Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 57 (2015) 606 – 613

Procedia Computer Science

3rd International Conference on Recent Trends in Computing 2015 (ICRTC-2015)

# Key Author Analysis in Research Professionals' Relationship Network Using Citation indices and Centrality

Anand Bihari[a,*], Manoj Kumar Pandia[b,**]

[a]*Department of CSE, Silicon Institute of Technology, Bhubaneswar, Odisha, India*
[b]*Department of MCA, Silicon Institute of Technology, Bhubaneswar, Odisha, India*

**Abstract**

In social network analysis, the importance of an actor can be found by using the centrality metrics. There are many centrality metrics available e.g. degree, closeness, betweenness, eigenvector etc. In research community authors forms a social network, which is called Research Professionals' Collaboration Network. This is similar to social network where each author is an actor and an article written together by some authors establishes collaboration between them. Each author acquires a certain value based on the citation of their articles. There are many citation indices are available such as citation count, h-index, g-index, i10-index etc. To analyze the Research Professionals' collaboration Network and for finding the key author, the citation indices can be used. In this paper, we compare and combine both social network analysis metrics and the citation indices to get better result in finding the key author.



Peer-review under responsibility of organizing committee of the 3rd International Conference on Recent Trends in Computing 2015 (ICRTC-2015)

*Keywords:* Social Network; Researcher Relationship; Centrality; citation count; h-index; g-index; i10-index;

## 1. Introduction

A research professionals' relationship network is a set of researcher's which has connection in pair to represent their relationship. Two researchers are considered in a relationship if they have published articles in journal, conference and publish & edit books together. In such type of network, a researcher is called as ”node” or ”vertex” and the connection is an ”edge”. Research professionals are dedicated to the advancement of the knowledge and practice of professions through developing, supporting, regulating and promoting professional standards for technical and ethical competence. Research professionals' relationship network is represented by undirected weighted graph. This network represents only the collaboration of research professionals and no where related to how many papers published by each of these authors. Research professionals' relationship network is a type of social network. Different type of

* Corresponding author. Mob.: +919835842013.
** Corresponding author. Mob.: +919338845287.
*E-mail address:* csanandk@gmail.com

*E-mail address:* manoj.pandia@gmail.com


Peer-review under responsibility of organizing committee of the 3rd International Conference on Recent Trends in Computing 2015 (ICRTC-2015)
doi:10.1016/j.procs.2015.07.414

social network analysis metrics and citation indices are available for finding key actors in the network. In this paper we use different type of social network analysis metrics and citation indices for finding key author in the research professionals' relationship network.

## 2. Social Network Analysis (Methodology)

Social Network Analysis (SNA) views social relationships in terms of network theory, consisting of nodes, representing individual actors within the network, and ties which represent relationships between the individuals, such as friendship, kinship, organizations and sexual relationships [1,3,4]. These networks are often depicted in a social network diagram, where nodes are represented as points and ties are represented as lines. Social network analysis[7,14,29] is the mapping and measuring of relationships and flows between people, groups, organizations, computers, URLs, and other connected information/knowledge entities. The nodes in the network are the people and groups while the links show relationships or flows between the nodes. SNA provides both a visual and a mathematical analysis of human relationships [8,9].

In general, the benefit of analyzing social networks is that it can help people to understand how to share professional knowledge in an efficient way and to evaluate the performance of individuals, groups, or the entire social network [3].

## 3. Social network analysis metrics and citation indices for key author analysis

Key or prominent actors are those that are linked or involved with other actors extensively. In the context of an organization, a person with extensive contacts (links) or communications with many other people in the organization is considered more important than a person with relatively fewer contacts.

Thus, several types of social network analysis metrics and citation indices are defined to find key actor in the social network. Here, we discuss degree centrality, closeness centrality, betweenness centrality, eigenvector centrality, frequency, citation count, h-index, g-index and i10-index.

**Degree Centrality:** Degree centrality of a node refers to the number of edges that are adjacent to this node [1,2,4]. Degree centrality represents the simplest instantiation of the notion of centrality since it measures only how many connection tie authors to their immediate neighbors in the network [7,10]. Therefore important nodes usually have high degree. Degree centrality is an indicator of an actor's communication activity. The normalized degree centrality is defined as the number of links of an actor divided by the maximal possible number [21,27,32]. The normalized degree centrality $d_i$ of node i is given as

$$d_i = \frac{d(i)}{n-1} \tag{1}$$

**Closeness Centrality:**Closeness measures how close an individual is to all other individual in a network, directly or indirectly [1,3]. This metric can only be used in the connected network which means every node has at least one path to the other [11,13,24]. The closeness centrality $C_c(n)$ of a node n is defined as the reciprocal of the average shortest path length and is computed as follows:

$$C_c(n) = \frac{1}{avg(L(n,m))} \tag{2}$$

where L(n,m) is the length of the shortest path between two nodes n and m. The closeness centrality of each node is a number between 0 and 1. Closeness centrality is a measure of how fast information spreads from a given node to other reachable node in the network.

**Betweenness Centrality:**Betweenness centrality is an indicator of an actor's potential control of communication within the network [1,2,7]. Betweenness centrality is defined as the ratio of the number of shortest paths (between all pairs of nodes) that pass through a given node divided by the total number of shortest paths. The betweenness centrality $C_b(n)$ of a node n is computed as follows

$$C_b(n) = \sum_{s \neq n \neq t} \frac{\sigma_{st}(n)}{\sigma_{st}} \tag{3}$$

where s and t are nodes in the network different from n, $\sigma st$ denotes the number of shortest paths from s to t, and $\sigma st$(n) is the number of shortest paths from s to t that n lies on[4,23].
The betweenness value for each node n is normalized by dividing by the number of node pair excluding n: (N-1)(N-2)/2, where N is the total number of nodes in the connected component that n belongs to. Thus, the betweenness centrality of each node is a number between 0 and 1[25,26,32].
**Eigenvector Centrality :**Eigenvector centrality is a measure of the influence of a node in a network. Eigenvector centrality is extension of degree centrality [12]. In degree centrality, degree centrality of a node is simply count the total no of nodes that are adjacent to that node, but in eigenvector centrality, not only consider the total no of adjacent nodes also consider the importance of the adjacent node. In eigenvector centrality, all connections are not equal. In general, connections with influence person will lend a person more influence than connection with less influence persons. In eigenvector centrality not only the connections are important also the score (eigenvector centrality) of the connected node. Eigenvector centrality is calculated by assessing how well connected an individuals to the parts of the network with the greatest connectivity[15]. Individuals with high eigenvector scores have many connections, and their connections have many connections, and their connections have many connections out to the end of the network. Eigenvector centrality is simply dominant of Eigenvector of the adjacency matrix. Philip Bonacich proposed eigenvector centrality in 1987 and Google's PageRank is a variant of it [16,19].
**Eigenvector centrality based on adjacency matrix :** For a given graph G:=(V,E) with |V| number of vertices. Let A = ($a_{v,t}$ ) be the adjacency matrix of a graph G is the square matrix with rows and columns labeled by the vertices and edges [16,20]
i.e

$$a_{v,t} = \begin{cases} 1, & \text{if vertex v is linked to vertex t,} \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

If we denote the centrality score of vertex v by $X_v$, then we can allow for this effect by making $X_v$ proportional to the average of the centralities of v's network neighbors:

$$X_v = \frac{1}{\lambda} \sum_{t \in M(v)} X_t = \frac{1}{\lambda} \sum_{t \in G} a_{v,t} X_t \tag{5}$$

Where M(v) is a set of the neighbors of v and $\lambda$ is a constant. With a small rearrangement this can be written in vector notation as the Eigenvector equation

$$AX = \lambda X \tag{6}$$

Hence we see that X is a eigenvector of the adjacency matrix of A with eigenvalue $\lambda$ and $\lambda$ must be the largest eigenvalue (using the Perron-Frobenius theorem) of adjacency matrix A and X is the corresponding eigenvector [16,20].
The gist of the eigenvector centrality is to compute the centrality of node as a function of the centralities of its neighbors [16].
**Eigenvector and Eigenvalue :** An Eigenvector of a square matrix A is a non-zero vector v that when the matrix is multiplied by v, yields a constant multiple of v, the multiplier being commonly denoted by $\lambda$. That is

$$Av = \lambda v \tag{7}$$

The number $\lambda$ is called eigenvalue of A corresponding to v.

High eigenvector centrality individuals, however, cannot necessarily perform the roles of high closeness and betweenness. They do not always have the greatest local influence and may have limited brokering potential.

**Frequency:** Frequency of an author is the total number of publication.
**Citation count:** A citation can represent many types of links, such as links between authors, publication, journals and conferences. When a researcher refers to another author's works in their own publication work, they cite it. A citation index is a compilation of all the cited references form articles published during a particular year or period. A citation index allows determining the researcher impact of publication ac-cording to the number of times it has been cited by other researchers. Self citation is not included in such citation count. However, using citation count alone to judge the quality of research contributions can be unfair to some researchers [35,39].
**H-index:** A popular method to measure the centrality of academic papers. H-index = number of your papers h that have been cited at least h times[36,38]. The use of h-index aims at identifying researchers with more papers and relevant impact over a period of time. For any general set of papers one can arrange these papers in decreasing order of the number of citations they received. The h-index is then the largest rank h = r such that the paper on this rank (and hence also all papers on rank 1 to h) has h or more citations [34,35], . Hence the papers on ranks h + 1, h + 2, have not more than h citations.
**G-index:** The g-index is introduced as an improvement of the h-index of Hirsh to measure the global citation performance of a set of articles and effort to give some weightage to the highly cited paper[37]. This set is ranked in decreasing order of the number of citations that they received, the g-index is the (unique) largest number such that the top g articles received (together) at least $g^2$ citations. The g-index of an author is greater or equal to h-index (g >= h) [34,36] .
**I10-index:** The i10-index indicates the number of academic publication an author has written that have at least 10 citations from other[33]. It is only used by Google scholar.

## 4. Research Professionals' Relationship Network

In recent years, there has been a sharp increasing number of collaboration between research professionals. By publishing articles in association with other authors, researchers show their knowledge sharing activities, which are essential for knowledge creation. Most of the research activities are undertaken in the collaborative way because the research projects are too large for an individual researcher to handle, and also the domain of the project spans across multiple research areas. So it often needs collaboration from multiple researchers who are working prominently in those research areas.

An important result of research professionals' collaboration is the creation of new scientific/research knowledge, exchange knowledge, finding new research query, new publications, new invention, more number of articles publication. The product of research professionals' relationship is papers. These papers will impact the researchers who are read these articles, and researchers may found new ideas and express these ideas in an article. It is a cyclic process of spreading knowledge and innovation in research community.

## 5. Data collection, Cleansing and Analysis

In the process of publishing articles in journals and conferences, publishing and editing books, knowledge does not only exist in a particular researcher's mind but also kept in the papers. Currently, it is not clear which collaboration data is useful for evaluating the research community. Although, there is a large set of potential collaboration data and different source, the collaboration of data source is like joined conference organization, joined research proposal submissions, joined publications, joined conference attendance and teacher-student relationship and the data source is like IEEE Xplore, Google scholar, DBLP etc. For this analysis, we only considered joined publication as a measure and the data source is IEEE Xplore. IEEE Xplore provides the facility to search articles using various dimensions like date, publication type etc. The search result can be downloaded into CSV (Comma Separated Values) format. We search for articles detail topic wise for the period Jan-2000 to Jan-2014 and export the search result. Exported search result is in CSV format. The CSV file has 33 fields. The field names are as follows, "Document Title, Authors, Author Affiliations, Publication Title, Publication Date, Publication Year, Volume, Issue, Start Page, End Page, Abstract, ISSN, ISBN, EISBN, DOI, PDF Link, Author Keywords, IEEE Terms, INSPEC Controlled Terms, INSPEC Non-Controlled Terms, DOE Terms, PACS Terms, MeSH Terms, Article Citation Count, Patent Citation Count, Reference Count, Copyright Year, Online Date, Date Added To Xplore, Meeting Date, Publisher, Sponsors, Document Identifier". We have filtered the data and select only Document Title, Authors, Publication Title, Article Citation Count, Publisher and Document Identifier. Authors' name of an articles are stored in a single column which is separated by a ";". Author and its co-author are extracted based on delimiter ";" and store each author name in individual column.

After extraction of author's name we found that some of author's name is unable to read, so data cleaning has become necessary to clean such type of names by replacing actual name. We omitted the name of authors and marked each of them a personal number from 1 to 61546.After cleaning of publication data, 26802 publication and 61546 authors were finally available for analysis.We found that a paper has been written by minimum one author and maximum of 59 authors. Averagely each of these papers is written around two to four authors. In the dataset there are two papers that were composed by 59 authors which has the most co-published paper and 3230 papers is written by only one authors which has the least co-published paper. The most prolific writer is **Lau, Y.Y.**, who has published 54 papers. Author **Sasaki, M.** have maximum number of co-authors is 63 and some of authors work individually or have one or two co-authors.

## 6. Generating Relationship Network of Research Professionals'

Research professionals' relationship network is drawn based on co-author; means an author have connection with those authors who have published articles in journal and conference together, publish and edit books together, publish articles in transaction together which can indicate individuals' status and the scientific collaboration. Based on the available publication data of researcher, we can build a network matrix which is representing the relationship between researchers according to following technique: Lets, there are three papers: P1, P2 and P3. P1 and P2 are conference paper and P3 is journal. P1 has two authors' a1 and a2 and citation is 12, P2 has three authors' b1, a3 and b2 and citation is 2. P3 has five authors' c1, a3, b2, c2 and b1 and citation is 11. It shows like this:P1-{a1, a2} P2-{b1, a3, b2} P3-{c1, a3, b2, c2, b1}

After that we extract author and its co-author which are involved in their research work and calculate the relationship weight according to their total citation count of all papers which is published together. We link the author {a1-a2}, {b1- a3, b1 - b2, a3 - b2}, {c1 - a3, c1 - b2, c1 - c2, c1 - b1, a3 - b2, a3 - c2, a3 - b1, b2 - c2, b2 - b1, c2 - b1} [5,7,31,18,30]. Here a node represents researchers and a link between two nodes represents the publication co-authorship relationship between nodes. After generating a network, we can start the analysis.

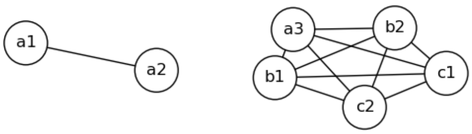


Fig. 1. Researcher relationship network an example

## 7. Key author analysis and result

In this part our aim is to find out key author in the research professionals' relationship network. We generate research professionals'relationship network based on available publication data and convey the frequency, citation count, h-index, g-index, i10-index, degree centrality, closeness centrality, betweenness centrality and eigenvector centrality. Now, we calculate frequency, citation count, h-index, g-index and i10-index of every researcher and also calculate normalized degree centrality, closeness centrality, betweenness centrality and eigenvector centrality by using python and networkx [17]. After that we export all results in MYSQL database and arrange the authors in descending order, then select top 10 authors from each measures and combined it and obtain 51 important authors shown in Table 1, in whichthe sequence of the authors are listed firstly by the decreasing order of frequency and then by citation count, h-index, g-index, i10-index, degree centrality, betweenness centrality and eigenvector centrality of authors. These 51 authors' data are available for analysis. After analysis, we found that only three authors Lau,Y.Y( Author id: 16947), Hirzinger,G.(Author id: 61365) and Gilgenbach, R.M (Author id: 8935) are present in all measure but they do not having high value in all measures. According to frequency Lau, Y.Y is the key author, according to citation count Candes, E.J (author id: 189)becomes a key author, according to h-index, g-index and i10-index author Giannakis, G.B (author id: 29358) is the key author, according to degree centrality Hirzinger, G (author id: 61365) is the key author, according to closeness centrality Wang, Y. (author id: 13513)is the key author, Kim J. (author id: 41078) is the key author based on betweenness centrality and author Mizuno, T. (author id: 48688) is key author based on eigenvector centrality.

Table 1. Top 51 Authors By centrality and citation indices in Research professionals' relationship network

| No. | Author ID | Frequency | Citation | h-index | g-index | i10-index | Degree Centrality | Closeness Centrality | Betweenness Centrality | Eigenvector Centrality |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 16947 | 54 | 159 | 8 | 12 | 6 | 0.0015584245 | 0.0597302006 | 0.0018587107 | 0.0000000013 |
| 2 | 61365 | 53 | 692 | 15 | 26 | 19 | 0.0017552781 | 0.0619001330 | 0.0046731385 | 0.0000000001 |
| 3 | 8935 | 47 | 152 | 8 | 12 | 7 | 0.0013287620 | 0.0597217937 | 0.0013977647 | 0.0000000013 |
| 4 | 56493 | 45 | 533 | 13 | 23 | 15 | 0.0006889877 | 0.0642255172 | 0.0035885079 | 0.0000000001 |
| 5 | 51188 | 42 | 720 | 17 | 26 | 23 | 0.0005085385 | 0.0632034440 | 0.0023709283 | 0.0000000017 |
| 6 | 29358 | 41 | 2223 | 23 | 41 | 32 | 0.0006889877 | 0.0644432554 | 0.0048778461 | 0.0000000005 |
| 7 | 53986 | 38 | 247 | 9 | 15 | 8 | 0.0008694368 | 0.0631794619 | 0.0013450811 | 0.0000036042 |
| 8 | 19686 | 34 | 223 | 9 | 14 | 7 | 0.0002952804 | 0.0565234339 | 0.0012528466 | 0.0000000000 |
| 9 | 10748 | 30 | 7 | 1 | 2 | 0 | 0.0005577519 | 0.0682436522 | 0.0006515796 | 0.0000000085 |
| 10 | 24476 | 29 | 354 | 11 | 18 | 12 | 0.0006397743 | 0.0588157407 | 0.0014161240 | 0.0000000001 |
| 11 | 51980 | 28 | 713 | 12 | 26 | 13 | 0.0008858413 | 0.0578717700 | 0.0040615117 | 0.0000000000 |
| 12 | 20117 | 27 | 182 | 6 | 13 | 4 | 0.0011647173 | 0.0653012831 | 0.0070899167 | 0.0000000002 |
| 13 | 42040 | 26 | 518 | 13 | 22 | 15 | 0.0005085385 | 0.0642873534 | 0.0033032907 | 0.0000000033 |
| 14 | 36257 | 24 | 965 | 15 | 24 | 17 | 0.0003773028 | 0.0502485463 | 0.0003132740 | 0.0000000000 |
| 15 | 1949 | 24 | 156 | 6 | 12 | 5 | 0.0010006726 | 0.0682135710 | 0.0098147929 | 0.0000000554 |
| 16 | 53589 | 23 | 978 | 14 | 23 | 16 | 0.0005577519 | 0.0589337943 | 0.0031531625 | 0.0000000058 |
| 17 | 21811 | 23 | 563 | 15 | 23 | 16 | 0.0005577519 | 0.0663969389 | 0.0041624579 | 0.0000000000 |
| 18 | 11373 | 22 | 672 | 8 | 22 | 7 | 0.0005085385 | 0.0453370786 | 0.0008980776 | 0.0000000000 |
| 19 | 56331 | 20 | 1284 | 15 | 20 | 18 | 0.0003444938 | 0.0546222400 | 0.0017207479 | 0.0000000000 |
| 20 | 25877 | 19 | 1188 | 12 | 19 | 15 | 0.0004265162 | 0.0598356220 | 0.0006995759 | 0.0000000001 |
| 21 | 58514 | 19 | 339 | 13 | 18 | 15 | 0.0005249430 | 0.0648566383 | 0.0006914223 | 0.0000000000 |
| 22 | 13513 | 16 | 112 | 5 | 10 | 2 | 0.0011647173 | 0.0739435385 | 0.0123000916 | 0.0000000871 |
| 23 | 56899 | 16 | 70 | 6 | 8 | 2 | 0.0013287620 | 0.0645821540 | 0.0051525725 | 0.0024905640 |
| 24 | 33095 | 15 | 201 | 7 | 14 | 7 | 0.0010334815 | 0.0718858696 | 0.0118138227 | 0.0000001434 |
| 25 | 46389 | 15 | 50 | 4 | 7 | 2 | 0.0012795485 | 0.0718250289 | 0.0134299745 | 0.0000000004 |
| 26 | 41078 | 14 | 111 | 4 | 10 | 1 | 0.0014764022 | 0.0725775985 | 0.0168249725 | 0.0000000200 |
| 27 | 25174 | 11 | 25 | 4 | 5 | 0 | 0.0009678636 | 0.0704856631 | 0.0067295431 | 0.0000000824 |
| 28 | 3667 | 10 | 52 | 4 | 7 | 2 | 0.0014107843 | 0.0645067771 | 0.0028558503 | 0.0113793075 |
| 29 | 24461 | 9 | 18 | 3 | 4 | 0 | 0.0007217966 | 0.0717564614 | 0.0076363362 | 0.0000000072 |
| 30 | 16530 | 7 | 89 | 3 | 7 | 2 | 0.0006561787 | 0.0703484486 | 0.0027333870 | 0.0000000432 |
| 31 | 27968 | 7 | 4 | 1 | 2 | 0 | 0.0006069653 | 0.0707822660 | 0.0079069144 | 0.0000000425 |
| 32 | 20621 | 6 | 11 | 2 | 3 | 0 | 0.0010334815 | 0.0719995838 | 0.0052392052 | 0.0000000833 |
| 33 | 48688 | 6 | 10 | 2 | 3 | 0 | 0.0016404469 | 0.0628135023 | 0.0013225435 | 0.1334739264 |
| 34 | 5340 | 5 | 6 | 2 | 2 | 0 | 0.0015912335 | 0.0624567989 | 0.0006998918 | 0.1333693085 |
| 35 | 56856 | 5 | 4 | 1 | 2 | 0 | 0.0015584245 | 0.0630196426 | 0.0012480845 | 0.1295166699 |
| 36 | 11029 | 4 | 2350 | 3 | 4 | 2 | 0.0001312357 | 0.0475522970 | 0.0003905120 | 0.0000000000 |
| 37 | 33989 | 4 | 55 | 2 | 4 | 2 | 0.0012795485 | 0.0721460914 | 0.0061896466 | 0.0000003761 |
| 38 | 36605 | 4 | 7 | 2 | 2 | 0 | 0.0011975262 | 0.0579120265 | 0.0000201996 | 0.1288715982 |
| 39 | 17849 | 4 | 6 | 2 | 2 | 0 | 0.0011319083 | 0.0579110066 | 0.0000020302 | 0.1284210454 |
| 40 | 189 | 3 | 2426 | 3 | 3 | 3 | 0.0000492134 | 0.0431370853 | 0.0000152403 | 0.0000000000 |
| 41 | 17217 | 3 | 1878 | 3 | 3 | 2 | 0.0000656179 | 0.0404585477 | 0.0000457187 | 0.0000000000 |
| 42 | 25056 | 3 | 1877 | 3 | 3 | 2 | 0.0000328089 | 0.0372173858 | 0.0000152403 | 0.0000000000 |
| 43 | 17718 | 3 | 1493 | 3 | 3 | 3 | 0.0000820223 | 0.0534990176 | 0.0000304795 | 0.0000000002 |
| 44 | 1370 | 3 | 5 | 2 | 2 | 0 | 0.0011319083 | 0.0579110066 | 0.0000020302 | 0.1284210454 |
| 45 | 11735 | 3 | 5 | 2 | 2 | 0 | 0.0011319083 | 0.0579110066 | 0.0000020302 | 0.1284210454 |
| 46 | 20758 | 3 | 5 | 2 | 2 | 0 | 0.0011319083 | 0.0579110066 | 0.0000020302 | 0.1284210454 |
| 47 | 22469 | 3 | 5 | 2 | 2 | 0 | 0.0011319083 | 0.0579110066 | 0.0000020302 | 0.1284210454 |
| 48 | 23930 | 3 | 5 | 2 | 2 | 0 | 0.0011319083 | 0.0579110066 | 0.0000020302 | 0.1284210454 |
| 49 | 18993 | 1 | 2333 | 1 | 1 | 1 | 0.0000328089 | 0.0431369438 | 0.0000000000 | 0.0000000000 |
| 50 | 16409 | 1 | 1449 | 1 | 1 | 1 | 0.0000328089 | 0.0479742511 | 0.0000000000 | 0.0000000000 |
| 51 | 33284 | 1 | 1449 | 1 | 1 | 1 | 0.0000328089 | 0.0479742511 | 0.0000000000 | 0.0000000000 |

Author with high frequency are key or prominent since they have published more number of papers with the others. But if an author published more number of papers with less quality or does not have any impact in research community then we can't say an author is prominent or key in the network. According to citation indices like citation count, h-index, g-index and i10-index the authors are prominent or key based on citation of their publication but mostly research work is done in the collaborative way and citation of each publication is equal to all authors means a paper has been written by 4 authors and citation is 12 then each author uses same citation count for their index calculation. So we can't say authors having high citation count, h-index, g-index and i10-index value are prominent. Author with high degree centrality are important since they have more relationship with the others but in this case we can't say an author having high degree centrality are important because if an author having a connection with more no of students then the degree centrality is high but their productivity is less. So, degree centrality is not suited for finding key author in the network. Authors who have a good closeness centrality are also significant means an author is nearest to average no. of authors. Authors who have a good betweenness centrality are also prominent means an author is frequently control information flow in the network and an author who have a good Eigenvector centrality is also prominent in the network means an author have a connection with a high score eigenvector centrality authors. Closeness and betweenness centrality are useful in information flow network but research professionals' relationship network is a co-authors relationship network. The product of research professionals' relationship network is research articles and the research articles quality is depends on authors of article, if the authors are highly qualified and they have relation with other high qualified researchers then have a probability to publish a good quality papers. So eigenvector centrality is more suited for finding key author in the network.

## 8. Conclusion

In this paper we have investigated the research professionals relationship network and find out the key author in the network based on different social network analysis metrics like normalized degree centrality, closeness centrality, betweenness centrality, eigenvector centrality and citation indices like frequency, citation count, h-index, g-index and i10-index. We found that different social network analysis metrics and citation indices gives different value for each researcher but some of the author having high value in all h-index, g-index and i10-index. It means a researcher who is the key author in network based on degree centrality; he/she is not guarantee to prominent based on other centrality and citation indices but some of the authors is present in all measure in top 10. So, we can say that these authors are key author in research professionals relationship network but they dont have high value in all measures. If we compare citation indices and all centrality measures, eigenvector centrality is more suited for finding key author in the network because the eigenvector centrality of nodes is depends on the neighbors node eigenvector centrality. If the neighbors eigenvector centrality is high then have a probability to gain high eigenvector centrality value of a node. A node having high eigenvector centrality than other is called key node in the network.

## References

1. Liu, Bing. Web data mining.2007;Springer.
2. Kas, Miray and Carley, L Richard and Carley, Kathleen M. Monitoring social centrality for peer-to-peer network protection.*IEEE Communications Magazine*; vol. 51,no. 12, p. 155 - 161. 2013.
3. Abbasi, Alireza and Altmann, Jörn. On the correlation between research performance and social network analysis measures applied to research collaboration networks. *44th Hawaii International Conference on System Sciences (HICSS), 2011* ; p. 1-10, 2011,IEEE.
4. Friedl, Dipl-Math Bettina and Heidemann, Julia and others. A critical review of centrality measures in social networks. *Business & Information Systems Engineering*; vol 2,no. 6, p. 371-385. 2010,Springer.
5. Wang, Bing and J. Yang. To form a smaller world in the research realm of hierarchical decision models. *Proc. PICMET'11*; 2011,PICMET.
6. Wang, Bing and Yao, Xiaotian. To form a smaller world in the research realm of hierarchical decision models. *International Conference on Industrial Engineering and Engineering Management (IEEM), 2011 IEEE*; p. 1784-1788, 2011,IEEE.
7. Liu, Xiaoming and Bollen, Johan and Nelson, Michael L and Van de Sompel. Herbert, Co-authorship networks in the digital library research community. *Information processing & management*; vol. 41, no. 6, p. 1462-1480, 2005, Elsevier.
8. Tang, Jie and Zhang, Duo and Yao, Limin. Social network extraction of academic researchers. *Seventh IEEE International Conference on Data Mining, 2007. ICDM 2007.*; p. 292-301, 2007, IEEE.
9. Said, Yasmin H and Wegman, Edward J and Sharabati, Walid K and Rigsby, John T. RETRACTED: Social networks of author-coauthor relationships. *Computational Statistics & Data Analysis*; vol. 52, no. 4, p. 2177-2184, 2008, Elsevier.
10. Umadevi, V. Automatic Co-authorship Network Extraction and Discovery of Central Authors. *International Journal of Computer Applications*; vol. 74, no. 4,p. 1-6, 2013, Citeseer.

11. Jin, Jianzhi and Xu, Kaihua and Xiong, Naixue and Liu, Yuhua and Li, Guoqiang. Multi-index evaluation algorithm based on principal component analysis for node importance in complex networks. *Networks, IET*; vol. 1, no. 3, p. 108-115, 2012, IET.
12. Spizzirri, Leo. Justification and Application of Eigenvector Centrality.
13. Correa, Carlos D and Crnovrsanin, Tarik and Ma, Kwan-Liu and Keeton, Kimberly. The derivatives of centrality and their applications in visualizing social networks. 2009, Citeseer.
14. Correa, Carlos and Crnovrsanin, Tarik and Ma, Kwan-Liu. Visual reasoning about social networks using centrality sensitivity; *IEEE Transactions on Visualization and Computer Graphics*; vol. 18, no. 1,p. 106-120, 2012, IEEE
15. Bonacich, Phillip and Lloyd, Paulette. Eigenvector-like measures of centrality for asymmetric relations. *Social Networks*; vol. 23, no. 3, p. 191-201, 2001, Elsevier.
16. Newman, Mark EJ. The mathematics of networks. *The new palgrave encyclopedia of economics*; vol. 2, p. 1-12, 2008, Citeseer.
17. D. A. S. Aric A. Hagberg, P. J. Swart. Exploring network structure, dynamics,and function using networkx. *proceedings of the 7th python in science conference (scipy 2008)*; 2008.
18. He, Yulan and Cheung Hui, Siu. Mining a web citation database for author co-citation analysis. *Information processing & management*; vol. 38,no. 4,p. 491-508, 2002, Elsevier
19. Ding, De-wu and He, Xiao-qing. Application of eigenvector centrality in metabolic networks. *2nd International Conference on Computer Engineering and Technology (ICCET), 2010* ; vol. 1, p. V1-89, 2010,IEEE.
20. Straffin, Philip D. Linear algebra in geography: Eigenvectors of networks. *Mathematics Magazine*;, p. 269 -276, 1980, JSTOR.
21. Deng, Qiuhong and Wang, Zhiping. Degree centrality in scientific collaboration supernetwork. *International Conference on Information Science and Technology (ICIST), 2011* ; p. 259-262, 2011, IEEE.
22. Kato, Yukiya and Ono, Fumie. Node centrality on disjoint multipath routing. *73rd Conference on Vehicular Technology (VTC Spring), 2011 IEEE* ; p. 1-5, 2011, IEEE.
23. Estrada, Ernesto and Rodriguez-Velazquez, Juan A. Subgraph centrality in complex networks. *Physical Review E*; vol. 71, no. 5, p. 056-103, 2005, APS.
24. Borgatti, Stephen P and Everett, Martin G. A graph-theoretic perspective on centrality. *Social networks*; vol. 28, no. 4, p. 466-484, 2006, Elsevier.
25. Wang, Gaoxia and Shen, Yi and Luan, Enjie. A measure of centrality based on modularity matrix. *Progress in Natural Science*; vol. 18, no. 8, p. 1043-1047, 2008, Elsevier.
26. Gómez, Daniel and Figueira, José Rui and Eusébio, Augusto. Modeling centrality measures in social network analysis using bi-criteria network flow optimization problems. *European Journal of Operational Research*; vol. 226, no. 2, p. 354-365, 2013, Elsevier.
27. Erkan, Günes and Radev, Dragomir R. LexRank: Graph-based lexical centrality as salience in text summarization. *J. Artif. Intell. Res.(JAIR)*; vol. 22, no. 1, p. 457-479, 2004.
28. Lohmann, Gabriele and Margulies, Daniel S and Horstmann, Annette and Pleger, Burkhard and Lepsien, Joeran and Goldhahn, Dirk and Schloegl, Haiko and Stumvoll, Michael and Villringer, Arno and Turner, Robert. Eigenvector centrality mapping for analyzing connectivity patterns in fMRI data of the human brain. *PloS one*; vol. 5, no. 4, 2010, Public Library of Science.
29. Wasserman, Stanley. Social network analysis: Methods and applications. vol. 8, 1994, Cambridge university press.
30. Said, YH and Wegman, EJ and Sharabati, WK and Rigsby, JT. Social networks of author-coauthor relationships (Retraction of vol 52, pg 2177, 2008). *COMPUTATIONAL STATISTICS & DATA ANALYSIS*; vol. 55, no. 12, p. 3386-3386, 2011,ELSEVIER SCIENCE BV PO BOX 211, 1000 AE AMSTERDAM, NETHERLANDS.
31. Newman, Mark EJ. Coauthorship networks and patterns of scientific collaboration. *Proceedings of the National Academy of Sciences*; vol. 101, no. suppl 1, p. 5200-5205, 2004, National Acad Sciences.
32. Borgatti, Stephen Peter. Centrality and AIDS. *Connections*; vol. 18, no. 1, p. 112-114, 1995.
33. *http://ieeexplore.ieee.org/xpl/opac.jsp*
34. Majdi Maabreh and Izzat M.Alsmadi. A Servey of impact and citation indices: Limitations and isssues. *International Journal of Advanced Science and Technology*; vol. 40, p. 35-54, 2012.
35. Kumar, M Jagadesh. Evaluating Scientists: Citations, Impact Factor, h-Index, Online Page Hits and What Else?. *IETE Technical Review*; vol. 26, no. 3, p. 165, 2009, Medknow.
36. Huang, Mu-hsuan and Chi, Pei-shan. A comparative analysis of the application of h-index, g-index, and a-index in institutional-level research evaluation ; *Journal of Library and Information Studies*; vol. 8, no. 2, p. 1-10, 2010.
37. Egghe, L.. Theory and practice of the G-index. *Scientometrics*; vol. 69, no. 1, p. 131-152, 2006.
38. Hirsch, Jorge E. An index to quantify an individual's scientific research output. *Proceedings of the National academy of Sciences of the United States of America*; vol. 102, no. 46, p. 16569-16572, 2005, National Acad Sciences.
39. Egghe, Leo and Rousseau, Ronald. An h-index weighted by citation impact. *Information Processing & Management*; vol. 44, no. 2, p. 770-780, 2008, Elsevier.