Exploring metrics of analyzing citation networks

Invited Paper

Suhas Gundimeda
Indian Institute of Technology Madras

Karnaty Rohit Reddy
Indian Institute of Technology Madras

Abstract—We describe some measures and metrics proposed by various papers in scientometrics to analyze citation and co-citation networks.

I. INTRODUCTION

Scientometrics deals with the measuring and analysing science, technology and innovation. In this paper we look into analysis of scientific contributions in peer-reviewed publications using their citation, co-citation and co-authorship networks.

II. PREVIOUS WORK

A. Citation count

Citation count is the sum of all published citations to a scientific paper. This is the most widely used measure of impact, and is the standard metric against which all of the others are compared.

B. PaperRank

The PageRank algorithm has the problem of nodes with no inbound edges causing it to have no solution. To solve this problem the algorithm is modified to assign rank of 1 to all initial nodes and reset it to zero after computation.[2]

C. h-index

An author metric whose value is the maximal number h such that the author has at least h papers with at least h citations.

D. PaperRank-Hirsch index

This is a proposed alternative to h-index which uses the calculated PaperRanks. The PaperRanks are normalized to represent more meaningful quantities, namely the most probable number of citations which the node might receive, the Q-value. The index is then defined as the integer number h such that an author has at least h papers with Q-value greater than or equal to h.

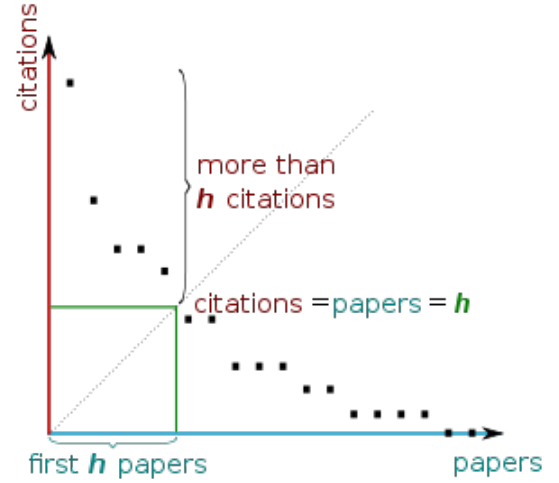### E. CoRank

1) CoRank-LinkCount:
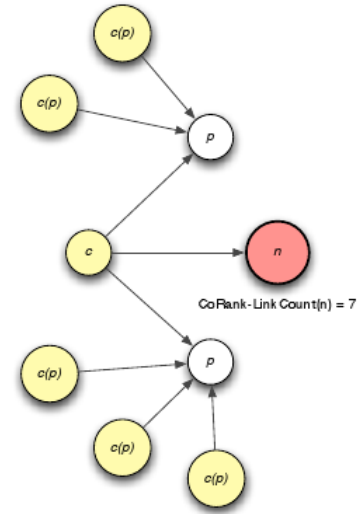

Fig. 1. h-index illustrated

Fig. 2. Network of Co-Cited publications from publication n[1]

F. Betweeness centrality

G. Closeness centrality on co-authorship networks

Closeness centrality (Freeman, 1979) emphasizes the distance of a vertex to all others in the network through the geodesic distance from each vertex to all others. It is a measure

of how long it will take for information to spread from a particular node (Yin et al., 2006).

$$C_c(n_t) = \sum_{i=1}^{N} \frac{1}{d(n_i, n_j)}$$

Where $d(n_i, n_j)$ is the distance from node i to node j and C_c is the closeness centrality.

REFERENCES

[1] DAVID C TARRANT (2011) "A Study of Early Indication Citation Metrics ", University of Southampton, Electronics and Computer Science, PhD Thesis.
[2] Mikalai Krapivin, Maurizio Marchese and Fabio Casati "Exploring and Understanding Scientific Metrics in Citation Networks"
[3] Erjia Yan, Ying Ding, "Applying centrality measures to impact analysis: A coauthorship network analysis"

bla bla