INSTITUTO DE COMPUTAÇÃO
UNIVERSIDADE ESTADUAL DE CAMPINAS

Invisible Work in Standard Bibliometric Evaluation of Computer Science

C. Billa *J. Wainer* *S. Goldenstein*

Technical Report - IC-09-20 - Relatório Técnico

June - 2009 - Junho

Invisible Work in Standard Bibliometric Evaluation of Computer Science

Cleo Billa Jacques Wainer Siome Goldenstein

Abstract

Science is a competitive business, and researchers from different fields are constantly compared in terms of funding and for promotion evaluations. This article analyzes the quantity of computer scientists' work which is not duly recognized when using Web of Science and/or Scopus bibliometrics. We randomly selected CS faculty from highly-ranked US Computer Science departments and determined that, on average, 67% of their published work is not accounted for within Web of Science searches and 44% is absent when searching with Scopus. We defined this parameter as the invisible work rate. We compared these figures to similar samples from Mathematics, Physics, and Electrical Engineering faculty. While CS and EE have significantly distinct publishing patterns as compared to Mathematics and Physics, the variance of the invisible work rate for CS is high, which suggests that different subareas of CS also have different publication practices – a possible indication that it will be difficult for computer scientists to agree on one bibliometric evaluation criterion.

Keywords: research evaluation, bibliometrics, bibliographic repositories

1 Introduction

Everyday multidisciplinary committees make strategic decisions, rule about subjects ranging from faculty promotion to grant awarding, and rank and compare scientists from different areas. Although these committees may use different criteria to evaluate researchers in subjects as disparate as History and Medicine, it seems logical to group the subjects of Mathematics, Computer Science (CS) and Electrical Engineering (EE) together for these committees' evaluations.

These comparative evaluations become more frequent as the number of scientists and papers grow. Since funding sources do not grow as fast, and research practices vary among different subjects, using the same criteria in different areas may bring about large injustices. We believe that the recent discussion in the Communications of the ACM about CS research evaluation [4, 6] builds the case for the CS community to defend itself from these expected injustices in future comparative evaluations.

The most traditional assessment criteria are based on Thomson-Reuters' Web of Science (WoS) indexing service, which quantifies not only the production of, but also the number of citations by individual scientists, university departments (for example [5]), whole universities [2]), countries [7], and different scientific areas.

We, as computer scientists, have an intuitive understanding that these assessment criteria are unfair to the area as a whole. The goal of this paper is to provide some quantitative evidence of such unfairness. In this paper, we do not question or justify computer science's research culture of publishing articles in conferences prior to (or instead of) publishing them in journals, but we show

how and why these evaluation metrics are inappropriate to compare computer science researchers to researchers in other closely related fields, such as Mathematics and Electrical Engineering.

We define a researcher's *invisible work* as all of his or her scientific publications that are not indexed by WoS or by Scopus. Thus, this work is not counted as part of the scientist's standard bibliometric evaluation.

In this paper we compare CS's invisible work to Physics, Mathematics, and Electrical Engineering, as similar scientific areas. Physics is typically an area whose publication practices are in accordance with the standard bibliometric assessment. Mathematics is closer to CS as a subject than Physics, but should have less invisible work than CS because they publish less in conferences. Electrical Engineering should have publication practices similar to CS based on it's work in conferences. In our comparison, we focus on publication count - that is, the number of papers produced by a scientist.

In order to evaluate the invisible work of Physics, Math, EE, and CS, we generated a controlled sample of 50 scientists from each of these fields from top US universities, collected their "true" production (the publications the scientists themselves list in their personal web pages and/or publicly available *curriculum vitae*), and counted their production in WoS and Scopus. The proportion of the unindexed production to the true production is their invisible work.

1.1 WoS and Scopus

Thomson-Reuters' Web of Science[1] is the most traditional citation indexing service, and it is standard in most bibliometric research. Regarding computer science, WoS divides the field into six subfields: Artificial Intelligence (with 105 journals indexed), Cybernetics (19 journals), Hardware and Architecture (48), Information Systems (114), Interdisciplinary Applications (96), and Software Engineering (91). However, the same journal can be classified into different categories.

The CS community generally identifies at least three problems regarding WoS.

- WoS does not include conferences other than the ones published in the Springer's Lecture Notes in Computer Science (LNCS) series.
- WoS does not include journals considered important in the field.
- WoS includes some journals that most of the community does not consider to be CS journals.

Scopus[2] is a recent competitor to WoS, developed by Elsevier. Scopus not only includes in it's set more CS journals (for example all of the ones published by ACM and Elsevier), but also includes some conferences in it's indexing set. At the time of this writing, Scopus indexed 849 CS journals, and 162 conferences[3]. The CS community generally feels that Scopus better represents the venues where the field publishes it's scientific works.

There have been some articles that compare the differences in coverage and the resulting differences in citation counts when one or another service is used [1, 3], but these works are not specific to the CS area.

[1] `http://www.isiknowledge.com`
[2] `http://www.scopus.com`
[3] `http://www.scopus.com/scopus/source/browse.url?subjectArea=1700&sourceType=&entitlementType=&x=17&y=8`

2 Methodology

2.1 Selection of scientists and their "true" production

Based on a ranking of the best graduate schools in the US as determined by *USNews and World Report*[4], we selected those Universities ranked between 6th and 10th from Physics, Math, CS (all taken from the Sciences section), and Electrical, Electronic, and Communications Engineering (from the Engineering section). The rational for avoiding the top five departments is that they could have specific features in their research culture that place them as outliers amongst the rest of the area. We believe that well-ranked departments probably have a "correct" or a "good" balance between the different subareas within Physics, Math, EE, and CS, and they are, hopefully, representative of the balance of these subareas in the general research population. Table 1 shows the universities we used in the sampling.

Computer Science	**Mathematics**	**Electrical Engineering**	**Physics**
Cornell	U. of Chicago	Georgia Tech	Cornell
Princeton	Cal Tech	U. of Michigan–Ann Arbor	U. of Chicago
U. of Washington	Yale	Carnegie Mellon	U. of Illinois–Urbana-Champaign
Georgia Tech	Columbia	Cornell	UC–Santa Barbara
U. of Texas–Austin	New York U.	Purdue	Yale

Table 1: Selected Universities

We randomly ordered the faculty of each department and selected the first 10 scientists with a complete and up-to-date "Publications" section in their personal web pages or publicly available curriculum vitae, to a total of 200 scientists, 50 from each area. We did not consider researchers whose pages refer to "Selected Publications", and we did not consider as up-to-date web pages without publications listed for 2008.

For those 50 researchers, we counted the number of publications from 2000 to 2007, inclusive. This time window should capture the current publication practices (which may have been different in previous decades). This data was collected in March and April 2009.

Many of the Physics and Math scientists do not maintain up-to-date personal web pages or public CVs. Therefore, it was not always possible to select 10 researchers from each university in table 1. In these cases we selected more researchers from the other universities, or in the case of Physics, we had to select researchers from the 6th to the 12th top-ranked universities.

Twenty of the CS researchers and 32 of the EE researchers divided their production into journal and conference papers. For these we recorded the two partial totals as well.

2.2 WoS data

To collect the WoS data we used the basic search form using the last name of the scientist and his/her initials. After a primary search, WoS retrieved a list of papers. Another option: the `View Distinct Author Sets`, allowed us to view a list of authors that matched the search. Using this tool, we manually selected the set of authors that matched the respective scientist. From the new list of papers, we used a refine tool to select only the papers published between 2000 and 2007.

[4] Available in `http://www.usnews.com/sections/rankings`.

In `Document Type`, WoS shows how many documents are (journal) articles, proceedings papers, meeting abstracts, notes, reviews, editorials, corrections, letters, and others. We recorded only articles and proceedings papers and considered the total researcher production the sum of both figures.

2.3 Scopus data

Scopus provides an advanced search that allows the use of boolean operators and the selection of an author set. We built a query with `pubyear aft 1999 and pubyear bef 2008` and the `Author Name` for each scientist.

In the `Select Author` window, we used the scientist's last name and his or her complete first name. If more than one author was retrieved, we manually selected set of "author items" that corresponded to the particular scientist. We used the subject area and affiliation to support these decisions.

Scopus also distinguishes documents by type. Again, the only types of documents recorded were (journal) articles and conference papers, and total production was defined as the sum of these two numbers.

Regarding CS, Scopus is problematic in it's classification of LNCS as both a conference and a journal publication. The issues of LNCS are mainly proceedings of different conferences, and most CS researchers would classify papers published in LNCS as conference papers. For Scopus that same paper is counted twice, as a conference and a journal paper, which would generally increase both the total and the journal publications count of CS researchers. Thus, to obtain CS researcher journal article production, we excluded all papers published in LNCS.

2.4 Metric

We define the invisible work rate for a researcher as:

$$\text{invisible work rate} = \frac{\text{true number of papers} - \text{number of papers in indexing service}}{\text{true number of papers}}$$

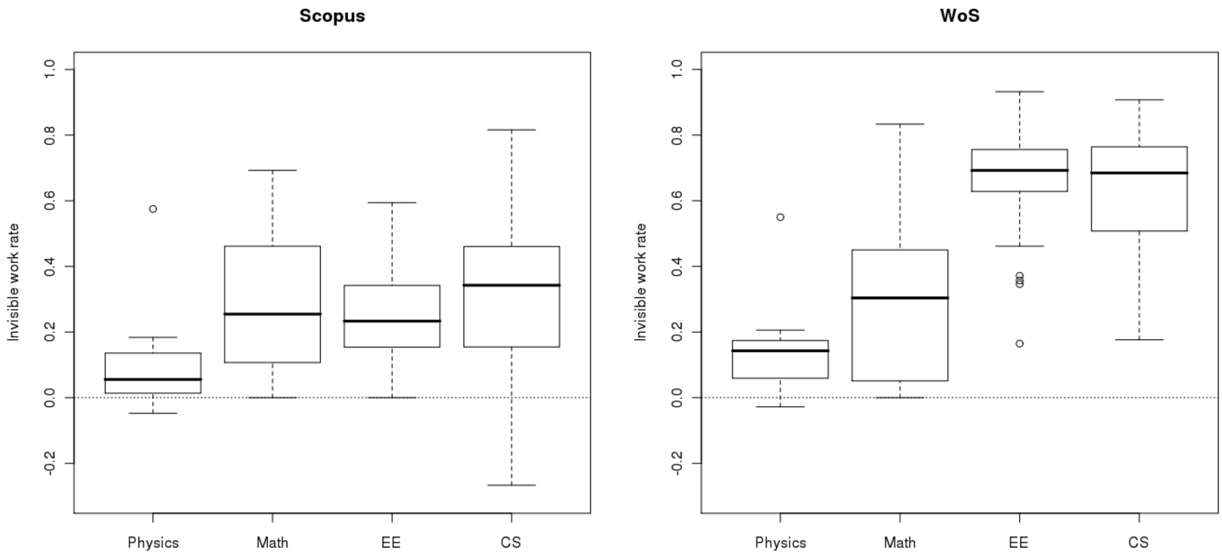
and we analyze the distribution of invisible work rates for the different scientific areas.

An invisible work rate of 1 means that all of the scientist's publications are not accounted for in the indexing service; a rate of 0 means that all of his/her publications are represented in the indexing service.

It is possible to obtain a negative rate if there are more works in the indexing service than the ones declared by the scientist. There are three main reasons for a negative rate. Our search strategy may have resulted in more than one researcher being selected and all their indexed work would be attributed to a single researcher. There were a lot of cases where our query returned more than one author (specially in WoS where only the initials are used to disambiguate authors with the same last name). Although we exercised care in these cases, we cannot be sure that there were no examples of two or more authors being counted together.

Negative rates may also occur if researchers do not list all of their work in their public list of publications, such as poster or conference papers presented by students. Also, some researchers may not frequently update their web pages and CVs.

Finally, negative rates may indicate that the indexing service incorrectly attributes papers or wrongly classifies papers. This is more relevant when we distinguish the invisible work rate for



(a) Scopus (b) WoS

Figure 1: Distribution of Invisible Work

conferences and journal articles. WoS in particular has some errors regarding the classification of a work as conference or article.

3 Results

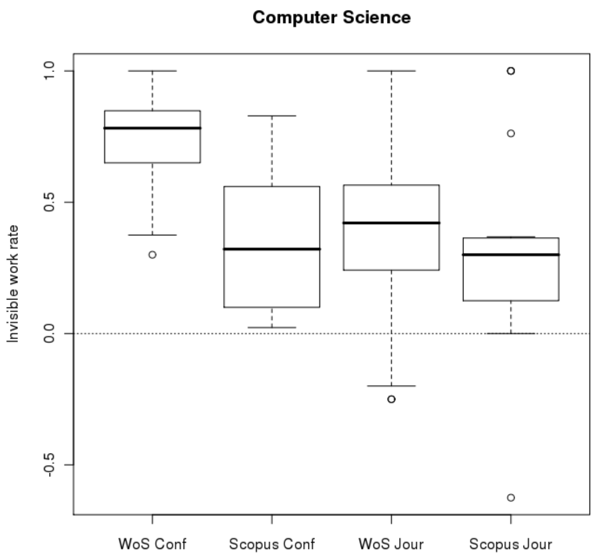
The boxplots in Figure 1, show the distribution of invisible work rate for each area.

If we consider WoS, on average 67% (standard deviation = 22%) of the CS's work is not represented in the indexing service, whereas for EE this rate is 67% (sd = 14%), for Math, 29% (sd = 24%), and for Physics 20% (sd = 23%). Using Scopus, the rate improves for CS: only 44 % (sd = 27%)) of the work is not represented, as compared to 25% (sd = 13%) for EE, 30% (sd = 22%) for math, and 15% (sd = 22%) for Physics.

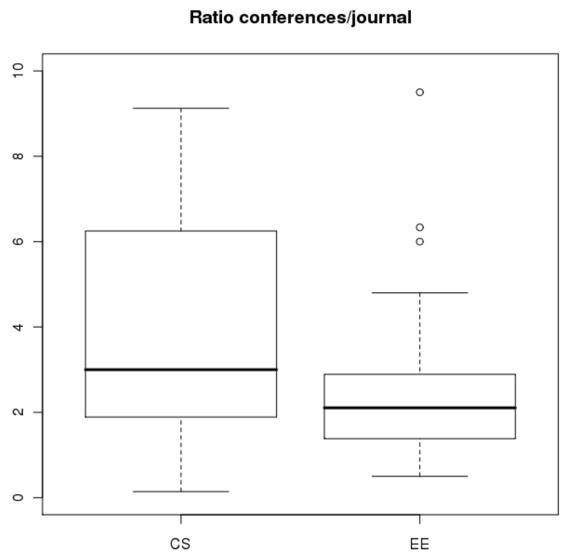
As expected, less of the Physics and Math researchers' work is unaccounted for using both WoS and Scopus. The invisible work rates for CS and EE are surprisingly similar within WoS. Using Scopus, there is an improvement in the invisible work rate for both CS and EE, but not uniformly.

An important characteristic of the CS distribution is the large standard deviation. There are CS researchers for whom most of their work is represented in WoS or Scopus, and there are some for whom most of their work is not represented in these services. We speculate that these differences are due to different publication practices in the many subareas of CS. Other areas, including EE, seem more homogeneous in their publication practices.

These differences in publication practices in the subareas of CS may lead to tension and injustices in bibliometric evaluations between CS researchers. Researchers in subareas which publish more in conferences may feel that both WoS and Scopus bibliometric evaluations are unfair.



(a) CS invisible work separated by venue.



(b) Conference/journal ratio (see text for outlier not shown).

Figure 2: Conference and journal

3.1 Conferences versus journals

An expected reason for the under-representation of CS work in both WoS and, less intensely, in Scopus is the volume of CS work that is published in conferences. Conferences usually are not indexed in WoS and only a few are indexed in Scopus. This could explain the high invisible work rate in CS (and EE).

Using the 20 CS researchers that distinguished between journal and conference papers, we calculated the invisible work rate for conferences and journals separately. Figure 2.a shows the distribution of invisible work for conferences and journals.

The CS invisible work rate for journals is 43% (sd = 39%) for WoS and 39% (sd = 36%) for Scopus; it is 77% (sd = 16%) and 45% (sd = 28%) respectively, for conferences.

Notice that the invisible work rate for journals ranges from one to a negative number. That is, there are researchers for whom none of their journal publications are accounted for by the bibliographic service, and there are researchers for whom there are more journal papers in the service than in their CVs. Again, this may be due to the different practices of the subareas of CS, but it also shows how much using such a service can be unfair.

If we compare the ratio of conference papers to journal papers of the 20 CS researchers and the 32 EE researchers that separately listed their publications in their personal web pages or CVs, again the inhomogeneity of the CS area is clear. Figure 2.b shows both distributions, but there is an outlier (more than 20 conference/journal) for CS not shown in the graph. Different researchers in CS publish differently in journals and conferences. Some do not publish in conferences at all, while one published more than 20 conference papers for each journal paper. The average rate for CS is 4.7 (sd = 5.9) whereas for EE, the average is 2.6 (sd = 1.8). The differences in the standard deviations are statistically significant by the Levene test based on means (p-value=0.01).

4 Conclusions and Discussion

When CS is classified as a science (as *US News and World Report* does), the standard bibliometric evaluations are unfair to CS as a whole. On average 67% of the published work of a computer scientist is not accounted for in the WoS bibliographic service, a rate much higher than that for Math and Physics scientists. In this case, a computer scientist should lobby for the use of Scopus as the bibliometric service used.

CS has similar mean and median values of invisible work rate as EE, but there is an interesting difference. The standard deviation of the invisible work rate in CS is statistically significantly larger (Levene test, WoS p-value=0.03, Scopus p-value=0.001). That is, for different members of the CS community there are large variations on how well their work is represented in the indexing services. We speculate the the different publication practices in different subareas of CS are the reason for the differences. It would be interesting to study the different rates for subareas of CS.

The practice of CS researchers in publishing their work in conferences is an important component of the invisible work rate of the area. On average, 77% of the conference publications are not indexed in WoS and 45% are not in Scopus, and computer scientists publish on average 4.7 papers in conferences for each journal paper. This rate is not only higher than that for EE, but it also has a large standard deviation, again indicating the inhomogeneity of the CS area.

These high values in the distribution variances, which we hypothesize come from the field's inhomogeneity, are a clear indication that it will be difficult for computer scientists to agree on one bibliometric evaluation criteria.

4.1 Limits on this research

The main source of error in this research is the retrieval of WoS and Scopus data. With at least half of the researchers, there was a decision on whether to join different "author items" as representing a single researcher, or to break a single author item in order to refine the entries for the researcher. In this case, we used information such as affiliations and subject area to make the decision.

Only one of this paper's authors collected the 200 researchers' data from WoS and Scopus. Thus, there is consistency in this data collection, but there can be some systematic errors. In order to evaluate the error of this procedure, another of this paper's authors independently collected the data for the 50 CS researchers again. Between these two evaluations, while there were two cases in which the difference between the total number of publications for a researcher was larger than 10 papers, the average difference for Scopus was 1.6 publications, and for WoS it was 1.8 publications. In terms of relative error, it averaged 4.5% of the researcher's publications for Scopus, and 5.0% for WoS. Hence, different forms of querying the bibliographic services resulted in a 5% difference in the number of publications attributed to the researcher.

The strength of the conclusions of this paper are dependent on our sample's representation of the selected scientific fields communities, and the CS community, in particular. There are sources of potential bias in our samples.

Most obviously, all researchers in the sample work in the US, therefore some of the conclusions may not be valid for other countries. In countries other than the US, the reward and evaluation structures may encourage researchers to concentrate their efforts in publishing their results in journals or in conferences, and thus both the conference to journal ratio and the invisible work rate may be different.

As we discussed, we selected the 6th to 10th ranked departments in an effort to avoid the bias inherent in the evaluation criteria used by *USNews and World Report*. However, it may be that the 6th to 10th ranked schools also carried some of the same bias. On the other hand, a more highly-ranked department should have enough faculty to achieve a balance between the different subareas of CS.

These highly-ranked departments may have a bias related to the distribution of tenured versus untenured faculty, or more experienced researchers versus younger academics. More highly-ranked departments probably have a higher percentage of experienced, tenured faculty, and these researchers probably distribute their publishing efforts differently than younger, untenured researchers.

Finally, the last source of bias is not relevant to the CS community, but may have been important in the analysis of the other areas. Our sample was not a random selection of the faculty in the department, but a random selection of faculty that maintains an up-to-date personal web page. In EE and CS, almost all faculty does this, but that was not true in Math and Physics.

References

[1] J. Bar-Ilan. Which h-index? a comparison of WoS, Scopus and Google Scholar. *Scientometrics*, 74(2):257–271, 2008.

[2] R. Ma, C. Ni, and J. Qiu. Scientific research competitiveness of world universities in computer science. *Scientometrics*, 76(2):245–260, 2008.

[3] L.I. Meho and K. Yang. Impact of data sources on citation counts and rankings of LIS faculty: Web of Science vs. Scopus and Google Scholar. *Journal of the American Society for Information Science and Technology*, 58(13):2105–2125, 2007.

[4] B. Meyer, C. Choppy, J. Staunstrup, and J. van Leeuwen. Research evaluation for computer science. *Communications of the ACM*, 2009.

[5] D. Torres-Salinas, E. Lopez-Cozar, and E. Jimenez-Contreras. Ranking of departments and researchers within a university using two different databases: Web of science versus scopus. *Scientometrics*, accepted, 2009. doi:10.1007/s11192-008-2113-9.

[6] M. Y. Vardi. Conferences vs. journals in computing research. *Commun. ACM*, 52(5):5–5, 2009.

[7] J. Wainer, E. Xavier, and F. Bezerra. Scientific production in computer science: A comparative study of brazil and other countries. *Scientometrics*, accepted, 2009. doi:10.1007/s11192-008-2156-y.