ISSN 2249-3875

International Journal of Pharmaceutical Erudition

Review Article

Scientific Journal: Quality or Quantity

Goyal Anju*, Jain Apoorva

B. N. P. G. Girls College of Pharmacy, Udaipur (Rajasthan), India 313001

More influential authors are more likely to produce high quality and thus highly cited papers, and well-cited papers can bring greater acknowledgments to their authors. Simply counting the number of citation cannot represent well the true prestige. Without distinguishing between citations, the citation from a good paper with high impact will have the same weight as citations with lower impact. This article describes the nature and motivations for citations in the scientific literature, and the explicit and implicit knowledge. It also describes computational methods that may be applied to large quantities of citation data in order to understand and characterize the literature, its key actors and events, and identified a number of techniques to be used to characterize the literature.

Key Words: *Scientific Journal, citation impact, Citation frequency, citation index, page rank*

Introduction

A reader would never be able to reliably establish the rationale behind a piece of research, understand where it came from, or identify its flaws. Citations in science thus become not only useful, but essential. It may be used as a form of recognition of the work of others, or means of explanation or contextualization of research..Cynics suggest that citations are motivated more by self-interest, and are primarily used as a persuasive tool to support arguments by associating the author with established work. Citations in scientific literature are sign-posts to important prior works, and by following this trail a researcher can track the origins of a subject back to its roots, identify key authors and ideas and construct a mental picture of how the field has developed and where its limits lie.[1] Online scientific literature provides abundant of reputation evidence in terms of researchers' (author) publications, as well as the citation relationships among these publications, both of which can be taken advantage of in evaluating researchers reputations. To evaluate the reputation of especially within one scientific domain, there are typically two basic approaches. One is called the content-based approach, in which relevant documents representing expertise of a re-searcher can be considered, and information retrieval models can be

***Address for Correspondence**

goyalanjugoyal@yahoo.co.in

applied to evaluate the relevance between these documents and thus authors with the query topic. [2][3][4][5]

A **citation** is a reference to a published or unpublished source (not always the original source). More precisely, a citation is an abbreviated alphanumeric expression (e.g. [Newell84]) embedded in the body of an intellectual work that denotes an entry in the bibliographic references section of the work for the purpose of acknowledging the relevance of the works of others to the topic of discussion at the spot where the citation appears. A prime purpose of a citation is intellectual honesty: to attribute prior or unoriginal work and ideas to the correct sources, and to allow the reader to determine independently whether the referenced material supports the author's argument in the claimed way[6]. Bibliographies, and other list-like compilations of references, are generally not considered citations because they do not fulfill the true spirit of the term: deliberate acknowledgment by other authors of the priority of one's ideas.[7]

Scientific citation is the process by which conclusions of previous scientists are used to justify experimental procedures, apparatus, goals or theses. In patent law the citation of previous works,or prior art, helps establish the uniqueness of the invention being described .The focus in this practice is to claim originality for commercial purposes, and so the author appears to be strongly motivated to avoid citing works that cast doubt on its uniqueness.[8] Scientific citations that appear in the reference list of articles from a large number of scientific journals. This database is produced by the Institute for Scientific Information [ISI] in, Philadelphia. The references are listed to show by whom, within a given period of time, and the results are published in the Science Citation Index (SCI). [9][10]

Citation impact recent development in research on citation impact is the discovery of universality, or citation impact patterns that hold across different disciplines in the sciences, social sciences, and humanities.[11] This finding has suggested a *Universal Citation Impact Measure* that extends the H-index by properly rescaling citation counts and resorting publications; however the computation of such a universal measure requires the collection of extensive citation data and statistics for every discipline and year.[12] It is the number of times a paper has been cited. ISI indexes articles in a core set of journals. So citations from publications like conference proceedings and books may not be represented in the citation count for a particular paper. [13][14]

Citation index Citation Indexes track

ISSN 2249-3875

these citations and compile statistics that allow researchers to see which articles have cited which previous articles. Using a citation index, a researcher can discover which authors and which journals are most often cited in works in the field. This is helpful in determining who is having the biggest effect, or impact, on that field.[15] Automated citation indexing[16] has changed the nature of citation analysis research,[14] A citation index is a compilation of all the cited references from journal articles indexed in the database. We can look up a published paper in order to find journal article that have cited it.

Citation index database The institute for Scientific Information (ISI) publishes the Science Citation Index (SCI), which provides access to current bibliographic information and cited references.The online version of SCI, called WEB OF SCIENCE. [13]

Citation frequency Modern scientists are sometimes judged by the number of times their work is cited by others this is actually a key indicator of the relative importance of a work in science. [8]

Citation Analysis

Citation analysis is a branch of bibliometrics, a science that uses numerical methods to draw conclusions about the role of publications and their authors in the growth of knowledge. The main source of raw data for this science is the counting of books journals, authors, institutions, themes and so on and the counting of links that can be established between all of these through citations.

- Zipf's 'Productivity of Words' law says that if the words and phrases in an article are sorted in decreasing order according to the number of occurrences, then the rank of a word or phrase will be inversely proportional to its number of occurrences.
- Lotka's 'Productivity of Authors' law notes that the number of articles produced by a small number of authors tends to account for the majority of the publications.
- Bradford's 'Productivity of Journals' or 'Scattering' law suggests that the majority of articles produced that are relevant to a particular field will be published in a small number of journals.

Theoreticians and library scientists have been seeking ways to exploit these and other principles for the purposes of automating the characterization of the literature via computer analysis.[1] Automated citation analysis enables the estimation of importance or reputations of publications and authors, citation analysis mainly focused on counting the number of citations. Under this scheme, an author

will have higher reputation if he can be cited by many other authors.[2] The process of acknowledging or citing the author, year, title, and locus of publication (journal, book, or other) of a source used in a published work. Such citations can be counted as measures of the usage and impact of the cited work. This is called citation analysis or bibliometrics. Among the measures that have emerged from citation analysis are the citation counts for:

* an individual article (how often it was cited);

*an author (total citations, or average citation count per article;

*a journal (average citation count for the articles in the journal).[14]

Citation Content

Citation content can vary depending on the type of source and may include:

* *Book:* author(s), book title, publisher, date of publication, and page number(s) if appropriate.[17][18]

**Journal:* author(s), article title, journal title, date of publication, and page number(s).

**Newspaper:* author(s), article title, name of newspaper, section title and page number(s) if desired, date of publication.

**Web site:* author(s), article and publication title where appropriate, as well as a URL, and a date when the site was accessed.

**Play:* inline citations offer part, scene, and line numbers, the latter separated by periods: 4.452 refers to scene 4, line 452. For example, "In Eugene Onegin, Onegin rejects Tanya when she is free to be his, and only decides he wants her when she is already married" (Pushkin 4.452-53).[19]

**Poem:* spaced slashes are normally used to indicate separate lines of a poem, and parenthetical citations usually include the line number(s).

Unique Identifiers

Along with information such as author(s), date of publication, title and page numbers, citations may also include unique identifiers depending on the type of work being referred to.

*Citations of books may include an International Standard Book Number (ISBN).

*Specific volumes, articles or other identifiable parts of a periodical, may have an associated Serial Item and Contribution Identifier (SICI).

*Electronic documents may have a digital object identifier (DOI).

*Biomedical research articles may have a PubMed Identifier (PMID). A citation number, used in some citation systems, is a number or symbol added inline and usually in superscript, to refer readers to a

footnote or endnote that cites the source. In other citation systems, an inline parenthetical reference is used rather than a citation number, with limited information such as the author's last name, year of publication, and page number referenced; a full identification of the source will then appear in an appended bibliography. citation systems Broadly speaking, there are two citation systems.[20]

Note Systems involve the use of sequential numbers in the text which refer to either footnotes (notes at the end of the page) or endnotes (a note on a separate page at the end of the paper) which gives the source detail. The notes system may or may not require a full bibliography, depending on whether the writer has used a full note form or a shortened note form.In the humanities, many authors use footnotes or endnotes to supply anecdotal information. In this way, what looks like a citation is actually supplementary material, or suggestions for further reading.[21]

Parenthetical Referencing also known as Harvard referencing where full or partial, in-text citations are enclosed within parentheses and embedded in the paragraph, as opposed to the footnote style.Depending on the choice of style, fully cited parenthetical references may require no end section. Alternately a list of the citations with complete bibliographical references may be included in an end section sorted alphabetically by author's last name.[6]

Patent in Citation

Patent document that follows WIPO standards contain a section titled 'List Prior Art Documents'. Under this section there is a typically a list of citations. The 'Cited Patents' and a list of 'non-patent citations'. The 'cited patents' as the name suggests, are patents that are either the party seeking the patent(applicant) or the Examiner(from the Office of the relevant country) consider relevant for understanding either the background in which the patent is set or the scope of the patent disclosed. Patient citation analysis is the study of these patient citations for potentially determining patent's value or identifying potential licensees or the other leads based on the citation trends of a corporation's patents by other companies in the same or different market/industry.

The categories of citations may be present in both 'forward cites' and 'backward cites' of a patent: both are indicators of the importance/value of the patent. forward cites are citations by patents granted/applied for after publication/grant of the patent under investigation that cite to Pi..Backward cites are the patents that Pi itself cite to. Forward cites are instinctively relevant in determining the

value of patent. This data can be further analyzed to find the average number of forward citations received by a patent per year. Backward citations are considerably important in determining the value and scope of a patient because they help delimit the scope of claims.[22] The number of times a patent is cited to be a significant metric of patent quality and thus of innovation [8].

Citation Networks

Network science is a set of mathematical techniques for describing pair-wise relationships (paths or 'edges') between entities (nodes or 'vertices') in a network. Networks can be defined using a number o f metrics that are common to all network 'Graphs'. In the scientific literature, citations can be understood as a form of contact (paths), between articles (nodes). the concept of a '*research front*' ,is the collection of articles that research scientists are actively citing. The research front is a changing collection of concepts and research issues that defines the current 'limits' of a field. Current research and new publications try to move the research front forward, and so tend to cite the articles that authors perceive as comprising that front. Over time the front moves on and the articles being actively cited change. This explains the pattern of 'transient' articles. The combination of classic and transient articles makes up what Price called the *'intellectual base'* of a field, i.e. the collection of research work and insight that defines it.[1] To increasing web based access and use of scholarly literature, several others have developed and proposed innovative methods and tools to rank scholarly journals and refine the evaluation of both science and scientists both within and outside the citation-based systems.[23]

The citation network is one form of social network in which scientific factors, like authors and papers, can be represented as nodes, and their mutual interactions, like co-authorship and citation, can be modeled as edges.[24][25] Different kinds of citation network, including coauthor network for authors and citation reference network for papers and take advantage of their mutual rein performance to improve reputation ranking performance.

Multi-Type Citation Network Framework

Two versions of the framework have been considered, reflecting different relationships among factors.

Notation and Preliminaries

In the multi-type citation network, different kinds of social factors, as well as their mutual relationships have been considered and integrated. The citation network can be formally denoted as G =

ISSN 2249-3875

International Journal of Pharmaceutical Erudition

(V, E), where

• V is a set of nodes, representing social factors. In our current integrated network, V is combination of four different types of social actors: authors, papers, affiliations and venues. • E is a set of directed edges, representing relationships among every pair of social actors. All the possible relationships we may have are the relationship between authors, papers, affiliations and venues. Due to different relationships among the four types of social actors we can consider, we construct two versions of the multi-type citation network, to which we refer as 4-T graph version-1 (4-T) and 4-T graph version-2 (4-TV2) respectively.

Framework version-1

In 4-T graph version-1,the citation relationship among every pair of social factor types. The graph (shown in Fig. 1) can be viewed as a combination of subgraphs, including those representing each of the types of social factors:

1. Author Graph GAu. There would be one edge from author aui to author auj if they coauthored at least one paper or if author aui cites author auj . We say that author aui cites author author auj if and only if there is at least one publication of aui that cites one of the publications of auj . We do not count the number of co-authorship or citationship in the current framework, and thus there would be only one edge between two authors even though they coauthored more than once. The same mechanism works for other subgraphs defined in the following.

2. Paper Graph GP. There would be one edge from paper pi to pj if pi cites pj in its references.

3. Affiliation Graph GAf . There would be one edge from afi to afj if two authors, each of which comes from afi and afj respectively, coauthor in at least one paper, or there is at least one paper produced in affiliation afi that cites one of the publications from afj.

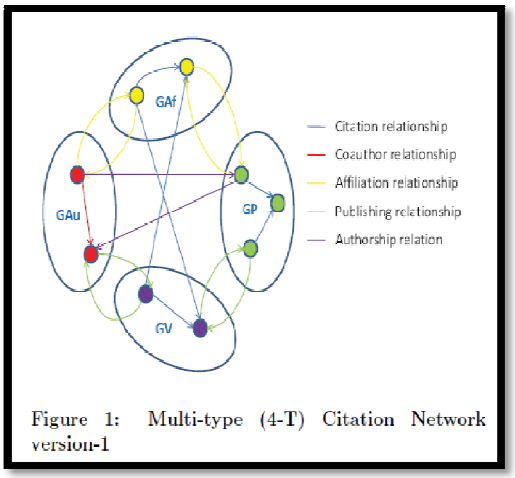

Figure 1: Multi-type (4-T) Citation Network version-1

4. Venue Graph GV . One edge will be drawn from vi to vj if there is at least one paper which is published in vi that cites one of the papers published in vj. As well as graphs that relate one type of social actor to another:

(i) Bipartite AuthorPaper Graph GAuP . There would be one edge from aui to pj , if aui is one of the authors of pj .

ISSN 2249-3875

International Journal of Pharmaceutical Erudition

Correspondingly, there would one edge from pj to aui, indicating that it is written by aui.

(ii) Bipartite Author Affiliation Graph GAuAf . One edge would be drawn from aui to afj and afj to aui, if aui belongs to the affiliation of afi. One distinct author may belong to different affiliations in different period of time; thus it is possible for one author node to point to several affiliation nodes.

(iii) Bipartite Author Venue Graph GAuV. If there is at least one paper written by aui and published in vj, there would be a corresponding edge from aui to vj and from vj to aui.

(iv) Bipartite Paper Affiliation Graph GPAf . One edge will go from paper pi to affiliation afj if pi is written by an author that belongs to afj .

(v) Bipartite Paper Venue Graph GPV . One edge will go from pi to vj and vj to pi if pi is published in vj .

(vi) Bipartite Affiliation Venue Graph GAfV . If there is one paper belonging to affiliation afi published in vj , there would be an edge from pi to vj and from vj to pi.

Framework version-2

There may exist redundant information within edges in version-1, since all relationships are generally inferred from the citations among papers. As a result, we introduce a simplified version of the graph. In this, we only considered the coauthor relationship between authors, while ignoring the citation relationship between them.Affiliation nodes will only be connected with author nodes, and venue nodes will only be connected with paper nodes. There are no direct edges within the affiliation graph and venue graph. The relationships between authors and venues can be related by firstly relating authors to papers, and then papers to venues. A similar process works when representing the relationship between affiliations and papers.

Heterogeneous PageRank

In the original homogeneous PageRank, each node evenly distributes its authority score among its children. Using such an even propagation in the multi-type citation network, author nodes will evenly distribute its authority to other authors, papers, affiliations, and venues (under framework version-1), which may not well

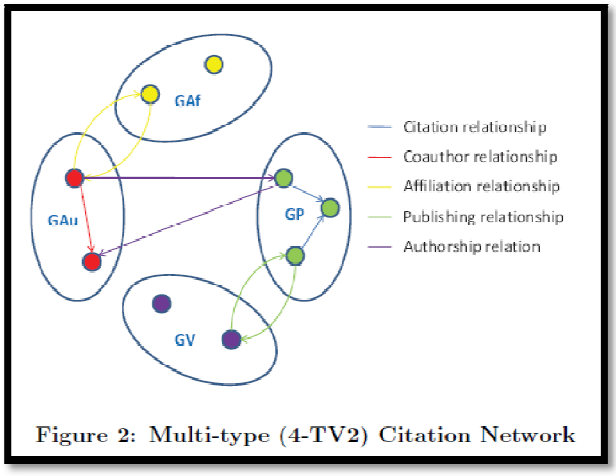

Figure 2: Multi-type (4-TV2) Citation Network

represent the actual interaction possibilities among nodes of different entities. In order to better represent the different impact among multiple types of social actors, we propose a heterogeneous Page Rank algorithm based on the assumption that where there would be a different propagation probability for a node to follow different kinds of out-going links (links to different types of nodes).

$$PR(i) = (1-d)\sum_{j:j\to i}\beta_{ji}\frac{PR(j)}{O(j)_{type(i)}} + d\frac{1}{N} \qquad (1)$$

where:

- j and i are two nodes of any types, where j has out-going link to i.
- d: random jump.
- β_{ji}: is the parameter determining the propagation probability from node j to i. β_{ji} is equal to β_{jk} if node i and node k are of the same type. $\sum_{type(i)}\beta_{ji} = 1$, where node j has an out-going link to i.
- $O(j)_{type(i)}$ is the number of outlinks j has to the nodes of the same type with i.
- N: total number of nodes in the network.

This heterogeneous Page Rank can be described as

The above relation describes the Heterogeneous Page Rank. [2]

The Influence of Journals, Authors and Articles

There are a wide array of metrics and measures used by different researchers and databases to compare the relevant impact of journals, authors and articles. All reflect the basic idea that the more citations garnered, the greater the impact. [1]Citation data continue to be used to assess science whether it is the impact of individual articles, journals, researchers, research Departments, and even countries [23].

Journal and Article Status- ISI IF, Page Rank and The Y-Factor

This gives a reasonable measure of the medium-term impact of a collection of articles. The same formula could be applied to the set of articles published by an author or institution. A newer method of measuring both article and journal importance, and one with growing popularity, is Page Rank. The degree of prestige conferred is calculated by dividing the 'Page Rank' score of a citing article by the number of articles that it cites. Each cited article is then assigned this score, and the scores summed to give the Page Rank of that article .[1]

The **impact factor**, often abbreviated **IF**, is a measure reflecting the average number of citations to articles published in science and social science journals. It is frequently used as a proxy for the relative importance of a journal within its field, with journals with higher impact factors deemed to be more important than those with lower ones. In a given year, the impact factor of a journal is the average number of citations received per paper published in that

journal during the two preceding years. [26] Journal Impact Factor is from Journal Citation Report (JCR), a product of Thomson ISI (Institute for Scientific Information). JCR provides quantitative tools for evaluating journals. [27] Impact factor is an index based on the frequency with which a journal's articles are cited in scientific publications, a marker of journal quality [28]. The citation rate of a journal is quantified as its impact factor, definition of the impact factor of a journal is that it is in essence the number of times the articles it publishes are cited divided by the number of articles that could be cited [8].The higher the impact factor, the more citations a journal has generated.[15] A journal's impact factor is based on 2 elements: the numerator, which is the number of citations in the current year to any items published in a journal in the previous 2 years, and the denominator, which is the number of substantive articles (source items) published in the same 2 years. [14]

The IF, index and other citations-based parameters for measuring and ranking of journals, institutions and individuals through number crunching have become inextricable part of science and technology.[30] The higher the IF, the better rated is the journal encouraging authors to publish paper.[31]

The Thomson Reuters Impact Factor

The *JCR* provides quantitative tools for ranking, evaluating, categorizing, and comparing journals. The impact factor is one of these; it is a measure of the frequency with which the "average article" in a journal has been cited in a particular year or period. Thus, the impact factor of a journal is calculated by dividing the number of current year citations to the source items published in that journal during the previous two years. The impact factor is useful in clarifying the significance of absolute (or total) citation frequencies [32]

Thomson Reuters does not depend on the impact factor alone in assessing the usefulness of a journal, and neither should anyone else. The impact factor should not be used without careful attention to the many phenomena that influence citation rates, Some of the journals listed in the *JCR* are not citing journals, but are cited-only journals. This is significant when comparing journals by impact factor because the self-citations from a cited-only journal are not included in its impact factor calculation. Self-citations often represent about 13% of the citations that a journal receives. The cited-only journals with impact factors in the *JCR* Journal Rankings and Subject Category Listing may be ceased or suspended journals,

superseded titles, or journals that are covered in the science editions of *Current Contents*®, but not a citation index.

A title change can affects the impact factor for two years after the change is made. The old and new titles are not unified unless the titles are in the same position alphabetically. In the first year after the title change, the impact is not available for the new titles unless the data for old and new can be unified. In the second year, the impact is split. The new title may rank lower than expected because only one year of the source data is introduced in its calculation.[33]

Y Factor

The Y-factor[34][35] is a simple combination of both the ISI IF and the weighted PageRank. Significantly, the authors claim that the resulting journal rankings correspond well to a general understanding of journal status[23].

The Y-Factor, based on multiplying the ISI IF for a collection of articles by the results of weighted version of the Page Rank Algorithm for the same collection of articles. This combined score would take into account both the popularity measured by the IF, and the prestige measured by Page Rank. [1]

Page Rank

Page Rank™ is a software system for ranking web pages developed by Google and has also been used to rank publications and later in the system of weighted Page Rank. The Page Rank algorithm[34] makes a clear distinction between popularity and expert appreciation or prestige of published research. . Thus the weighted version of the popular PageRank algorithm can be used to obtain a metric that reflects prestige. Analysis of journals according to their ISI IF and their weighted Page Rank shows significant overlaps and differences .[23]

The results list from a query with Page Rank will perhaps have a higher rate of precision than other page retrievalsystems, the pair argue that one link isn't necessarily equal to a vote in a website's favor--that same link could be reciprocal deals or even spam links on a page. the biggest weakness of Page Rank, though, is its ability to be manipulated. While manipulation of Page Rank may not be ethical.[15]

Article and Author Status – Normalised Citation Impact Index

The normalised citation impact index (NCII) calculates a score based on the average number of citations per article published, divided by the number of years that the article has been in print[1].

Author Status --- H-Index and M-Value

H-Index quantifies both the actual scientific productivity and the apparent

scientific impact of a scientist based on the set of the scientist's most quoted papers and the number of citations that they have received in other people's publications. The index can also be applied to study the productivity and impact of a group of scientists, such as a department, university or country. [23][36] The H-index and M-value were metrics developed by Jorge Hirsch (Hirsch, 2005) as a means of assessing the impact of individual scientists. . By starting with all papers published by an author it is possible to see if all of these have been cited at least as many times as the number of papers. If not, then the paper with the least number of citations is eliminated, and the process repeated until the number of articles remaining matches the number of citations for the least cited article. This number is the H-index.

M-Value is an extension of the h-index that takes into account the length of a scientist's career, and is simply the h-index divided b y the number of years between the author's first publication and their most recent. Researchers who maintain a consistently high output will have a higher m-value, as will researchers who have been working for a shorter time. This has the advantage of allowing the status of newer authors to be compared with more established figures. The m-value is the same regardless of whether either author is still actively publishing or ceased.[1]

Citation Network Analysis

The application of Topical Page Rank into citation network link analysis due to the fact that researchers may be experts in different scientific domains. A heterogeneous link analysis of the citation network, exploring the impact of weighting various factors. Another important approach, which is also our main focus in this paper, is via the social network analysis approach.graph-theoretic approaches in ranking network entities, researchers have begun to introduce link analysis approaches like PageRank[37][38] and HITS[39]into citation network analysis.

Topical Link Analysis in Citation Networks

(a) Topical Page Rank

The basic idea of Topical Page Rank is to incorporate a topic distribution into the representation of each webpage as well as the importance score of each page. Therefore, there are at least two vectors associated with each page: the content vector Cu : [C(u1),C(u2), ..., C(uT)], which is a probability distribution used to represent the content of u across T topics, and the authority vector, Au : [A(u1),A(u2), ...,A(uT)], which is used to measure the importance of the page, where A(uK) is the importance score on topic K.

ISSN 2249-3875

International Journal of Pharmaceutical Erudition

Topical Page Rank is also a random surfer model. Thus, the surfer's behavior can be modeled by a set of conditional probabilities:[40]

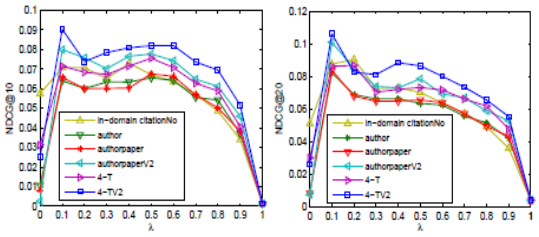

Figure 4: Comparison among different levels of citation network (NDCG@10 and NDCG@20 for ACM members) as the BM25 weight (λ) is varied.

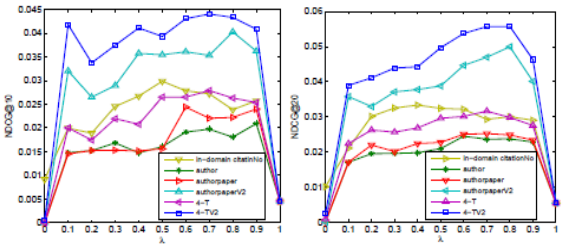

Figure 5: Comparison among different levels of citation network (NDCG@10 and NDCG@20 for PC members) as the BM25 weight (λ) is varied.

(b) Topical Citation Analysis

A topical link analysis approach into authors' reputation ranking. Similar to web pages, publication papers may also cover different topics. The same is true for authors' authority propagation. Believing in the prestige of a person on one aspect (eg. on data mining) does not mean that this person also owns a high reputation on other aspects (e.g., networking). When authors choose to collaborate and coauthor with each other, they may have mutual trust and interests on some certain aspect (topic). Publishing venues are normally more focused on certain research areas than others. SIGIR, has a high prestige in the information retrieval research field, the networking domain.

Exeperimental Results

Multi_Type Citation Network:

Figures 4-5 indicate the NDCG results for different kinds of citation network analysis approaches using original uniform Page Rank as propagation mechanism and using ACM members and PC members as evaluation dataset respectively.To reduce the amount of manual labelling, we only gave to judges the results when combined with BM25 with parameter _ set to 0.5. We a introduced the ranking method of in-domain citation count as one of our compared approach. We took the 23 categories provided by Libra as domains, and regarded it as in-domain citation if two papers are within onedomain and there is a citationship between them.

Topical Page Rank

A key issue in Topical Page Rank is to generate the static per-document content vector. We extracted the top level primary category and additional category for all the papers in the dataset and thus get 12 categories in total. We regard these categories as topics. Since each author is represented by a profile which is a concatenation of all the papers he has written, we can accumulate all the topics

ISSN 2249-3875

International Journal of Pharmaceutical Erudition

mentioned by each published paper, and then compute the topic distribution. The same mechanism works for computing the topic distribution of venues, for which we collected all the papers published in that venue, accumulated papers' topics, and then computed the corresponding distribution. The same thing for generating affiliations' topic distribution by collecting the papers writ-ten by authors from that affiliation, and taking use of the papers' topic distribution to compute the affiliations' topic distribution.

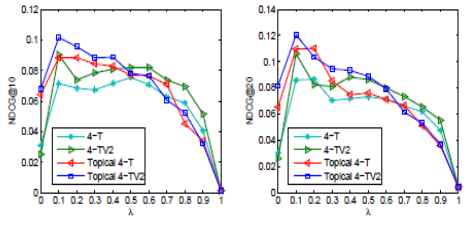

Figure 6: Topical PageRank Performance (NDCG@10 and NDCG@20 for ACM members) as the BM25 weight (λ) is varied.

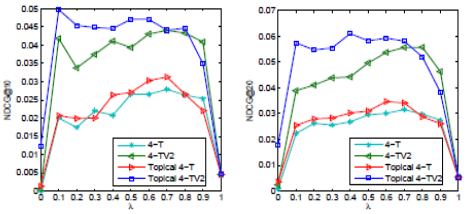

Figure 7: Topical PageRank Performance (NDCG@10 and NDCG@20 for PC members) as the BM25 weight (λ) is varied.

Figures 6-7 for the results of topical experiments. We set the _ value to be 0.85 in all experiments. We found once again a high consistency among the results from B different approaches , and that introducing Topical PageRank can improve the performance indeed. The improvement of the best performance of Topical 4-TV2 over that of 4-TV2 is 12.9%.

Comparison with two baselines

One of the main characteristics of our multi-type citation network analysis approach lies in its combination of both content-based approach and graph-based approach.

The CoRank algorithm generates author and paper rankings by taking propagation between authors and papers into account. It is a graph-based approach. Instead of building a big graph for all the authors and papers in the dataset, they first rank authors in terms of their topic weights in a certain domain, retrieve the top 500 authors, and build the graph based on these authors and their publications.

Heterogeneous Page Rank

A heterogenous PageRank algorithm with the intention of exploring on the different impact among social factors. The parameter _ij indicates the authority propagation probability among factors i and j. It is actually a parameter optimization problem if we want to get the best performance by tuning the parameters. [2]

Citation Network Visualization

The visualisations were created using a software application that the authors developed themselves, called Citespace'. Two dimensional graphs can be created to

ISSN 2249-3875

visualise a network, and a number of software tools exist that can do this, including a number specifically intended for use with citations, such as Pajek and CiteSpace. *CiteSpace* is Java application made freely available b y Chen (2006) that uses a variant of the SCI-Map algorithm to plot a graph of article co-citation in 2-d space.). The selection of article groups as clusters is based only on their positioning after the second step, and is independent of any other information, such as author, keywords, co-citations etc. [1] A novel integrated probabilistic model which combines a content-based approach with a multi-type citation network which integrates citations among papers, authors, affiliations and publishing venues in a single model. [2]

Critisism

Eugene Garfield himself addressed criticisms of impact factors while emphasizing their accuracy and significance. Responding to criticism that impact factors are based on a sampling of the literature, he reaffirms that a Journal Citation Report (including the impact factor) is based on every citation in every journal that it covers . [15]

Validity

The impact factor is highly discipline-dependent. The percentage of total citations occurring in the first two years after publication varies highly among disciplines from 1-3 percent .The impact factor could not be reproduced in an independent audit.

The impact factor refers to the average number of citations per paper, but this is not a normal distribution. It is rather a Bradford distribution, as predicted by theory. Being an arithmetic mean, the impact factor therefore is not a valid representation of this distribution and unfit for citation evaluation.

In the short term — especially in the case of low-impact-factor journals — many of the citations to a certain article are made in papers written by the author(s) of the original article. This means that counting citations may be independent of the real "impact" of the work among investigators. Garfield, however, maintains that this phenomenon hardly influences a journal's impact factor. Journal self-citation is common in journals dealing in specialized topics having high overlap in readership and authors, and is not necessarily a sign of low quality or manipulation. [26]

Citation Data Quality

The radius of a node is made proportional to the cube root of its citation frequency, thus relating the number of citations to the volume of a sphere. In Citespace the degree of proportionality is configurable, so that nodes can easily be made larger or

smaller in the final visualisation. CiteSpace can minimise complexity by using a method called ' *Pathfinder Network Scaling'* to eliminate the weakest links and form a simplified representation. The selection criterion to retain a link is that the weight of a path that connects two articles should not be greater than the weight of any alternative paths.

•Chronological Tree Ring Visualisation - Makes clear in the graph not just how popular an article was, but when it was popular.

•Burst-analysis - Identifies the sudden appearan ce of frequent references to novel concepts or phrases. [1]

ISI produces two tools for measuring impact factor: journal citation reports, which compares and ranks journal with similar journal covering a subject discipline; and University Science Indicators, which compares and ranks universities in various ways, including by field/subject discipline. [13]

There are two factors used in calculation, which uses data from the previous two years.

First, the number of cited articles for a journal are tallied for a given year. Second, the number of articles published in this same journal are figured for that year. Data from two years is added together, then the number of cited articles are divided by the number of published articles.

Cited Articles per Year:	Published Articles per Year:
2007 = 75	2007 = 48
2006 = 80	2006 = 48
Total: 155	Total: 96

A 2008 impact factor would look like this for a hypothetical journal.[15]

Merits and Limitations

1. The use of impact factor as an index of journal quality relies on the theory that citation frequency accurately measures a journal's importance to its end users.

2. It provides quantitative tools for ranking, evaluating, categorizing and comparing journals.

3. It eliminates some of the bias of such counts which favor large journals over small ones or frequently issued journals over less frequently issued ones and of older journals over newer ones. Particularly in the latter case such journals have a larger citable body of literature than smaller or younger journals.

4.There have been many innovative applications of journal impact factors. The most common involve market research for

publishers and others.

5.It provides librarians and researchers a tool for the management of library journal collections.

6.In market research, the impact factor provides quantitative evidence for editors and publishers for positioning their journals in relation to the competition—especially others in the same subject category.

7. It may also serve to advertisers interested in evaluating the potential of a specific journal.

8. Perhaps the most important and recent use of impact is in the process of academic evaluation.

9. The impact factor can be used to provide a gross approximation of the prestige of journals in which individuals have been published.

It has aims to create awareness about impact factor of a journal .Is it a metric of popularity or prestige is debatable, particularly in Indian contest. It has many merits and demerits which we think need considerations by Indian authors and time has come they should start submitting their valuable/best research to Indian journals particularly senior researchers, as young scientist/researchers have yet to achieve academic heights where as senior scientist have achieved such heights and it will not make much difference for them but will surly enhance impact of Indian journals.

Limitations

1. Review articles generally are cited more frequently than typical research articles because they often serve as surrogates for earlier literature.

2. It is widely believed that method articles attract more citations than other types of articles.

3. The practice of self-citation can be considered at many levels, including author self-citation, journal self-citation, and subject category selfcitation. This may increase the impact factor.

4. A title change affects the impact factor for two years after the change is made.

5. Different specialties exhibit different ranges of peak impact.

6. It does not distinguish between letters, reviews, or original research.

7. It has inadequate international coverage.

8. The coverage is very uneven.

9. Very few publications from languages other than English are included, and very few journals from the less-developed countries.

10. The number of citations to papers in a particular journal does not really directly measure the true quality of a journal much less the scientific merit of the papers within it.

11. It only reflects the intensity of publication or citation in that area and the

current popularity of that particular topic, along with the availability of particular journals.

12. Journals with low circulation, regardless of the scientific merit of their contents, will never obtain high impact factors in an absolute sense

13. By merely counting the frequency of citations per article and disregarding the prestige of the citing journals, the impact factor becomes merely a metric of popularity, not of prestige.

14.. A journal can adopt editorial policies that increase its impact factor. These editorial policies may not solely involve s improving the quality of published scientific work. Journals sometimes may publish a larger percentage of review articles. While many research articles remain uncited after 3 years, nearly all review articles receive at least one citation within three years of publication, therefore review articles can raise the impact factor of the journal.

15. Editorials in a journal do not count as publications. However, when published articles, often articles are cited, often from the same journal, those citations increase the citation count for the article.

16. An editor of a journal may encourage authors to cite articles from that journal in the papers they submit.

17. Many authors are biased to submit there research papers just on the basis of impact factor of a journal ignoring national journals and submitting only those research articles which have been rejected. This may worsen the situation for any local journals. [28]

Conclusions

The assumption is that different citation relationship can mutually reinforce each other, and thus can improve ranking performance. investigated the value of integrating author and paper information in citation network in ranking authors' reputations. In order to better evaluate the prestige of an author in different kinds of research topics, we incorporated topical link analysis into the citation network. We conclude from experimental results that:

• Multi-type citation networks can effectively improve ranking performance. Affiliation and publishing venues provide additional useful information in evaluating authors' reputations.

• Topical link analysis shows positive improvement in ranking authors' authority.

•Heterogeneous PageRank, with parameter tuning, can work even better than Topical PageRank.[2]

Quantitative measures of scientific output have therefore been sought as a more accessible alternative. Publication counts alone do not inform about the quality of scientific output. This has led to the

practice of evaluation by citation rates and impact factors. [9]

Impact factor does not reveal anything about the quality of the research presented in a journal, studies have been conducted that compare impact factors with perceptions of journal quality. [15]

By merely counting the frequency of citations per article and disregarding the prestige of the citing journals, the impact factor becomes merely a metric of 'popularity', not of 'prestige'[34]. IF and citations get weightage in recruitments, promotions, rewards and other recognitions. it is tough to define the quality of an academic publication using only non-quantifiable factors, such as its potential influence on the next generation of scientists. But citation-based data should at best be used to supplement peer judgement. It is about time that we seriously look for measures beyond the IF.As a starter, studies could be commissioned on the limitations of IF to come out with India-specific measures to evaluate science using open data sources like the Google Scholar. [23]

Highly criticized, especially when used to evaluate an individual investigator's scientific production (author impact), the IF has become the most widely used metric for assessing the overall quality of a biomedical journals internationally. In addition to the quality of material published in a journal, some factors that can influence IF include the frequency of publication, popularity among peer groups, number and nature of article published, circulation of the journal among international libraries and hence availability to different authors.[31]

References

1.http://www.lulu.com/items/volume_68/9010000/9010885/print/901885.pdf

2.http://www.cse.lehigh.edu/~brain/pubs/2010/RAIO/citation_network_analysis.pdf

3. Balog K., Azzopardi L., and de Rijke M. Formal models for expert finding in enterprise corpora. In Proc. of the 29th Annual Int'l ACM SIGIR Conf. on Research and Development on Information Retrieval, 2006, 43–50.

4. Fang H. and Zhai. C. Probabilistic models for expert finding. In ECIR, 2007, 418-430.

5. Macdonald C. and Ounis. I. Voting for candidates: adapting data fusion techniques for an expert search task. In Proc. of the 15th ACM Int'l Conf. on Information and knowledge management, 2006, 387-396.

6."Oxford Referencing System". http://www.lc.unsw.edu.au/onlib/refbib.html. Retrieved 18 January 2011.

7.http:// www. en. wikipedia. org/ wiki / citation.

8.http://www.en.wikipedia.org/wiki/scientific_citation.

9.Whitehouse GH *British Journal of Radiology*.2001;74, 1-3.

10. Seglen P. O. The skewness of science. *J Am Soc Information Sci* 1992, 43,628-38.

11 Kurtz M. J., Eichhorn G., Accomazzi A., Grant C. S., Demleitner M. and Murray S. S. The Effect of Use and Access on Citations, *Information Processing and Management* , 2004, 41 (6), 1395–1402,

12. Hoang D., Kaur J. and Menczer F. Crowdsourcing Scholarly Data, Proceedings of the WebSci10: Extending the Frontiers of Society On-Line, April 26-27th, 2010, 321

13.http://www.library.yale.edu/science/help/sciexpl.html.

14.http://www.en.wikipedia.org/wiki/citation_impact

15.http://fables&formulas.wetpaint.com/page/citation+index+impact+factor+and+page+rank.

16. Open access, readership, citations: a randomized controlled trial of scientific journal FASEB J. 201125 n7:2129-2134

17. Long Island University.

18. Duke University Libraries 2007.

19. Brigham Young University 2008.

20. University of Maryland, College Park 2006

21. How to Write Research Papers with Citations - MLA, APA, Footnotes, Endnotes.http://www.studenthandouts.com/citations.htm. Retrieved 2010, 01-31.

22. Kouchupillai M. and Smith M. A, Patent Valuation with Consideration for Emerging Technologies *Journal of intectual property righte*, 2007, 12, 154-164

23.Satayanaryana K. and Sharma A. Impact factor: time to move on. *Indian J Med. Res.* 2008, 127, 4-6

24. Garfield E. Citation analysis as a tool in *journal evaluation*. *Science,* 1972 178(60), 471–479.

25. Garfield E. Citation indexing: Its theory and application in science, technology, and humanities. John Wiley and Sons, Inc., New York, NY, USA, 1979.

26.http://www.en.wikipedia.org/wiki/impact_factor

27.http://www.en.wikipedia.org/wiki/impact_factor

28.http://www.jkscience.org/archive/volume93/journal%20knowledge/impact

29.Garfield Eugene,*CMAJ* 1999, 161(8)

30.Satyanarayana K. and jain N. C. The IF of IJMR is on the rise. *Indian J. Med. Res.* 2006, 123, 717-718

31 Mohopatra P.R. and Mishra B..Impact factor of the IJMR. *Indian J Med Res* 2008, 127, 198-199.

32. Garfield E. Citation analysis as a tool

in journal evaluation. *Science* 1972., 178, 471-9,

33.http://www.thomsonreuters.com/products_service/science/free/essays/impactfactor

34. Bollen Rodrigues M. A. and Van de Sompel H. Journal Status. *Scientometrics* 2006, 601.

35.www.wame.org/wame_listserve_discussions/list_Impact_factor&the_y_factor.

36. Hirsch J. E. An index to quantify an individual's scientific research output. *PNAS* 2005, *102,* 16569-72

37. Haveliwala T. H. Topic- sensitive PageRank. In in WWW'02: Proc. of the 11th Int'l Conf. on] World Wide Web, 2002,517–526.

38 L. Page. S. Brin, R. Motwani. and T. Winograd The.PageRank citation ranking: Bringing order to the Web. In Stanford InfoLab, Technical Report 1999, 66.

39. J. M. Kleinberg. Authoritative sources in a hyperlinked environment. *Journal of ACM*, 1999, 46(5), 604–632,

40. L. Nie, B. D. Davison, and X. Qi. Topical link analysis for web search. In Proc. of the 29th Annual Int'l ACM SIGIR Conf. on Research and Development in Information Retrieval, 2009, 91-98.

41. Pinski G. and Narin F. Citation influence for journal aggregates of scientific publications: Theory with applications to literature of physics. Inf. Proc. 1976.