Scholarometer: A System for Crowdsourcing Scholarly Impact Metrics

Jasleen Kaur, Mohsen JafariAsbagh, Filippo Radicchi, and Filippo Menczer
Center for Complex Networks and Systems Research
School of Informatics and Computing, Indiana University, Bloomington, USA

ABSTRACT

Scholarometer (`scholarometer.indiana.edu`) is a social tool developed to facilitate citation analysis and help evaluate the impact of authors. The Scholarometer service allows scholars to compute various citation-based impact measures. In exchange, users provide disciplinary annotations of authors, which allow for the computation of discipline-specific statistics and discipline-neutral impact metrics. We present here two improvements of our system. First, we integrated a new universal impact metric h_s that uses crowdsourced data to calculate the global rank of a scholar across disciplinary boundaries. Second, improvements made in ambiguous name classification have increased the accuracy from 80% to 87%.

1. SCHOLAROMETER

Scholarly classifications systems like Web of Science, MeSH for life sciences, and ACM CCS for computing, are based on a top-down approach in which the ontology is maintained by curators. As a result, these disciplinary categories do not accommodate the trend toward interdisciplinary scholarship and the continual emergence of new areas at disciplinary boundaries. Disciplinary boundaries create similar hurdles for measuring scholarly impact.

Crowdsourcing approaches can empower scholars to annotate each other's work. The crowdsourcing model has the added advantage that when combined with citation information, it can enable the collection of statistical data necessary for the computation of cross-disciplinary impact metrics. *Scholarometer* (`scholarometer.indiana.edu`) is a social tool for scholarly services developed in our lab with the dual aim of exploring the crowdsourcing of disciplinary annotations and developing cross-disciplinary impact metrics [4, 5]. These two aims are closely related and mutually reinforcing. The annotations enable the collection of discipline-specific statistics and therefore the computation of universal impact metrics. In turn, the service provided by computing these metrics works as an incentive for the users to provide the annotations.

WebSci'14, June 23–26, 2014, Bloomington, IN, USA.
ACM 978-1-4503-2622-3/14/06.
http://dx.doi.org/10.1145/2615569.2615669 .

Tools exist for both citation analysis (e.g., Publish or Perish [2]) and social management of bibliographic records (e.g., Mendeley). To our knowledge, Scholarometer is the first system that attempts to couple these two functions with the goal of achieving a synergy between disciplinary annotations and universal impact metrics [4, 5]. Social tagging of scholarly work has already been adopted in popular systems such as Bibsonomy [1] and many others. In the folksonomies that result from these social tagging systems, tags are assigned to papers. Tags have also been used to describe journals [3]. In Scholarometer, users tag authors instead. Recently, similar skill endorsements have been introduced by systems such as LinkedIn and ResearchGate. Currently, our system is the only one that makes these annotations publicly available.

The Scholarometer interface (Figure 1) lets users query and tag authors. The tagging interface implements a compromise between the use of a controlled vocabulary and free tagging; the user must enter at least one subject category from the Thomson-Reuters/ISI citation indices and can enter any free tags without additional constraints. Facilities are available for sorting, filtering, deleting, merging, and exporting records. Finally, a citation analysis panel reports on various impact measures.

As of March 2014, the queries submitted to Scholarometer resulted in a collection of citation data about 39,000 authors of 2.8 million articles in 2,400 disciplines. Further statistics for authors and disciplines are available on the Scholarometer website.

Scholarometer provides ways to share the crowdsourced data with the research community via an API and to explore the data through interactive visualizations of discipline networks and author networks. These visualizations can help identify potential referees, members of program committees and grant panels, collaborators, and so on.

2. GLOBAL RANKING BY IMPACT

Our tool has been used to evaluate many proposed impact metrics with the goal of providing "universal" metrics that allow for the comparison of author impact across disciplines [5, 6]. One of these metrics, h_s, has been shown to be able to remove discipline bias [6]. We defined h_s as the h-index of an author normalized by the average h of the author's discipline. We have integrated this universal impact metric in Scholarometer. However, it is not trivial to compute the global ranking of scholars who are annotated with multiple disciplines. A multi-disciplinary author will have

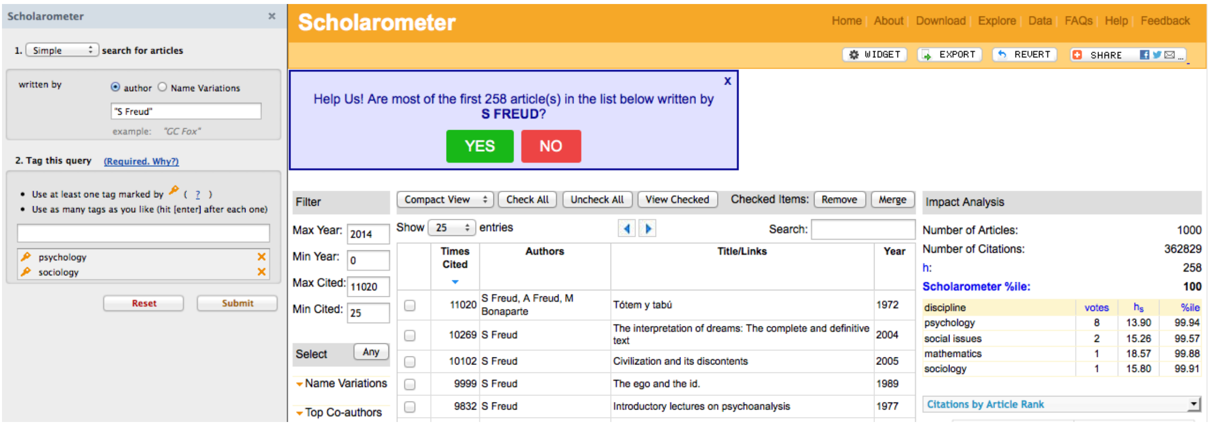

Figure 1: Illustration of the Scholarometer interface. The box at the top allows the user to provide binary feedback about whether the queried name is ambiguous.

multiple h_s values. To obtain a single Scholarometer rank (see Figure 1), we considered the following methods:

$$rank_s = rank(\sum_{d \in D} w_d.rank_d(h)) \quad (1)$$

$$rank_s = rank(h/\sum w_d.\langle h \rangle_d) \quad (2)$$

$$rank_s = rank(\sum_{d \in D} w_d.h/\langle h \rangle_d) \quad (3)$$

where d is a discipline in the set D of disciplines, $\langle h \rangle_d$ is the average h of the authors annotated with discipline d, w_d corresponds to the number of times an author is tagged with discipline d, and, finally, $rank_d$ is the rank of an author within discipline d.

We computed the ranks of all scholars based on all three methods. We rejected method (1) as it is sensitive to the local rank of the author, especially in those disciplines with few authors. Methods (2) and (3) produce the same author rankings, except when authors are tagged with more than one discipline, some of which are unreliable. Method (2) tends to penalize authors who are tagged with unreliable disciplines. Let us consider the example of an author with $h = 10$ who is tagged once with "computer science (cs)" ($\langle h_{cs} \rangle = 2$). Both methods (2) and (3) would produce the same value of $h_s(cs) = 5$: this implies that the author is 5 times above the average of authors in "cs"; this value will be used to obtain her/his global rank. Now, let us imagine that the author is also tagged once with the unreliable discipline "underwater basket weaving (ubw)", that has only this one author; the discipline's $\langle h_{ubw} \rangle$ would be 10 and the author's $h_s(ubw) = 1$. According to the two methods, the combined h_s value would be 1.67 (method (2)) and 3 (method (3). To avoid such penalizations, we adopted method (3) to generate the global Scholarometer rank.

3. NAME DISAMBIGUATION

The name ambiguity problem is especially challenging in the field of bibliographic digital libraries. The problem is amplified when names are collected from heterogeneous sources and leads to computation of biased impact metrics. This is the case in the Scholarometer system, which performs bibliometric analysis by cross-correlating author names in user queries with those retrieved from digital libraries. The uncontrolled nature of user-generated annotations is very valuable but creates the need to detect ambiguous names. We obtained promising results in the scholar name disambiguation problem by employing three kinds of heuristic features based on citations, publications, and crowdsourced topics [7]. However, ambiguous names remain a serious challenge for bibliometric analysis.

Here we report on improvements in the ambiguous name detection, achieved by integrating feedback from users. We added a button in the results window (see Figure 1) to obtain feedback about whether the query is ambiguous or not. This is the only feedback we collect from users, who cannot modify our database. We used the collected feedback to retrain the classifier. This way, the training set increased from approximately 500 manually labeled authors to 3,350 authors at the time of this writing. We used the same four features based on citations, crowdsourced topics, and publication metadata with the random forest algorithm, as it outperformed the previous logistic regression algorithm [7]. Our modified approach can detect ambiguous author names in crowdsourced scholarly data with an improved accuracy of 87% compared to the previously reported accuracy of 80%.

Acknowledgements. We thank Xiaoling Sun, Diep Thi Hoang, and Lino Possamai for their contributions to the Scholarometer project. This work is partly supported by the Lilly Endowment and NSF (award IIS-0811994).

4. REFERENCES

[1] D. Benz, A. Hotho, R. Jäschke, B. Krause, F. Mitzlaff, C. Schmitz, and G. Stumme. The social bookmark and publication management system bibsonomy. *The VLDB Journal*, 19(6):849–875, 2010.

[2] A.-W. Harzing. *The publish or perish book*. Tarma Software Research, 2010.

[3] S. Haustein and I. Peters. Using social bookmarks and tags as alternative indicators of journal content description. *First Monday*, 17(11), 2012.

[4] D. T. Hoang, J. Kaur, and F. Menczer. Crowdsourcing scholarly data. In *Proc. Web Science Conference: Extending the Frontiers of Society On-Line (WebSci)*, 2010.

[5] J. Kaur, D. T. Hoang, X. Sun, L. Possamai, M. JafariAsbagh, S. Patil, and F. Menczer. Scholarometer: A social framework for analyzing impact across disciplines. *PloS one*, 7(9):e43235, 2012.

[6] J. Kaur, F. Radicchi, and F. Menczer. Universality of scholarly impact metrics. *Journal of Informetrics*, 7(4):924—932, 2013.

[7] X. Sun, J. Kaur, L. Possamai, and F. Menczer. Ambiguous author query detection using crowdsourced digital library annotations. *Information Processing and Management*, 49(2):454–464, 2013.