Table 1: Number of Awards and Awardee given by each Society

Society	Awards	Awardees
AMS	10	296
IMU	5	72
LMS	8	289
NASL	1	14
Total	24	671

Figure 2 presents the organizational framework utilized to evaluate the two ranking indices. The steps are delineated in the following sub-sections.

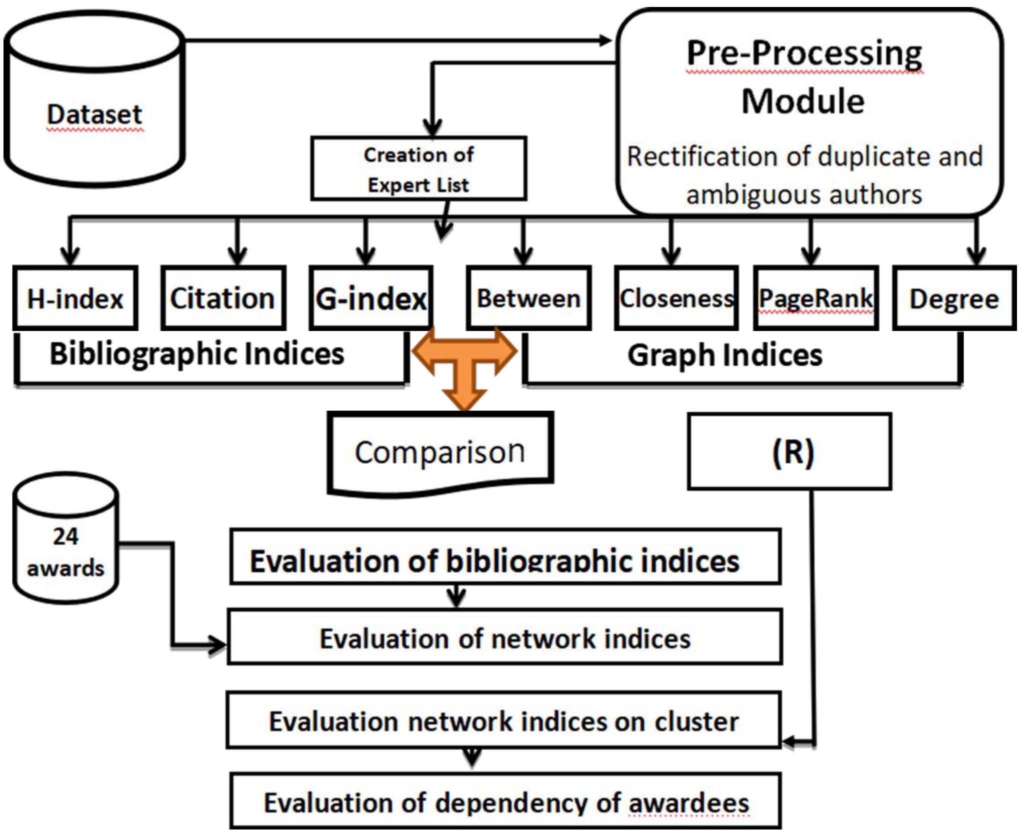

Figure 2: Organizational Framework to Evaluate the Two Ranking Indices

Selection and Pre-processing of Data

A data set comprising research articles from the mathematics domain has been utilized for this work. These articles are categorized using MSC (Mathematics Subject Classification) by two major mathematical reviewing databases, I.e Mathematical Reviews and Zentralblatt MATH. Several mathematics journals (such as *Conformal Geometry and Dynamics, Journal of the American Mathematical Society* and *Mathematics of Computation*) request the authors of research papers and expository articles to list subject codes from the Mathematics Subject Classification in their papers to assist the readers by providing them feasibility to retrieve the relevant content. The related extraction details are available in Ayaz and Afzal (2016).

There are 57,533 authors and among those 29,263 who have the same last name, have been disambiguated in Ayaz and Afzal (2016) . Whereas 9,403 authors do not have a co-author relationship, therefore the data related to those authors are omitted from the data set. This shows that there are 1.09 authors per paper which is relatively a low co-authorship ratio. Moreover, edge-list from this source has been extracted to construct an unweighted network. The biggest cluster from the resultant clusters has 23,903 authors and 15,602 publications. Other relevant details regarding the dataset are presented in Table 3.

Table 3: Summary of Dataset

	Original Dataset	**Co-authors dataset**	**The largest component in the dataset**
Authors	57,533	48,130	23,903
Publications	62,033	52,630	44,268
Radius	-	15	13
Diameter	-	26	20
Clustering Coefficient	-	0.2004	0.25486

Computation of Bibliographic Ranking Indices and Network Indices

The next step involves the computation of bibliographic ranking indices and network indices lists that are calculated by two different tools. To compute the ranking lists from traditional ranking indices (*g*-index, *h*-index), the total number of citations of all publications is calculated. To compute the ranking lists for co-author network indices, the igraph (Csardi and Nepusz 2006) library, that is supported by a widely used statistical tool (R), is used. An unweighted, undirected co-author network is constructed in R for computation of network centrality measures.

The following traditional bibliographic (citation-based) indices for ranking are used:

- Citation count: Citations have their own importance to rank the authors.
- *h*-index: It is a scientific measure, which is calculated by considering both the number of publications of an author and the number of citations. A scientist has index *h* if *h* of his/her papers have at least h citations each, and the other (Np−*h*) papers have no more than *h* citations each (Kelly and Jennions 2006).
- *g*-index: Similar to *h*-index, *g*-index is an author level metric (Burrell 2008) . It is used to measure the importance of top articles of authors. A scientist has index *g* if *g* is the largest integer such that his or her top *g* papers received together at least g^2 citations (Woeginger 2009).

The centrality measures are employed to identify the important, prominent, focal, and gatekeepers authors in a co-author network (Liu, Sidhu and Beacom 2017). Network analysts have varied opinions on how network centrality should be measured (Borgatti and Johnson 2013). Four commonly used network centralities are *degree, betweenness, closeness,* and *PageRank*. These network indices are described below.

- Degree**:** In a co-author network, degree refers to the number of edges connected to the selected author. It indicates the influential author based on the number of connected authors (Sarigoel et al. 2014).

$$\mathbf{C_D(n_i) = deg(n_i)} \qquad (1)$$

  In equation (1), $C_D(n_i)$ refers to the degree centrality of n_i representing the number of authors with whom n_i has co-authored research papers.
- Closeness**:** Closeness centrality focuses on geodesic distance between authors in a co-author network. It incorporates the distance that information from one author has to travel to reach another author (Abbasi, Altmann and Hossain 2011).

$$\mathbf{C_C(n_i) = \sum_{j=1}^{N} 1/_{d(n_i,n_j)}} \qquad (2)$$

  Equation (2) refers to the distance of an author to all other authors in a co-author network. $C_c(n_i)$ represents the closeness of an author n_i and $d(n_i,n_j)$ represents the distance between an author n_i and n_j (Erfanmanesh, Rohani and Abrizah 2017).
- Betweenness centrality: Betweenness is measured by counting the number of instances an author acts as a bridge between two other authors on the shortest path. Authors with high betweenness are deemed as experts in a co-author network in terms of knowledge (Abbasi, Altmann and Hossain 2011).

$$\mathbf{C_B(n_i) = \sum_{j,\,k\neq i} {}^{g_{ijk}}/_{g_{jk}}} \qquad (3)$$

  In equation (3), $C_B(n_i)$ refers to betweenness of n_i whereas g_{ijk} represents the numbers of the shortest paths between author n_j and n_k that contains n_i and g_{jk} represents the number of all paths between author n_i and n_j.
- PageRank centrality: The PageRank is proposed by (Brin and Page 1998) , which identifies important web pages. In the context of the co-author network, an author is said to be influential if he/she is associated with other influential authors.

$$\mathrm{PR(n_i)} = {}^{(1-d)}/_{N} + \mathrm{d.}\sum_{\mathrm{P_j}\,\in \mathrm{M(n_i)}} {}^{\mathrm{PR(P_j)}}/_{\mathrm{L(P_j)}} \qquad (4)$$

  In equation (4), N is the total number of authors in a co-author network. D is the damping factor that is used to control the behaviour of PageRank. Traditionally $d = 0.85$ is considered the default value.PR(p_i) is the PageRank of the author. *L(p_j)* is the number of outgoing edges from the author p_j and *M(p_i)* is the set of PageRank of the rest of the authors.

To evaluate the results, the research questions are addressed by considering both indices, i.e. the traditional citation indices and the commonly used network indices.

Awardees Evaluation in Authors' Ranking Lists

After acquiring the ranking lists, the presence of awardees is verified to ensure whether the awardees exist in the ranking lists or not. For this purpose, the dataset is divided in the form of percentage and the authors are searched in the distribution of the top 10%, 20%, 30% and so on.

Evaluation of Awarding Societies' Dependency on Network Indices

The next step involves the investigation of the dependency of prestigious awardees of mathematics on co-author network indices, which also addresses one of our research questions. The same percentages of authors as described earlier are considered to perform the evaluation. The results of this evaluation are detailed out and discussed in the next section.

RESULTS AND DISCUSSIONS

(a) The bibliographic index (citation count, *h*-index, *g*-index) that contributes most to bring the international awardees in the list of top-ranked authors

The performance of each bibliographic index is shown in Figure 3 to analyze the presence of awardees in top rankings. The ranking list is divided into ascending order slices of Top 50, Top 100, Top 200, Top 500 and Top 1000 authors to observe the ranking behaviour of each index. *h*-index has identified most authors in all slices whereas there exists a variation in behaviour of citation count and *g*-index. Up till the Top 200, citation count has identified more awardees as compared to *g*-index and this behaviour is reversed for remaining slices. These results show that awarding societies considered in this work are implicitly more dependent on *h*-index as compared to other bibliographic indices. As discussed earlier, *h*-index considers both publication count and citation count in tandem.

(b) The co-author network index (degree, closeness, betweenness, and PageRank) that contributes most to bring the international awardees in the list of top-ranked authors

The performance of each network index is illustrated in Figure 4 to depict the presence of awardees in top rankings. Among the network indices, betweenness centrality has outperformed all its counterparts. Betweenness centrality ranks those nodes higher that interlink communities, i.e., that are part of the maximum number of the shortest path. It can be deduced that awarding societies tend to award those authors who have contributed to multiple domains. These authors act as bridges between different communities of authors who work on any particular research area. This finding is in line with the press release by the London Mathematical Society[1].

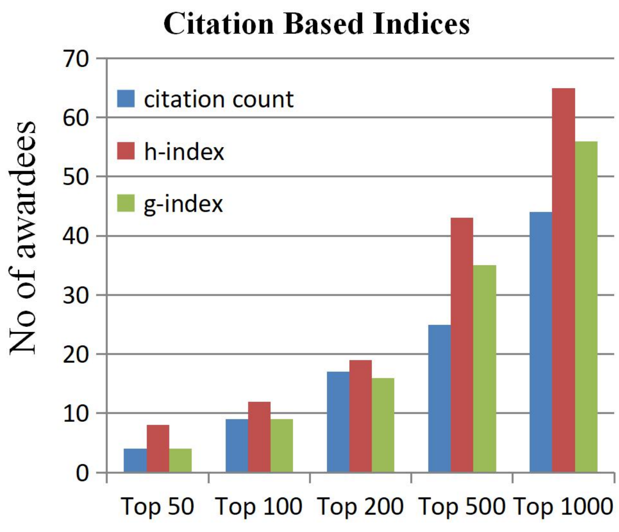

Figure 3: Awardees from Bibliographic Indices

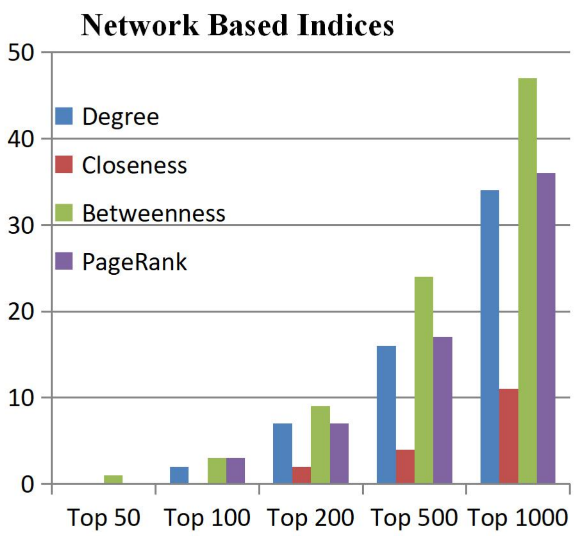

Figure 4: Awardees from Network Indices

The behaviour of closeness remained consistent on the downside, as it has identified the least number of awardees among the four network indices. The performance of degree

[1] https://www.lms.ac.uk/prizes/citations-lms-prize-winners

and PageRank centrality is comparable; both have identified more or less the same number of awardees. In general, we observed a mean correlation up to 0.68 when the top 100 authors list obtained by h-index and betweenness centrality is considered.

(c) Percent of G_A closer to the C_A

As shown in Figure 3 and Figure 4, *h*-index has outperformed in identifying the awardees of all societies considered in this study. However, it takes a considerable span of time that the author receives the citations for his publication. A paper may take a few years to obtain a significant number of citations, that will be reflected in his/her *h*-index after some time. Therefore, it is necessary to analyze, whether co-author network indices that are temporally independent, can be used to identify the awardees of the society.

Another experiment, considering the biggest cluster in the network has been performed. For this experiment, two indices that have produced the best result in their categories are considered, i.e., betweenness from network-based indices and h-index from citation-based indices.

A network cluster (community) is a grouping of network nodes that contain more edges among the cluster nodes. In order to compute the clusters in the co-author network, the algorithm proposed by Clauset, Newman and Moore (2004) has been used. We have analyzed several medium to small clusters in the graph and extracted the biggest cluster having approximately 28,000 authors. Afterward, betweenness centrality is computed for the clustered network. As shown in Figure 5, it can be observed that the disagreement with respect to the number of awardees identified by both betweenness centrality and *h*-index gradually reduces as the number of authors included in the list increases. The corresponding result in percentage is shown in Table 4.

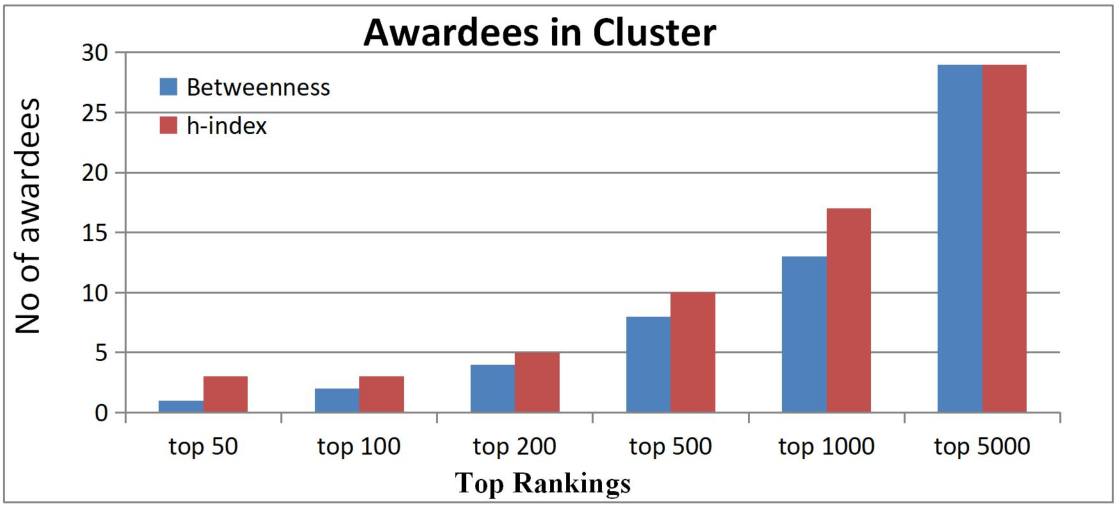

Figure 5: Comparison of Betweenness and *h*-index on Clustered Network

As discussed earlier, the betweenness centrality of a node represents the number of shortest paths that pass through that node. If authors, at the start of the career, have placed themselves in such a position that a significant number of shortest paths pass through their node, then it can be deduced that those authors have the potential to gain a high *h*-index particularly in the case of networks with high clustering coefficients.

Table 4: Agreement in Percentage of Awardees identified by Betweenness and *h*-index

	Agreement in Un-Clustered Network	**Agreement in Clustered Network**
Top 50	0.25	0.33
Top 100	0.33	0.33
Top 200	0.47	0.66
Top 500	0.55	0.8
Top 1000	0.71	0.76
Top 5000	0.90	1

The clustering coefficient is relatively low for the dataset considered in this study due to the long-time span of the dataset, i.e., more than 50 years. The clustering coefficient is also dependent on the field of the authors in a co-author network. Authors belonging to the same field tend to collaborate more than the other fields (Gazni and Didegah 2011) . Additionally, the trend of collaboration among authors is being increased in the recent times (Wu, Venkatramanan and Chiu 2015). Thus, betweenness centrality measure holds the potential to some extent in the identification of influential authors even if they are at the early stage of their careers.

(d) The mathematical awarding society that is more dependent on the co-author network indices

The awarding criteria of these societies are not publicly known. Therefore, through analysis of this research question, the implicit dependence of each society on a particular index can be identified. The dependence of each society on indices is shown in Figure 6. This dependency is investigated by computing the percentage of occurrence of awardees with respect to each awarding society. For this purpose, the results of the top 10% of the ranking list are considered to measure the dependency of awarding society upon network and bibliographic indices. For AMS and LMS awardees, *h*-index has performed well among other measures. For IMU awardees, the number of publications and *g*-index have performed well and for NASL, the number of citations has performed well with a significant margin. The following observations exhibit the contribution of each index.

In the case of AMS, from network indices, betweenness centrality has outperformed all other network indices by identifying 48 percent of the awardees in the top 10%. From bibliographic indices, *h*-index has performed well by identifying 57 percent awardees on top 10%. Degree and PageRank have performed almost in a similar manner by identifying 36 percent of the awardees from the top 10% of the ranking list. In bibliographic indices, citations and *g*-index have performed almost similar by identifying approximately 53 percent of authors. Whereas, the performance of closeness is low as compared to all other indices.

Degree, betweenness centrality and PageRank behaved similarly to bring the awardees on top with 47 percent in top 10% ranking lists for IMU. The performance of closeness remained low for all societies. The contribution of all non-bibliographic indices remained equal and have brought almost 55 percent of the authors in the top 10% of the ranking lists.

In the case of LMS, betweenness centrality has performed better than other bibliographic indices. It has identified almost 37 percent awardees in the top 10%. Closeness centrality has consistently performed worse. In non-network indices, the performance of citations is low as compared to *h*-index, *g*-index, and publication count.

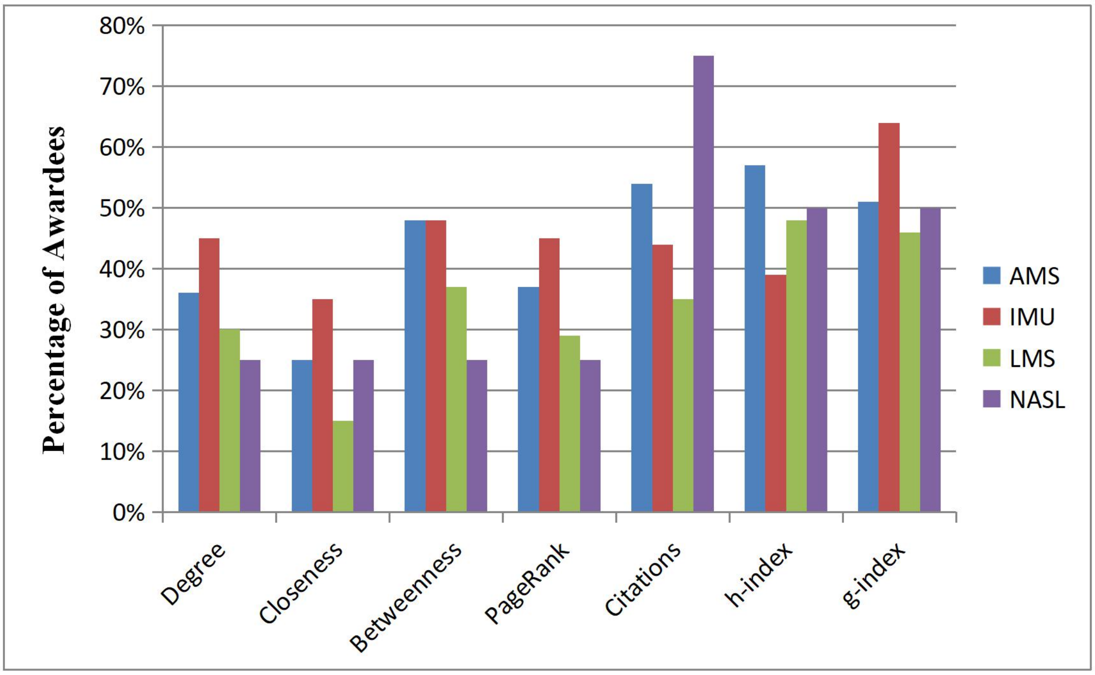

Figure 6: Dependency of Awardees on All Indices

(e) The age difference between the awardees identified by bibliographic and graph-based metrics

To answer this question, the top 10% of awardees obtained by *h*-index and betweenness centrality are examined because these two metrics have identified the maximum number of awardees in their respective categories. When the top 10% of the awardees identified by *h*-index are considered, *h*-index is able to identify awardees of only AMS. However, betweenness has been able to identified awardees of all selected 4 societies. Among the AMS awardees, 21 awardees are exclusively identified by *h*-index, while 16 awardees are exclusively identified by betweenness centrality. To evaluate the awardees' age, the difference between the year an author was awarded and the year of first publication of the corresponding author is calculated. The mean delay of more than 20 years was observed for the awardees identified by *h*-index. Whereas, the corresponding mean delay of less than 14 years for awardees identified by betweenness centrality is observed. Among the awardees identified by both metrics, biographies of approx. A total of 91 percent of the awardees are found on Wikipedia (www.wikipedia.org) containing information about their ages. Authors identified by betweenness centrality are on average 8.4 years younger than the authors identified by *h*-index. Both indicators show that there is a significant difference in the two evaluated metrics to infer that relatively younger authors have been identified by betweenness centrality as compared to *h*-index.

CONCLUSION

The in-depth analysis of experiments and evaluation revealed that network indices hold the potential to rank the authors with respect to their position in the co-author network. Among all other network indices, betweenness centrality has significantly performed well in ranking the authors and specified relatively strong association with the awardees of scientific societies. This association increased when the biggest cluster from the co-author network is considered for betweenness centrality. The centrality metrics are potential indices to effectively identify the authors that have contributed significantly to the community but have not yet received substantial citations to be identified by bibliometric indices. The limitation of the presented work is that solo authors cannot be ranked via analysis of the co-author network. However, the collaboration trend is gradually being increased (Sarigoel et al. 2014) and there would be more studies on collaborative research in the future wherein the proposed network-indices based ranking mechanism would definitely assist to rank the novice researchers. According to the results, low correlation coefficient likely to produce insignificant outcomes. These outcomes are centralities dependent which relies on dense co-author network. Hence in order to evaluate centrality-based measures along with bibliometric measures, we must opt with the dataset containing a high correlation coefficient.

In the future, we intend to stipulate the significance of the weighted network, i.e., network edges containing weights corresponding to the number of papers the authors have co-authored. Additionally, an important aspect of the position of an author in research publication should also be considered to obtain accurate ranking.

ACKNOWLEDGEMENT

No grant from any public, commercial, or non-profit funding agency was offered for the conduct of this research. We thank the reviewers for their enriching comments to improve the article.

REFERENCES

Abbasi, A., Altmann, J. and Hossain, L. 2011. Identifying the effects of co-authorship networks on the performance of scholars: A correlation and regression analysis of performance measures and social network analysis measure. *Journal of Informetrics,* Vol. 5, no.4: 594-607.

Adali, S., Lu, X. and Magdon-Ismail, M. 2013. Deconstructing centrality: thinking locally and ranking globally in networks. In Proceedings. *International Conference on Advances in Social Networks Analysis and Mining.* ACM. 418-425.

Alarfaj, F., Kruschwitz, U., Hunter, D. and Fox, C. 2012. 2012. Finding the right supervisor: expert-finding in a university domain. *Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: Student Research Workshop.* Stroudsburg: ACM. 1-6.

Alonso, S., F. J. Cabrerizo, E. Herrera-Viedma, and F. Herrera. 2009. h-Index: A review focused in its variants, computation, and standardization for different scientific fields. *Journal of Informetrics,* Vol. 3, no. 4: 273-289.

Asif, R., and Islam, M. A. 2016. Finding most collaborating mathematicians a co-author network analysis of mathematics domain. *International Conference on Computing, Electronic and Electrical Engineering (ICE Cube).* Quetta: IEEE. 289-293.

Ausloos, M. 2013. A scientometrics law about co-authors and their ranking: the co-author core. *Scientometrics,* Vol. 95, no. 3: 895-909.

Ayaz, S. and Afzal, M.T. 2016. Identification of conversion factor for completing-h index for the field of mathematics. *Scientometrics,* Vol. 109, no. 3: 1511-1524.

Beaver, D. and Rosen, R. 1978. Studies in scientific collaboration: Part I. The professional origins of scientific co-authorship. *Scientometrics* 1 (1): 65-84.

Biryukov, M. 2008. Co-author network analysis in DBLP: Classifying personal names. In: Le Thi H.A., Bouvry P., Pham Dinh T. (eds) *Modelling, Computation and Optimization in Information Systems and Management Sciences.* MCO 2008. Communications in Computer and Information Science, Vol 14. pp 399-408. Springer, Berlin, Heidelberg.

Borgatti, S. P., Martin, G. E., and Johnson, J. C. 2013. *Analyzing social networks.* SAGE Publications Limited.

Bornmann, L., R. Mutz, S. E. Hug, and H. D. Daniel. 2011. A multilevel meta-analysis of studies reporting correlations between the h index and 37 different h index variants. *Journal of Informetrics* Vol. 5, no.3: 346-359.

Bridle. H., Vrieling, A., Cardillo, M., Araya Y. and Hinojosa, L. 2013. Preparing for an interdisciplinary future: A perspective from early-career researchers. *Futures,* Vol. 53, no.1: 22-32.

Brin, S. and L. Page. 1998. The anatomy of a large-scale hypertextual web search engine. *Computer Networks and ISDN Systems,* Vol. 30, no.1: 107-117.

Burrell, Q. L. 2008. On Hirschś h, Eggheś g and Kosmulskiś h(2). *Scientometrics,* Vol. 79, no.1: 79-91.

Castro, R. D., and J. W. Grossman. 1999. Famous trails to Paul Erdos, Math. *Intelligencer* Vol. 3, no.21: 51-63.

Clauset, Aaron and Newman, Mark EJ and Moore, Cristopher. 2004. Finding community structure in very large networks. *Physcial Review E,* Vol. 70, no.6: 066111.

Csardi, G., and Nepusz, T. 2006. The igraph software package for complex network research. *InterJournal, Complex Systems* Vol. 1695, no.5: 1-9.

Dorta-Gonzalez, P. and Dorta-Gonzalez, M. I. 2013. Impact maturity times and citation time windows: The 2-year maximum journal impact factor. *Journal of Informetrics* Vol. 7, no.3: 593-602.

Dunaiski, M., Visser, W. and Geldenhuys, J. 2016. Evaluating paper and author ranking algorithms using impact and contribution awards. *Journal of Informetrics,* Vol. 10, no.2: 392-407.

Erfanmanesh, M., Rohani, V.A. and Abrizah, A. 2017. Co-authorship network of scientometrics research collaboration. *Malaysian Journal of Library & Information Science,* Vol. 17, no.3: 73-93.

Franceschet, M. 2011. Collaboration in computer science: a network science approach. *Journal of the American Society for Information Science and Technology,* Vol. 62, no.10: 1992-2012.

Freeman, L. C. 1978. Centrality in social networks conceptual clarification. *Social Networks,* Vol. 1, no.3: 215-239.

Gang, J., F. Liu, Z. Zhang, and S. Li. 2015. The application of improve authorś personal influence of network ranking algorithm. *Journal of Applied Science and Engineering Innovation,* Vol. 2, no. 4: 102-105.

Gazni, A. and Didegah, F. 2011. Investigating different types of research collaboration and citation impact: a case study of Harvard University's publications. *Scientometrics,* Vol. 87, no.2: 251-265.

Ghani, R., Qayyum, F., Afzal, T. M. and Maurer, H. 2019. Comprehensive evaluation of h-index and its extensions in the domain of mathematics. *Scientometrics,* Vol. 118, no.3: 809-822.

Glänzel, W. 2006. On the h-index-A mathematical approach to a new measure of publication activity and citation impact. *Scientometrics,* Vol. 67, no.2: 315-321.

Hicks, D. J., Coil, D. A., Stahmer, C. G. and Eisen, J. A. 2019. Network analysis to evaluate the impact of research funding on research community consolidation. *PloS One,* Vol. 14, no. 6: e0218273.

Hirsch, J. E. 2005. An index to quantify an individualś scientific research output. *Proceedings of the National academy of Sciences of the United States of America.* 16569-16572.

Jin, B., Liang, L., Rousseau, R. and Egghe, L. 2007. The R-and AR-indices: Complementing the h-index. *Chinese science bulletin,* Vol. 52, no.6: 855-863.

Jinsong, Z., Xue, Y., Xuan, H. and Li, T. 2019. Author cooperation network in biology and chemistry literature during 2014--2018: Construction and structural characteristics. *Information,* Vol. 10, no. 7: 236.

Kelly, C.D, and Jennions, M.D. 2006. The h index and career assessment by numbers. *Trends in Ecology & Evolution,* Vol. 21, no. 4: 167-170.

Li, Y., C. Wu, X. Wang, and P. Luo. 2014. A network-based and multi-parameter model for finding influential authors. *Journal of Informetrics* ,Vol. 8, no.3: 791-799.

Liu, W., Sidhu, A. and Beacom, A.M. 2017. Social network theory. In *The International Encyclopedia of Media Effects*, by Patrick Rössler, 12. John Wiley & Sons.

McCarty, C., J. W. Jawitz, A. Hopkins, and A. Goldman. 2013. Predicting author h-index using characteristics of the co-author network. *Scientometrics,* Vol. 96, no.2: 467-483.

Moreira, C., P. Calado, and B. Martins. 2015. Learning to rank academic experts in the DBLP dataset. *Expert Systems,* Vol. 32, no. 4: 477-493.

Ottaviani, J. 2016. The post-embargo open access citation advantage: It exists (probably), it's modest (usually), and the rich get richer (of course). *PLoS ONE,* Vol. 11, no. 10:: e0165166. Available at: https://doi.org/10.1371/journal.pone.0165166.

Petersen, A.M., Wang, F. and Stanley, H.E. 2010. Methods for measuring the citations and productivity of scientists across time and discipline. *Physical Review E,* Vol. 81, no. 3: 1539-3755.

Sarigöl, E., Piftzner, R., Scholted, I., Garas, A. and Schweitzer, F. 2014. Predicting scientific success based on coauthorship networks. *EPJ Data Science,* Vol. 3, no.1: 9.

Woeginger, G.J. 2009. Generalizations of Egghe's g-index. *Journal of the American Society for Information Science and Technology,* Vol. 60, no.6: 1267-1273

Wu, Y., Venkatramanan, S. and Chiu, D.M. 2015. "Research collaboration and topic trends in computer science: An analysis based on UCP authors." *Proceedings of the 24th International Conference on World Wide Web.* Florence: ACM. 1045-1050.

Zhang, L.X., Liu, Y.J. and Lu, X,Z. 2014. Using networks to measure influence and impact. Vol. 556–562, Trans Tech Publications, Ltd., May 2014, pp. 2668–2671. Available at: doi:10.4028/www.scientific.net/amm.556-562.2668.