Exploring and Understanding Scientific Metrics in Citation Networks

Mikalai Krapivin, Maurizio Marchese, and Fabio Casati

Department of Information Engineering and Computer Science,
University of Trento, Italy
{krapivin,marchese,casati}@disi.unitn.it

Abstract. This paper explores scientific metrics in citation networks in scientific communities, how they differ in ranking papers and authors, and why. In particular we focus on network effects in scientific metrics and explore their meaning and impact. We initially take as example three main metrics that we believe significant; the standard citation count, the more and more popular h-index, and a variation we propose of PageRank applied to papers (called PaperRank) that is appealing as it mirrors proven and successful algorithms for ranking web pages and captures relevant information present in the whole citation network. As part of analyzing them, we develop generally applicable techniques and metrics for qualitatively and quantitatively analyzing such network-based indexes that evaluate content and people, as well as for understanding the causes of their different behaviors. We put the techniques at work on a dataset of over 260K ACM papers, and discovered that the difference in ranking results is indeed very significant (even when restricting to citation-based indexes), with half of the top-ranked papers differing in a typical 20-element long search result page for papers on a given topic, and with the top researcher being ranked differently over half of the times in an average job posting with 100 applicants.

Keywords: Scientific metrics, Scientometric, Page Rank Algorithm, Paper Rank, H-index, Divergence metric in ranking results.

1 Introduction

The area of scientific metrics in scholarly social networks (metrics that assess the quality and quantity of scientific productions) is a relevant area of research [1,2,3] aiming at the following two objectives: 1) measuring scientific papers, so that "good" papers can be identified by the scientific community and so that researchers can quickly find useful contributions when studying a given field, as opposed to browsing a sea of papers, and 2) measuring individual contributions and related reputation, to determine the impact of a scientist and to help screen and identify candidates for hiring and promotions in industry and academia.

Until only 30 years ago, the number of researchers and of conferences was relatively small, and it was relatively easy to assess papers and people by looking at

J. Zhou (Ed.): Complex 2009, Part II, LNICST 5, pp. 1550 – 1563, 2009.

papers published in international journals. With small numbers, the evaluation was essentially based on looking at the paper themselves. In terms of quantitative and measurable indexes, the number of publication was the key metric (if used at all). With the explosion of the number of researchers, journals, and conferences, the "number of publications" metric progressively lost meaning. On the other hand, this same explosion increased the need for quantitative metrics at least to "filter the noise". For example, a detailed, individual, qualitative analysis of hundreds of applications typically received today for any job postings becomes hard without quantitative measures for at least a significant preliminary filtering.

Recently, the availability of online databases and Web crawling made it possible to introduce and compute indexes based on the number of citations of papers (citation count, from hereafter denoted as CC) and its variations or aggregations, such as the impact factor and the h and g indexes [4]) to understand the impact of papers and scientists on the scientific community. More and more, Universities (including ours) are using these indexes as a way to filter or even decide how to fill positions by "plotting" candidates on charts based on several such indexes.

Besides "traditional" metrics, novel metrics for papers and authors are being proposed [5]: they are inspired at how the significance of Web pages is computed (essentially by considering papers as web pages, citations as links, and applying a variation of the PageRank algorithm [5,6]). PageRank-based metrics are emerging as important complement to citation counts as they incorporate important information present in the whole citation network, namely the "weight" (the reputation or authority) of the citing paper and its density of citations (how many other papers it references) in the metric. From a computational point of view, PageRank is a statistical algorithm: it uses a relatively simple model of "Random Surfer" [6] to determine the probability to visit a particular web page. Since random browsing through a graph is a stochastic Markov process, the model is fully described by Markov chain stochastic matrix. The most intriguing question about PageRank is how to compute one for very large networks such the web. The inventors of PageRank, Brin and Page, proposed a quite effective polynomial convergence method [6], similar to the Jacobi methods. Since then, a significant amount of research has been done in the exploration of the meaning of PageRank and proposals for different computation procedures [5,7,8,9]. When the attention is shifted from web pages to scientific citations, the properties of the citation graph – mainly its sparseness – has been used to simplify the computational problem [10]. Some recent work has also started to analyze and compare the effectiveness of the different ranking algorithms [11].

In the present work, we have based our computations on a variation of Page Rank (called PaperRank) for ranking scholarly documents. In particular, this paper performs an experimental study of scientific metrics based on citation networks with the goal of (1) assessing the extent of differences and variations on the evaluation results when choosing a certain metric over another, and (2) exploring and understanding the reasons behind these differences. We performed the analysis on a dataset consisting of over 260K ACM (Association for Computing Machinery) publications, a social scientific network of 244K authors and ca. one million citations. The results of the analysis are rather significant, in that even if we restrict to citation-based indexes, the choice of the specific index rather than another changes the result of filtering and selection of papers and evaluation of people about half of the times.