FGEI International Journal of Multidisciplinary Research (FIJMR)
Vol. 01, Issue. 02, Dec, 2022, 1(2), PP-33-43
ISSN:3005-5628 (Online), ISSN: 3005-6470 (Print)

Social Network Centrality Measures for Ranking Academic Authors

Muhammad Ashraf Siddiqui[1*], Muhammad Qasim Khan[2]
[1*]MS Scholar, Allama Iqbal Open University, Islamabad, Email: forashraf@gmail.com
[2]Assistant Professor, AIOU, Islamabad, Email: qasim@aiou.edu.pk
[*]**Corresponding Author:** forashraf@gmail.com

KEYWORD
social academic networks, academic network analysis, social networking data analysis, centrality, node importance, influence, centrality measures, S2ORC dataset, network centrality methods.

ABSTRACT
The progression in the arena of IT particularly in IoT, social media and communication technologies, knowledge turn out to be available on finger tips in every field of life. Perceiving the world through the lens of overlapping networks that transfer information, knowledge, and power reveals insights into trends, technology, and interests. Analytics play vital role in the field academic network research and development. Analyzing social academic networks provides new perspectives on various interesting topics. The social connections have a significant impact on our actions, thoughts, and knowledge. However, standard statistical methods lack a reliable method for considering the impact of strong connections. This can only be achieved through academic network analysis and by comparing and contrasting relevant data. The social networking data analysis gives us tools to quantify the social network connections. Centrality can be used to quantify a node's importance and influence within the network as a whole. The concept of importance has various implications depending on the type of network being analyzed. Centrality indices ask the question, "What characterizes the significance of a node?" Different centrality measures can be used to demonstrate a node's importance. In this study, we analyzed and experimented various network centrality methods, their characteristics, and limitations using S2ORC dataset that consisting of 81 million heterogeneous academic objects with over 136 million nodes. The results have listed the top influential authors by using the Network Centrality measures such as an author Anand Radhakrishnan ranked 5th with a value of 30, 3rd with a value of 16.1 and 15th with a value of 0.386 in Degree, Betweenness and Closeness Centralities, respectively. The results are tested using performance matrices such as Spearman Correlation, Kendall Correlation and Similarity. The result of all the measures were consistent with each other. This study will also helpful in future researches to measure the semantic ranking of authors.

INTRODUCTION

Sharing of similar interests among the nodes is the basis of the composition of the social network. On daily basis; billions of users are forming social networks by interacting with social media channels. In the world the most popular social network is Facebook with over 2.70 billion monthly active users as on (*Facebook Statistics*, 2020). With over 330 million active users, the Twitter shares 500 million tweets per day (*Twitter Statistic*, 2020). Approximately 92 million people worldwide are using Flicker to share and store photos electronically (King, 2015). According to Alhaidari (Alhaidari et al., 2020) nowadays, massive data over the internet cause the difficulty in finding the relevant results quickly by the users. Social network sites attract more people from the target community and from businesses to explore various research articles. The impact of social networks can be used to measure influence, as a research topic. Retweets on Twitter likes on Facebook and shares on Flicker can be used to measure the social influence.

The Citation is one example, in which an academic network is represented by the association of authors and journals. Publications, citation, co-citation, authors, co-authorship, etc. are the entities of the academic networks such as scientometrics and bibliometric research, the h-index has a considerable influence and impact on them (Chandra et al., 2020). The entire academic network depends on these entities. Publications, journals, and authors; represent a node in the academic network. The links between each other are represented by citations among these nods. In neuroimaging study by (Baek et al., 2022) have used Degree Centrality to measure the relationship between the position of the social network among university students and concluded that neural processing of external stimuli is similar in highly-central individuals. A paper's significance increases when it is referred to in additionally explore publications (Mingers, 2009). Because of the publication of the paper, the worth of the author and journal also increases (Yan et al., 2011). Based on EigenVector Centrality, Arcagni proposed a novel method to predict tennis match (Arcagni et al., 2022). In this approach the system allows the ratings of the whole set of players to vary every time there is a new match. Hirsch introduced a way to estimate the work of a scientist using the h-index (Hirsch, 2005). It is used to rank the authors. The Publication count and citation count are factors to estimate the H-index. Although H-index is only a factor, as described by Kumar et al. It was observed that for ranking of authors more than 37 variants with the H-index were also used (Chandra et al., 2020). A researcher with an h-index has published h papers. In the other's scholarly paper each of these h papers has been cited at least h times. Citation C_d is computed by placing the document in decreasing order. A rankR_d is assigned to each document after the arrangement. In incremental way the rank should be kept in. When R_d<=C_d the h-index is the rank R_d but R_d+1>C_d. When a paper of a researcher is chosen in h-index, then the one disadvantage of h-index is at that point advance references got by that paper don't take an interest in expanding the h-index of that researcher. Too overcome the issues in h-index, the use of G-index was proposed by Egghe (L. Bornmann et al., 2011). In g-index authors having diverse citation can be separated. In the study by (Karlovčec et al., 2022) have proved that by using the group degree centralization that the groups are not much central. In very small as well as in very big groups, the group is expanding but its starts decreasing at some point.

METHODS AND TECHNIQUES

Different social network analysis measures are discussed in the related literature to list out the leading influential authors. For social networks, various network measurements can be applied. Central analysis and network algorithms are mainly used to identify the dominant author. To rank and give weightage to a web-page, the PageRank method can be used based on calculating web pages those are linked to it (inbound links). If the popular and important pages are linked to a web-page, this makes a web page important or if several links are coming from several web pages (Alhaidari et al., 2020). The most influential person adopts the same product and gives confidence to companions is the utter assumption of the social influence.

On social networks, the centrality analysis ranks the user by their location. Over the social network, the centrality describes the importance of the user in spreading information. It demonstrates that in the social network, the centrality is vital property (Weng et al., 2010). Using different techniques, the importance of centrality can be analyzed, the majority of centrality measures use structural information to point out the top ranked significant nodes of the network (Samad et al., 2020). The Commonly used centrality metrics is Betweenness centrality that has a drawback of calculating expensive shortest distance in many ways (M. Then et al., 2017). The shortest path using the Closeness centrality to the other nodes can be calculated (Freeman, 1978). Quick Proliferation of data from a vertex through the system; give pivotal importance to a vertex that it can achieve different vertices in fewer advances; closeness centrality and subsequently. The author uses the complex centrality and eigenvector centrality on Twitter (Maharani et al., 2014). For measuring the importance a web page; The PageRank is one of the most important method. PageRank is applied on Twitter for calculating influence (Ma et al., 2017).

Many domains can be used to measure Network Centrality. Urban traffic flow is analyzed by using network centrality measures (Zhao et al., 2017). The Closeness Centrality, PageRank and the Degree Centrality has been in application by the author to analyze the urban traffic flow. The centrality techniques were used by the author in smart cities to impact the performance and choices of citizens (Kaple et al., 2017a).

In the study of various link analysis methods, the results are published by Gupta (Gupta et al., 2020). The merits and demerits of the traditional ranking algorithm were discussed includes PageRank, Hyperlink-Induced Topic Search (HITS), and citation count. On link analysis methods, another comparative study was published (Jiang et al., 2018). Link analysis methods were studied by them and also compares summation of papers ranks and citation counting.

Degree centrality was used to identify the significant paper from relevant papers (Waheed et al., 2019). They have proved that a node having more neighbors can be measured as a greater influential node of the network. In other words, according to the metric, in a social network a person with a higher friends count, the more citations is the more central one.

We can summed-up the intrinsic limitations of Betweenness into two primary issues (Sarlas et al., 2020). The potential interaction between nodes and non-spatial uniformity of the population. To deal with this, a weighing factor can be included in formulation of Betweenness centrality that permit the analyst to deal with the non-uniform effect on the centrality of all pairs of the authors in the network.

The PageRank is used to ascertain the significance of an author in an academic network. The formulation is mentioned at equation 1, below(Ishfaq et al., 2016):.

$$PR(b_i) = \sum_{i=1}^{n} \frac{PR(b_i)}{OD(b_i)} + \frac{1-c}{N}$$

PageRank of a node can be measured as $PR(b_i)$, whereas, c is the damping factor, $OD(b_i)$ characterizes the out-Degree of an author i.

RESULT AND ANALYSIS

The framework depicted in figure 1 has been used to determine the most influential authors from the dataset. The measures of Network Centrality were applied and to check the prestige of an author that will lead us to the significance of an author, the comparison of measures such as PageRank and EigenVector with the base index i.e. the h-index and the citation count has been conducted. Prestige measure values and the values of centrality measures are confirmed against academic measures. OSim, Kendall, and Spearman's correlations and also with h-index and citation-count to validate the results

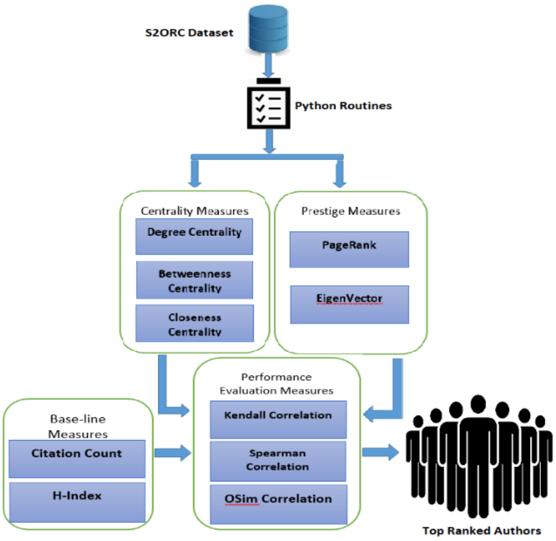


Figure 1. RESEARCH FRAMEWORK

As depicted in Table 1 that there are a total of 306,889 authors in the network. The paper per author ratio is 1:1.05 papers. Whereas, 2.3 is the average number of co-authors in the overall network. This network is a mix of single and co-author papers, even though results show that co-authorship has been increased from 2.24 (Yan et al., 2009) to 2.73 in the last ten years. It means that collaborative networks are increasing as compared to the decade-old data.

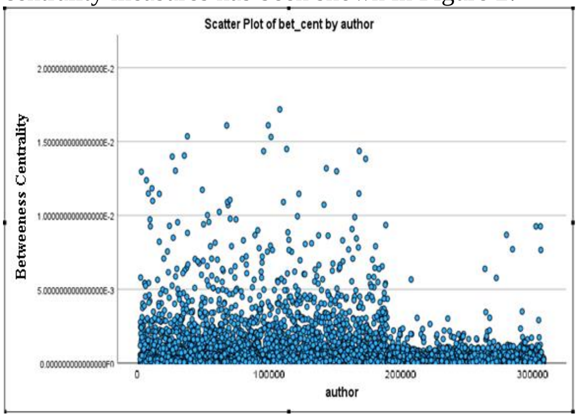
Table 1 Properties of S2ORC Network

Description	Value
Total Papers	101,576
Total Authors	306,889
Co-authors	2.73
Authors per paper	3.18
Papers per author	1.05
Largest Single Component	25.2%,
Degree (average)	25.6
Diameter	18
Path length(Average)	7.85

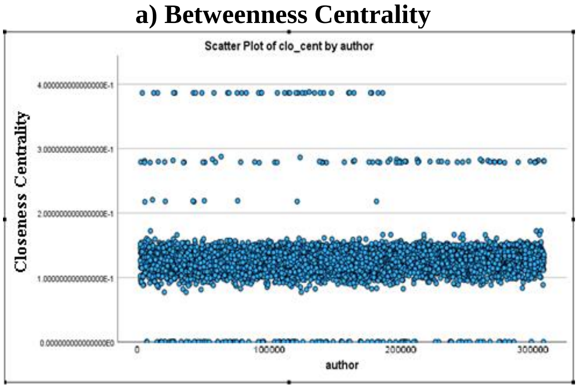
The most substantial node is the largest connected single node of the graph that fulfills the criteria of the most prominent node of the volume of the graph. As indicated in the results, the total number of authors in the network is the most significant component of the network and has a value of 25.2%. It depicts that S2ORC is not the largest graph of the connected components. SIGMOD, (Nascimento et al., 2003) reported a value of 60% of all nodes in the network that is highest constitute of the whole network. These results show that the factor largest connected network has a high value. The high value of the factor is due to a special interest group of common interest i.e. nature of the bibliography.

The average path length in the network is 7.85. The study conducted in 2007 (Yan et al., 2009) shows that the average path length was 9.68. This shows that at present time, collaborative networks are increasing as compared to previous results. Moreover, it can infer that in this era there is an increase of collaborative networks as compared to the previous research. The diameter of the resulted graph is 18.

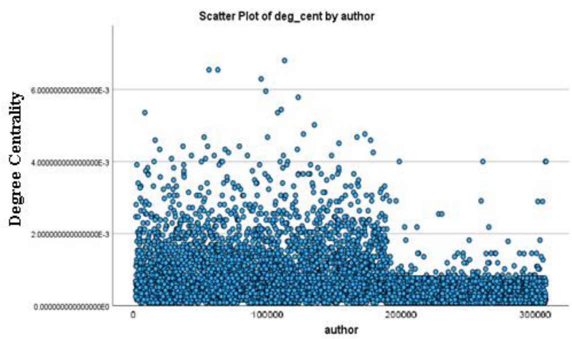
The frequency distribution of different network centrality measures has been shown in Figure 2.



a) Betweenness Centrality



b) Closeness Centrality



c) Degree Centrality

Figure 2. Frequency Distributions of Betweenness, Degree and Closeness Centralities

As indicated by the outcomes, the frequency distribution of degree and Betweenness centrality adheres to the fact that maximum authors have low value of centrality and a few authors have a high value of centrality called power-law distribution as shown in Figure 3. A normal curve is followed by the Closeness centrality, with a few exceptions.

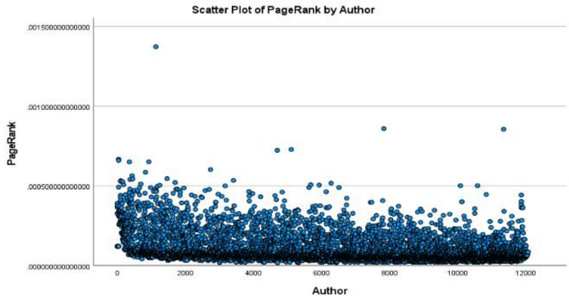


Figure 3. Frequency Distribution of PageRank

Frequency distributions of PageRank and EigenVector are shown in Figure 3 and 4 respectively. As per the results, Eigenvector and PageRank frequency distributions most of the authors have a low value of centrality and a few authors have a high value of centrality, therefore, a power-law distribution.

In Table 2, the list depicts the top twenty (20) authors calculated by using network centrality measures such as the Closeness centrality, the Degree centrality and the Betweenness centrality.

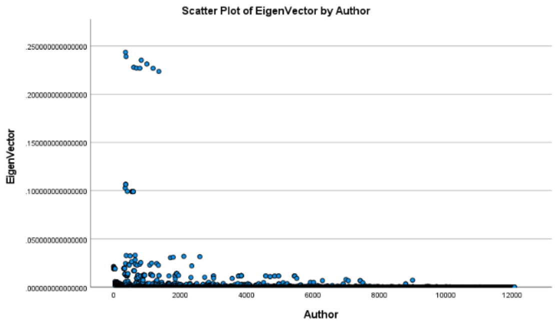


Figure 4. Frequency Distribution of PageRank

Authors in prominent bold font show consecutive appearances in all three centralities. While prominent bold and italic font shows the authors appeared in two centralities.

Table 2 Ranking of Authors

Rank	Degree		Betweenness		Closeness	
1	Andrew McLennan	34	Qi Wang	18.83	Amy Ellison	0.390
2	*Alessandro Rizzi*	33	Andrei Chapoval	17.17	Amrita Saha	0.388
3	*Albert Liao*	33	Anand Radhakrishnan	16.10	A P Strelnikova	0.387
4	Amy Ellison	31	Alexander Tsatsoulas	16.09	Anthony Comuzzie	0.387
5	Anand Radhakrishnan	30	*Ahmet Toksoy*	15.35	A Lugo	0.386
6	*Anil Adapa*	29	Anders B phgr rglum	15.31	*Alessandro Rizzi*	0.386
7	*Andrew Durham*	27	Andrew McLennan	14.49	*Albert Liao*	0.386
8	Andrei Chapoval	27	Ase Uttenthal	14.36	*Andrew Durham*	0.386
9	*Abdelfattah Attallah*	27	Amy Ellison	14.35	Apul Goel	0.386
10	Anne Tjonneland	25	Philip Kass	14.11	*Abdelfattah Attallah*	0.386
11	*A Stasiak Barmuta*	24	Ahmad Kamaei	14.05	*Anil Adapa*	0.386
12	*Annabel Chen*	24	A Dunker	13.98	Andrei Chapoval	0.386
13	A Landt	23	*A Stasiak Barmuta*	13.83	*Annabel Chen*	0.386
14	Anastasios Tefas	23	Antonello Nicolini	13.19	*Ahmet Toksoy*	0.386
15	A Scanu	23	*Atsushi Iwakura*	13.03	Anand Radhakrishnan	0.386
16	*Atsushi Iwakura*	23	Aparecido Pereira	12.99	Hyun u2010Taek Kim	0.386
17	A Carvalhais	23	A Akkas	12.95	Bong u2010Yeon Cho	0.386
18	*Apul Goel*	23	A Bailey Farchione	12.38	Eric Daliri	0.386
19	A L Kahler	22	Jun Zhao	12.09	Hyeon Jo	0.386
20	Alastair Fitter	22	Seth Owusu Agyei	11.86	Andrew McLennan	0.386

We can observe from the data shown in Table 2 that few authors are having ranking in two centrality measures. That means few authors have low closeness centrality as compared to the Betweenness centrality and degree. For instance, Andrew McLennan with a degree centrality value of 34, showing that he has collaborated with eighty (80) authors, whereas, the closeness centrality of Andrew McLennan is relatively low; for that reason, the author is not listed as top 20. Andrew Mclennan's co-author Johannes Berg is situated in Germany; therefore, he is close to USA authors.

In table 3, the measured values of the top forty (40) authors calculated on the basis of citation count along with network centrality measures are shown. We can observe that degree centrality is more in line with citation count (baseline) as compared with other centrality measures.

Table 3 Citation Count Values of the Top 40 Authors.

Authors	Citation Count		The Rank of Centrality Measures		
	Counts	Ranking	Degree	Closeness	Betweenness
A Jones	168	1	11	12.59	0.00
A Singh	118	2	9	10.22	0.00
A Smith	116	3	5	10.28	0.00
Arch Woodside	105	4	0	0.02	0.00
A Khan	101	5	6	13.21	0.04
A Sharma	99	6	6	11.57	0.05
Anthony Barnett	94	7	4	9.69	0.00
A Berger	92	8	12	12.73	0.14
Andrew Burroughs	91	9	9	13.95	0.38
Allan Ropper	89	10	8	14.64	0.83
Andreas Schedl	82	11	8	13.20	0.09
Annette Hammes	82	12	15	13.57	0.30
Andrea De Lucia	81	13	11	13.27	0.06
Attila Csendes	79	14	8	13.30	0.11
A Martin	68	15	8	13.98	0.22
A Wong	66	16	12	12.39	0.03
A Gupta	65	17	12	11.31	0.00
Alexander Maass	65	18	5	11.34	0.19
Alessandra Mallei	64	19	8	11.29	0.00
Andras Vasy	62	20	8	11.43	0.00
Alexander Dobrovic	61	21	10	14.89	0.06
Alan MacDiarmid	61	22	13	14.56	0.96
Aaron Livingston	59	23	12	14.90	0.39
A El Naggar	58	24	10	11.56	0.00
A Dobson	58	25	4	11.30	0.05
A Khisti	57	26	4	13.83	0.69
Alan Hargens	56	27	6	14.12	0.53
A Shah	55	28	17	13.65	0.17
A Schwartz	54	29	18	14.12	0.38
Anthony Kim	53	30	11	15.01	0.06
Anatoly Zhitkovich	53	31	7	11.80	0.00
Alan Francis	53	32	13	11.99	0.11
Andrew Fisher	51	33	13	11.64	0.00
A Korczyn	51	34	8	12.73	0.00
Andr u00e9 Schiper	50	35	14	13.54	0.23
Alec Wolman	50	36	6	11.53	0.02
A M Niknejad	50	37	13	12.32	0.00
Andrew Gordon	49	38	3	10.90	0.00
Alexander Hartemink	49	39	5	12.16	0.00
A Katz	48	40	14	12.27	0.05

Though, in a few cases, a high value of citation count has a low ranking in centrality measures. The case in point, Arch Woodside, Anthony Barnett, Zhitkovich, A. El Naggar, and Andrew Gordon have a large number of citation counts but his degree, closeness, and Betweenness centrality values are lower. Results show that A Khan has four co-authors, all four co-authors are situated in the UK, and most of them are not cut-points (Bastian et al., 2009), thus A Khan does not has high value of centrality. Nodes whose expulsion builds the number of components are called the cut points. A Khan has five co-authors H. H. Dang from the USA, P. Grondona from Italy, D. P.

Edwards from the UK, S. M. Andreani from the UK. Moreover, in 2005, A Khan has co-authored a paper that has been cited 38 times. His citation count is high because of the high number of citations of the paper, but he has limited co-authorship. Thus, he has a low rank of centrality, in Degree (6), Betweenness (0.04), and Closeness Centralities (13.21). The same is the case with P. Grondona. He has a value of 38 i.e. the publication count and some of the publications are highly cited and he only has 4 co-authors in the data set. Moreover, D. P. Edwards has a total of two papers in the dataset and all of them are co-authored. Whereas, some authors have a low centrality value for Betweenness and Closeness but have a high value of Degree Centrality. Such as the case with Anthony Barnett, A Wong, A Gupta, and Alessandra Mallei. Although their ranking of centrality is linked to their ranking of citation. In the dataset only a few publications of the authors are incorporated that has a chances to affect the ranking of the results.

EigenVector and PageRank of the top 20 authors are calculated, as shown in Table 4. The names of authors shown in bold are having a consecutive appearance in both the prestige measures.

Table 4 Top Twenty (20) Authors Measured on their Prestige

Rank	EigenVector		PageRank	
1	Andrew McLennan	0.243569	Amy Ellison	0.001373
2	Anand Radhakrishnan	0.239097	Amrita Saha	0.000860
3	Anil Adapa	0.235344	A P Strelnikova	0.000855
4	Anne Tjonneland	0.231416	Anthony Comuzzie	0.000728
5	Atsushi Iwakura	0.228006	A Lugo	0.000723
6	A L Kahler	0.227128	Alessandro Rizzi	0.000668
7	Alastair Fitter	0.227007	Albert Liao	0.000659
8	Apul Goel	0.226940	Andrew Durham	0.000651
9	Alexander Persterer	0.223681	Andrew McLennan	0.000650
10	Anton Simorov	0.106509	Abdelfattah Attallah	0.000602
11	Abhijit Shaligram	0.106509	Anil Adapa	0.000585
12	Dmitry Oleynikov	0.106509	Andrei Chapoval	0.000564
13	Eugene Boilesen	0.106509	Annabel Chen	0.000545
14	Jon Thompson	0.106509	Ahmet Toksoy	0.000543
15	Valerie Shostrom	0.106509	Anand Radhakrishnan	0.000540
16	Angharad Marks	0.102751	Annemiek J Linn	0.000534
17	Renee Hsia	0.099284	Bong u2010Yeon Cho	0.000530
18	Anne Tomolo	0.099026	Eric Daliri	0.000530
19	Amy O u2019Shea	0.099026	Hyeon Jo	0.000530
20	Brad Wright	0.099026	Hyun u2010Taek Kim	0.000530

The top 40 authors calculated on their h-index value are listed in Table 5. The H-index is a measure by which we can observe both; the citations received by an author or received by a publication and also the productivity of the paper.

To check the prestige of an author, the comparison of measures such as PageRank and EigenVector with the base index i.e. the h-index has been conducted. In table 5, the ranking of top 40 authors with respect to their value of centrality are typed in in bold font with their rank of h-index. There are some difference between values of network centralities and h-index that is shown in Table 5. The Dong Kim, C Chen, and C Lee are the three top authors in the ranking of h-index, whereas, they are having a low value of centrality. Dong Kim's h-index is 20 also has a large number of inbound citations i.e. 182, but as far as the co-authors are concerned, he has a low value. Due to the fewer number of co-authors, the results of prestige measures are also low as 0.798636 for PageRank and 0.079864 for EigenVector. The same is the case with C Liu and Feng Zhang have 12 and 21 co-authors, respective but having a low value of prestige.

Table 5 H-Index - Top Forty (40) Authors

	H-index		Prestige Measures	
Author	Counts	Ranking	PageRank	Eigenvector
Dong Kim	20	1	0.798636	0.079864
C Chen	20	2	0.898636	0.229864
C Lee	18	3	0.585396	0.05854
Hui Zhang	17	4	0.685396	0.20854
H Li	15	5	0.512286	0.051229
Feng Liu	15	6	0.612286	0.201229
Bo Wang	14	7	0.712286	0.351229
Feng Li	14	8	0.812286	0.501229

H Li	14	9	0.512286	0.051229
D Kim	13	10	0.612286	0.201229
Bo Wang	13	11	0.712286	0.351229
Feng Li	12	12	0.812286	0.501229
C Smith	12	13	0.912286	0.651229
D Kim	12	14	1.012286	0.801229
Bin Chen	11	15	1.269024	0.126902
D Smith	11	16	1.369024	0.276902
Hong Li	11	17	0.764738	0.076474
A Jones	11	18	0.261283	0.026128
C Liu	11	19	0.361283	0.176128
Feng Zhang	11	20	0.346802	0.03468
David Williams	11	21	0.685438	0.068544
Fei Wang	10	22	0.850373	0.085037
Bin Liu	10	23	0.950373	0.235037
David Williams	10	24	0.685438	0.068544
Bin Zhang	10	25	1.034705	0.10347
Gang Wang	10	26	1.134705	0.25347
Gang Chen	10	27	1.234705	0.40347
Dae Kim	9	28	1.334705	0.55347
C Yang	9	29	1.434705	0.70347
Fan Zhang	9	30	1.534705	0.85347
C Rao	9	31	1.634705	1.00347
Hao Liu	9	32	0.97507	0.097507
Feng Liu	9	33	1.07507	0.247507
David Williams	9	34	0.685438	0.068544
Bin Zhang	9	35	1.034705	0.10347
E Smith	9	36	1.134705	0.25347
Fang Chen	9	37	1.234705	0.40347
D Brown	9	38	0.59846	0.059846
A Khan	8	39	1.049484	0.104948
Feng Wang	8	40	0.58827	0.058827

The Tables 6, 7, and 8, show the results of correlation with prestige measures and with centrality measures. As per the results, both the prestige measures have a value of high correlation at 0.01 value of p. Where PageRank is having a low value of correlation coefficient than EigenVector. This shows that there is potential in PageRank and EigenVector to rank authors with a correlation of h-index.

Table 6 Spearman's Correlation

Centrality	H-index	PageRank	Eigenvector
PageRank		1	0.83*
Eigenvector			1
H-index	1	0.81*	0.79*

The value of correlation is substantial at level of 0.01

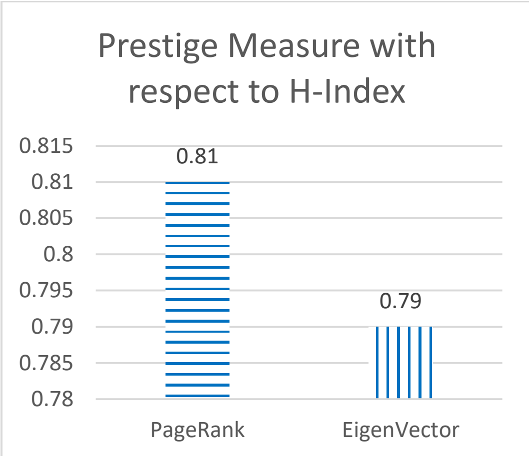


Figure 5 Spearman's Correlation of Prestige Measures with respect to H-Index

Table 7 Kendall Rank Correlation among EigenVector, H-index and PageRank

Centrality	H-index	PageRank	Eigenvector
PageRank		1	0.88*
Eigenvector			1
H-index	1	0.69*	0.9*

The value of correlation is substantial at level of 0.01

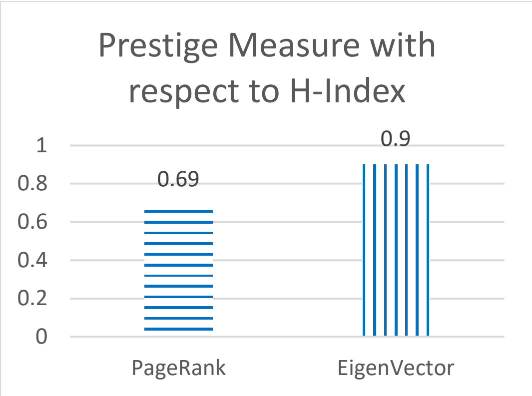


Figure 6. Kendall Correlation of Prestige Measures with respect to H-Index

Table 8 OSim Correlation (k=250)

Centrality	H-index	PageRank	Eigenvector
H-index	1	0.50	0.65
PageRank		1	0.61
Eigenvector			1

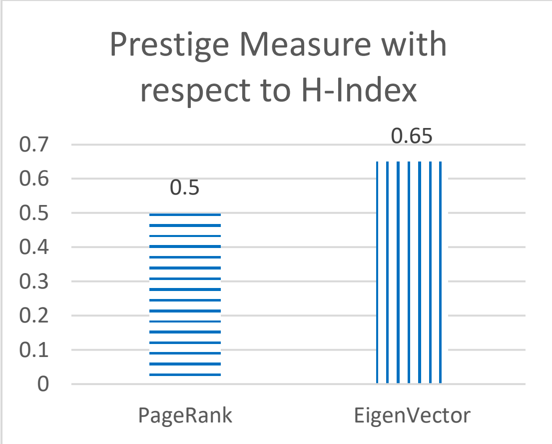


Figure 7 **OSim Correlation of Prestige Measures with respect to H-Index**

Similarly, the results of all three correlation measures are described in Tables 6, 7, and 8. The correlation of coefficient in-degree centrality is having a higher correlation as compared with two other centralities. While, h-index at p-value 0.01, has a significantly high correlation with centrality measures. We can say that a high value of correlation with citation count also has a high potential to rank authors that can be observe in the Table 9 below:

Table 9 **Spearman Correlation**

Centrality	Citation Count	Degree	Betweenness	Closeness
Degree		1	0.84*	0.84*
Betweenness			1	0.80*
Closeness				1
Citation Count	1	0.61*	0.42*	0.53*

The value of correlation is substantial at level of 0.01

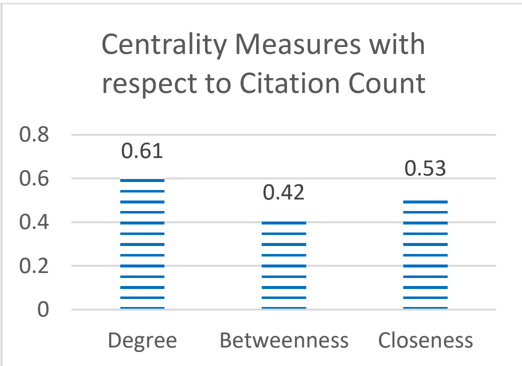


Figure 8 Spearman's Correlation with respect to Citation Count

Table 10 Kendall Rank Correlation

Centrality	Citation Count	Degree	Betweenness	Closeness
Citation Count	1	1.0*	0.21*	0.23*
Degree		1	0.70*	0.93*
Betweenness			1	0.60*
Closeness				1

The value of correlation is substantial at level of 0.01

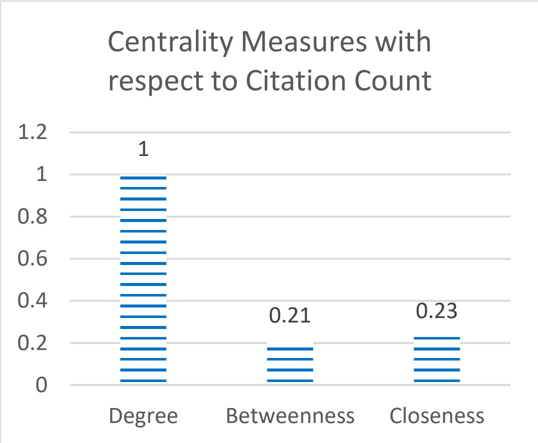


Figure 9 Kendall Correlation with respect to Citation Count

Table 11 OSim (k=250) [Centralities

Centrality	Citation Count	Degree	Betweenness	Closeness
Degree		1	0.62	0.69
Betweenness			1	0.53
Closeness				1
Citation Count	1	0.61	0.39	0.50

Figure 10 OSim Correlation with respect to Citation Count

CONCLUSION

The technical contribution in this research is the use of Network Centrality Measures to measure the influence of a publication. The three network centrality measures the Degree, Betweenness, and Closeness, are all different ways to quantify the importance of an author in a network. The use of these measures, along with prestige measures such as EigenVector and PageRank, provides a more comprehensive approach to measuring the influence of a publication and its authors. The influence of a journal publication can be measured by using the academic indexes which gives insight into the different aspects of the paper such as author, year of publication, inbound citation count, outbound citation count, h-index for a given author, etc.

The result shows that the ranking of authors computed using all three network centrality measures including Degree, Betweenness and Closeness are consistent with each other. It is further verified from the results that the ranking of authors is consistent with two prestige measures such as EigenVector and PageRank.

Academic measures are also validated in the results for network centrality and prestige of an author. The outcomes show that the betweenness centrality, closeness centrality, and degree centrality measures are altogether associated with citation count. Between these three estimates the degree centrality that is having a connection with the citation count.

It confirms that in author ranking measures, degree centrality has the capacity to be determined as a measure to rank authors. Nonetheless, sometimes, a few authors have a high citation count yet there is low rank of centrality. Results show that citation count measures the impact of articles and also quality.

However, in some cases, some authors have a high citation count but they have a low rank in the centrality measures. The influence and quality of articles can be measured using citation count. On the other hand, the impact and quality of author’s discipline can be measured using the network centrality. One of the reason for this is; the citations and the centralities measure different contents. Citation counts measure the quality and the significance of articles and social network are used to measures both article influence and author’s discipline impact. Results of correlation between the prestige measures and h-index show high correlation, among prestige measure eigenvector have a high correlation.

Subsequently, degree centrality estimates the capacity of the author’s co-authorship, Betweenness centrality ascertain the importance of an author concerning virtual communication with closeness centrality and the rest of the authors that computes the author’s rank in a network of co-authors and this also finds the shortest distance with further authors in the social network. Thus, centrality has its significance in the evaluation of impact, since it incorporates both article impact and authors' field impact.

Future work may include the semantic ranking of authors and further exploration to measure centrality of a node depending upon on the centrality of its neighbors' centrality and to rank authors to measure the impact of neighbors' ranking on a given author.

REFERENCES

Alhaidari, F., Alwarthan, S., & Alamoudi, A. (2020). User Preference Based Weighted Page Ranking Algorithm. 2020 3rd International Conference on Computer Applications & Information Security (ICCAIS), 1–6. https://doi.org/10.1109/ICCAIS48893.2020.9096823

Allen Institute for AI. (n.d.). Retrieved April 11, 2023, from https://allenai.org/

Amjad, T., Daud, A., & Aljohani, N. R. (2018). Ranking authors in academic social networks: A survey. Library Hi Tech, 36(1), 97–128. https://doi.org/10.1108/LHT-05-2017-0090

Arcagni, A., Candila, V., & Grassi, R. (2022). A new model for predicting the winner in tennis based on the eigenvector centrality. Annals of Operations Research. https://doi.org/10.1007/s10479-022-04594-7

Baek, E. C., Hyon, R., López, K., Finn, E. S., Porter, M. A., & Parkinson, C. (2022). In-degree centrality in a social network is linked to coordinated neural activity. Nature Communications, 13(1), 1118. https://doi.org/10.1038/s41467-022-28432-3

Bastian, M., Heymann, S., & Jacomy, M. (2009). Gephi: An Open Source Software for Exploring and Manipulating Networks. https://doi.org/10.13140/2.1.1341.1520

Bibi, F., Khan, H., Iqbal, T., Farooq, M., Mehmood, I., & Nam, Y. (2018). Ranking Authors in an Academic Network Using Social Network Measures. Applied Sciences, 8(10), 1824. https://doi.org/10.3390/app8101824

Brandao, M. A., & M. M. Moro. (2012). Affiliation Influence on Recommendation in Academic Social Networks. AMW, 230–234.

Chandra, P. K., Jha, V., & Abhishek, K. (2020). Effective Author Ranking Using Average of Different h-Index Variants. In M. Pant, T. Kumar Sharma, R. Arya, B. C.

Sahana, & H. Zolfagharinia (Eds.), Soft Computing: Theories and Applications (Vol. 1154, pp. 47–56). Springer Singapore. https://doi.org/10.1007/978-981-15-4032-5_6

Facebook Statistics. (2020, August). https://www.omnicoreagency.com/facebook-statistics

Fieller, E. C., Hartley, H. O., & Pearson, E. S. (1957). Tests for Rank Correlation Coefficients. I. Biometrika, 44(3/4), 470. https://doi.org/10.2307/2332878

Fletcher, J. M., & Wennekers, T. (2018). From Structure to Activity: Using Centrality Measures to Predict Neuronal Activity. International Journal of Neural Systems, 28(02), 1750013. https://doi.org/10.1142/S0129065717500137

Freeman, L. C. (1978). Centrality in social networks conceptual clarification. Social Networks, 1, 215–239.

Haveliwala, T. H. (2003). Topic-sensitive pagerank: A context-sensitive ranking algorithm for web search. IEEE Transactions on Knowledge and Data Engineering, 15(4), 784–796. https://doi.org/10.1109/TKDE.2003.1208999

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. Proceedings of the National Academy of Sciences, 102(46), 16569–16572. https://doi.org/10.1073/pnas.0507655102

Ishfaq, U., Khan, H. U., & Iqbal, K. (2016). Modeling to find the top bloggers using Sentiment Features. 2016 International Conference on Computing, Electronic and Electrical Engineering (ICE Cube), 227–233. https://doi.org/10.1109/ICECUBE.2016.7495229

Jianqiang, Z., Xiaolin, G., & Feng, T. (2017). A New Method of Identifying Influential Users in the Micro-Blog Networks. IEEE Access, 5, 3008–3015. https://doi.org/10.1109/ACCESS.2017.2672680

Jörgens, H., Kolleck, N., & Saerbeck, B. (2016). Exploring the hidden influence of international treaty secretariats: Using social network analysis to analyse the Twitter debate on the 'Lima Work Programme on Gender.' Journal of European Public Policy, 23(7), 979–998. https://doi.org/10.1080/13501763.2016.1162836

Kaple, M., K. Kulkarni, & K. Potika. (2017a). Viral Marketing for Smart Cities: Influencers in Social Network Communities. 9th IEEE International Workshop on Big Data Appications in Smart City Development.

Kaple, M., K. Kulkarni, & K. Potika. (2017b). Viral Marketing for Smart Cities: Influencers in Social Network Communities. 9th IEEE International Workshop on Big Data Appications in Smart City Development.

Karlovčec, M., Krnc, M., & Škrekovski, R. (2022). Evaluating group degree centrality and centralization in networks. Informatica, 46(5). https://doi.org/10.31449/inf.v46i5.3817

Khan, H. U., Daud, A., Ishfaq, U., Amjad, T., Aljohani, N., Abbasi, R. A., & Alowibdi, J. S. (2017). Modelling to identify influential bloggers in the blogosphere: A survey. Computers in Human Behavior, 68, 64–82. https://doi.org/10.1016/j.chb.2016.11.012

Khan, H. U., Daud, A., & Malik, T. A. (2015). MIIB: A Metric to Identify Top Influential Bloggers in a Community. PLOS ONE, 10(9), e0138359. https://doi.org/10.1371/journal.pone.0138359

King, D. L. (2015). Landscape of social media for libraries. *Library Technology Reports*, 51(1), 10–15.

L. Bornmann & W. Marx. (2011). The h index as a research performance indicator. *European Science Editing*, 37, 77–80.

Leydesdorff, L. (2007). Betweenness centrality as an indicator of the interdisciplinarity of scientific journals. *Journal of the American Society for Information Science and Technology*, 58(9), 1303–1319. https://doi.org/10.1002/asi.20614

Li, L., Wang, X., Zhang, Q., Lei, P., Ma, M., & Chen, X. (2014). A Quick and Effective Method for Ranking Authors in Academic Social Network. In J. J. Park, S.-C. Chen, J.-M. Gil, & N. Y. Yen (Eds.), *Multimedia and Ubiquitous Engineering* (Vol. 308, pp. 179–185). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-54900-7_26

M. Then, Günnemann, S., A. Kemper, & T. Neumann. (2017). Efficient Batched Distance, Closeness and Betweenness Centrality Computation in Unweighted and Weighted Graphs. *Datenbank-Spektrum*, 17, 169–182.

Ma, X., Li, C., Bailey, J., & Wijewickrema, S. (2017). Finding Influentials in Twitter: A Temporal Influence Ranking Model. ArXiv:1703.01468 [Cs]. http://arxiv.org/abs/1703.01468

Maharani, W., Adiwijaya, & Gozali, A. A. (2014). Degree centrality and eigenvector centrality in twitter. 2014 8th International Conference on Telecommunication Systems Services and Applications (TSSA), 1–5. https://doi.org/10.1109/TSSA.2014.7065911

Mingers, J. (2009). Measuring the research contribution of management academics using the Hirsch-index. Journal of the Operational Research Society, 60(9), 1143–1153. https://doi.org/10.1057/jors.2008.94

Nascimento, M. A., Sander, J., & Pound, J. (2003). Analysis of SIGMOD's co-authorship graph. ACM SIGMOD Record, 32(3), 8–10. https://doi.org/10.1145/945721.945722.

Nykl, M., Ježek, K., Fiala, D., & Dostal, M. (2014). PageRank variants in the evaluation of citation networks. Journal of Informetrics, 8(3), 683–692. https://doi.org/10.1016/j.joi.2014.06.005

Rehan Khan, Khan, H. U., Muhammad Shehzad Faisal, Iqbal, K., & Muhammad Shahid Iqbal Malik. (2016). An Analysis of Twitter users of Pakistan. International Journal of Computer Science and Information Security, 14, 855.

Samad, A., Qadir, M., Nawaz, I., Islam, M., & Aleem, M. (2020). SAM Centrality: A Hop-Based Centrality Measure for Ranking Users in Social Network. EAI Endorsed Transactions on Industrial Networks and Intelligent Systems, 7(23), 163985. https://doi.org/10.4108/eai.13-7-2018.163985

Sarlas, G., Páez, A., & Axhausen, K. W. (2020). Betweenness-accessibility: Estimating impacts of accessibility on networks. Journal of Transport Geography, 84, 102680. https://doi.org/10.1016/j.jtrangeo.2020.102680

Semenkovich, S. A., & Tsukanova, O. A. (2019). On the Algorithms of Identifying Opinion Leaders in Social Networks. Procedia Computer Science, 162, 778–785. https://doi.org/10.1016/j.procs.2019.12.050

Tang, J., Zhang, J., Yao, L., Li, J., Zhang, L., & Su, Z. (2008). ArnetMiner: Extraction and mining of academic social networks. Proceeding of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD 08, 990. https://doi.org/10.1145/1401890.1402008

Twitter Statistic. (2020, August). 2https://www.omnicoreagency.com/twitter-statistics/2020

Waheed, W., Imran, M., Raza, B., Malik, A. K., & Khattak, H. A. (2019). A Hybrid Approach Toward Research Paper Recommendation Using Centrality Measures and Author Ranking. IEEE Access, 7, 33145–33158. https://doi.org/10.1109/ACCESS.2019.2900520

Waltman, L., & van Eck, N. J. (2012). The inconsistency of the h-index. Journal of the American Society for Information Science and Technology, 63(2), 406–415. https://doi.org/10.1002/asi.21678

Weng, J., Lim, E.-P., Jiang, J., & He, Q. (2010). TwitterRank: Finding topic-sensitive influential twitterers. Proceedings of the Third ACM International Conference on Web Search and Data Mining - WSDM '10, 261. https://doi.org/10.1145/1718487.1718520

Yan, E., & Ding, Y. (2009). Applying centrality measures to impact analysis: A coauthorship network analysis. Journal of the American Society for Information Science and Technology, 60(10), 2107–2118. https://doi.org/10.1002/asi.21128

Yan, E., & Ding, Y. (2011). Discovering author impact: A PageRank perspective. Information Processing & Management, 47(1), 125–134. https://doi.org/10.1016/j.ipm.2010.05.002

Zhao, S., Zhao, P., & Cui, Y. (2017a). A network centrality measure framework for analyzing urban traffic flow: A case study of Wuhan, China. Physica A: Statistical Mechanics and Its Applications, 478, 143–157. https://doi.org/10.1016/j.physa.2017.02.069

Zhao, S., Zhao, P., & Cui, Y. (2017b). A network centrality measure framework for analyzing urban traffic flow: A case study of Wuhan, China. Physica A: Statistical Mechanics and Its Applications, 478, 143–157. https://doi.org/10.1016/j.physa.2017.02.069