Dynamic Ranking of IEEE VIS Author Importance

Daniel Witschard*
Linnaeus University

Ilir Jusufi†
Linnaeus University

Andreas Kerren‡
Linköping University
Linnaeus University

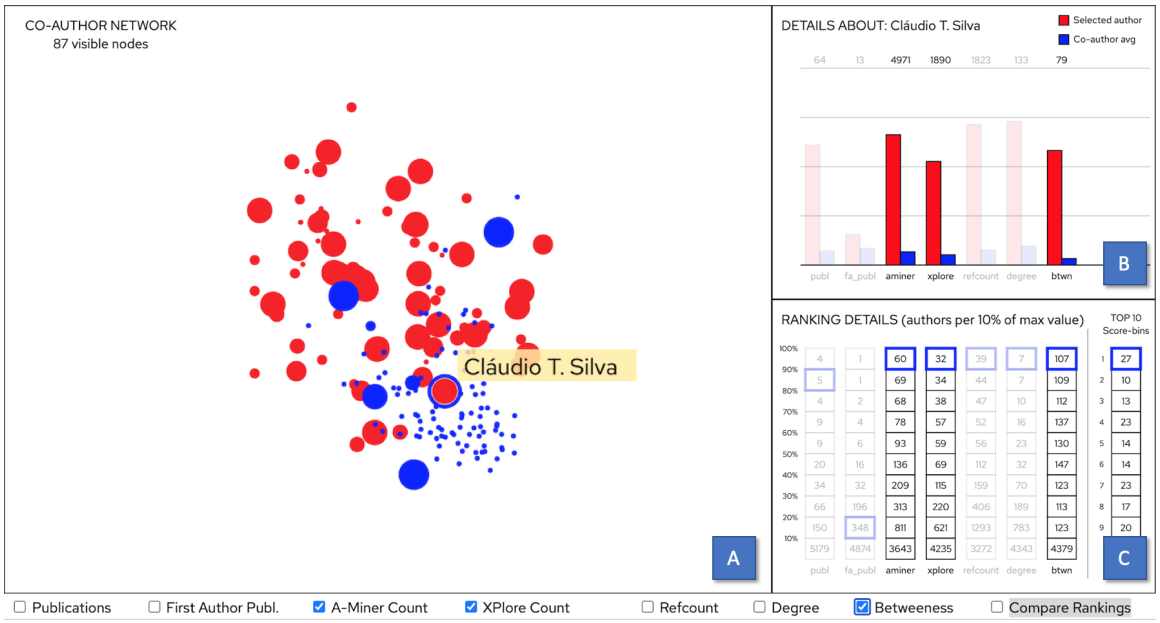

Figure 1: The visualisation consists of three linked views: (A) the Co-Author Network View, (B) the Author Details View, and (C) the Ranking View. When ranking criteria are activated, the node size is scaled to the current ranking so that only the most important authors are visible. The user may enable different ranking criteria and execute "on-hover-queries". When hovering an author node, the specific details of the author (together with the co-author average) is displayed as well as the author's position in the different rankings. The details of enabled criteria are highlighted and details of disabled criteria are faded, but still visible, to provide "information scent".

Abstract

The ranking of authors is an important task within the field of scientometrics, and several different methods and criteria exist. In this poster abstract, we present an interactive visualization approach for exploring combinations of several different ranking criteria for a given set of publications and its associated co-author network. Our visualization tool allows the user to gain insights into the relative importance of individual authors as well as into the interdependency of different ranking criteria.

Index Terms: Human-centered computing—Visualization—Visualization application domains—Visual analytics

1 Introduction

The concept of bibliometrics can be described as "the application of mathematical and statistical methods to books and other media", and within the subfield of scientometrics the focus lies on analyzing the quantitative aspects of scientific publications and their use. One central challenge is to try to establish relative rankings of importance/influence (for both publications and authors) since such data can be highly valuable for a better understanding of a scientific field. This since questions like "What are the most influential publications within the field of X?" or "Who are the most influential authors within the field of Y?" are natural starting points for performing tasks, such as exploring a field that was previously unknown or for summarizing a previously known one. Many aspects of the visualization of scientific publications and corresponding scientometrics data have been thoroughly explored in current research [2, 5] but some areas are still less covered. While there has been an extensive analysis of co-author networks [4] and also of the general concept of author ranking [1], the specific visualization of author rankings seems to have attracted less academic interest (or has been bundled together with other tasks). We therefore see an opportunity for a visualization specifically targeting this task and with a twofold aim to: (1) provide an easy-to-use tool for dynamically exploring the most influential authors of a document set, and (2) provide a purely visual method for assessing the interdependency of different ranking criteria.

2 Architecture

The data set that we use for our application is the IEEE VIS publication data set which contains information of articles published at the IEEE VIS conferences [3]. From this data, we have extracted about 3,000 articles published during the period 1990-2018. And

*e-mail: daniel.witschard@lnu.se
†e-mail: ilir.jusufi@lnu.se
‡e-mail: andreas.kerren@liu.se

Poster Abstract, IEEE Visualization and Visual Analytics (VIS '21), online conference, 2021.

in our implementation, we use the associated co-author network with about 5,500 authors. The visualization consists of an asynchronous backend implemented in Python and a web page frontend implemented in HTML/D3. The backend extracts statistics/data for the following 7 ranking criteria: (1) Number of publications, (2) Number of publications where the author appears as first author, (3) Aggregated Aminer citation count as per column value in the data set, (4) Aggregated IEEE Xplore citation count as per column value in the data set, (5) Aggregated count of outgoing references as per column value in the data set, (6) Degree centrality in the co-author network, and (7) Betweenness centrality in the co-author network. A full ranking table for all authors is created for each individual criterion (i.e., how the authors rank with regards to this specific criterion only), and each table is stored in a separate file. The pre-calculated ranking files are loaded into the visualization upon start which allows for high responsiveness since only minor calculations (typically merging several ranking tables into one combined result) are performed within the browser.

3 Visualization Approach

The approach is designed to be as responsive as possible, and therefore a major part of the functionality is implemented as "on-hover-queries" that are automatically executed when the cursor is moved over an object. As can be seen in Figure 1, the visualization consists of three linked views where the details of a specific author, and his/her co-authors are displayed with respect to the currently selected ranking criteria. When the visualization is loaded, all nodes of the co-author network are displayed in a circularly constrained force-based layout where the links have been suppressed to avoid visual clutter and to allow for efficient rendering. After enabling the different ranking criteria, a combined total ranking is calculated, and the size of the author nodes is automatically scaled to the current ranking so that only the most highly ranked authors remain visible. Hovering an author node displays the specific details of the author in relation to the average of his/her co-authors as well as the position in the different rankings, see Figure 1. Furthermore, hovering a cell in the Ranking View highlights the corresponding authors in the network view.

The application also includes functionality for assessing the level of interdependency of the ranking criteria which is important in order not to be misleading. For instance, we could suspect that the number of publications might be positively correlated to the citation counts and furthermore that the two different citation counts might show high positive correlation (i.e., a high value on one will most likely mean that the value is high also on the other). Therefore, scoring high on several dependent criteria can not be considered to be as significant as scoring high on several independent criteria. One well established design choice for assessing dependency would be to use a scatterplot matrix where the pairwise ranking orders of the different criteria are plotted against each other, and the level of correlation is calculated by fitting a regression line. However, this solution may not be very intuitive for someone who is not accustomed to statistical analysis, and we therefore opted to try a novel and purely visual solution. The idea is to display the Top 50 ranked authors from each criterion and use a color-encoding so that the independent criteria easily stand out on the display. Each author is represented by a "piechart glyph", and the colors encode the criteria identity, see Figure 2. An author who is among the Top 50 for only one criterion will have a glyph with only one color, and an author who is among the Top 50 for several criteria will have a piechart glyph with the corresponding number of colors. In a totally independent scenario (i.e., no author is among the Top 50 for more than one criterion), we would have n sets of 50 single-colored glyphs (each set having its own unique color and n being the number of criteria); and in a totally dependent scenario (i.e., all Top 50 lists are identical), we would have 50 piechart glyphs, each with n differently colored compartments. With this encoding, the most visible color corresponds to the most independent criteria, and the number of displayed glyphs gives an assessment of the overall interdependency of the chosen set of criteria (i.e., the fewer the dots the higher the dependency).

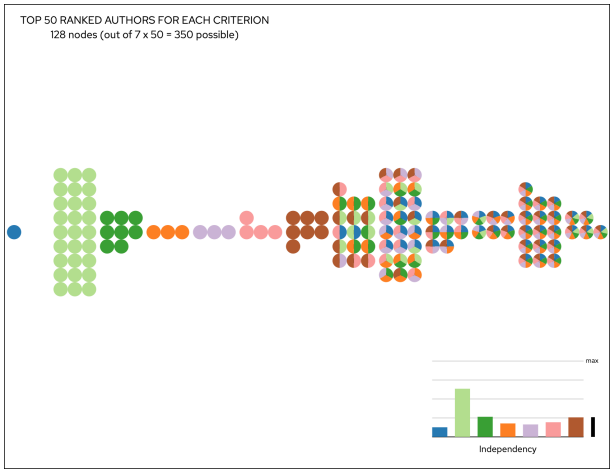

Figure 2: When in "Compare Rankings" mode, the visualization displays the top 50 ranked authors with regard to each individual criterion. Each criterion is color-coded with a unique color; and the more independent the criterion is, the more of its color is displayed. To facilitate the assessment, the glyphs are aligned in spatial substrates according to the number of colors that they contain. From the screenshot above, it is easy to perceive that the light green criterion (First Author Publications) is the most independent. Furthermore, we can see that the currently chosen set of criteria is highly interdependent since only 128 dots (out of the $7 * 50 = 350$ possible) are displayed and that many of them are multicolored.

4 Conclusion

In this poster abstract, we have presented a visualisation that targets dynamic ranking of author importance on the IEEE VIS publication data set. Our application can give insights into the relative importance of different authors and also into the interdependency of different ranking criteria. The concept is easily extendable to other data sets and/or other ranking criteria.

References

[1] T. Amjad, A. Daud, and N. R. Aljohani. Ranking authors in academic social networks: a survey. *Library Hi Tech*, 36(1):97–128, Jan 2018. doi: 10.1108/LHT-05-2017-0090

[2] P. Federico, F. Heimerl, S. Koch, and S. Miksch. A survey on visual approaches for analyzing scientific literature and patents. *IEEE Transactions on Visualization and Computer Graphics*, PP, 09 2016. doi: 10.1109/TVCG.2016.2610422

[3] P. Isenberg, F. Heimerl, S. Koch, T. Isenberg, P. Xu, C. D. Stolper, M. Sedlmair, J. Chen, T. Möller, and J. Stasko. Vispubdata.org: A metadata collection about IEEE Visualization (VIS) publications. *IEEE Transactions on Visualization and Computer Graphics*, 23(9):2199–2206, Sept. 2017. doi: 10.1109/TVCG.2016.2615308

[4] S. Kumar. Co-authorship networks: A review of the literature. *Aslib Journal of Information Management*, 67:55–73, 01 2015. doi: 10.1108/AJIM-09-2014-0116

[5] J. Liu, T. Tang, W. Wang, B. Xu, X. Kong, and F. Xia. A survey of scholarly data visualization. *IEEE Access*, PP:1–1, 03 2018. doi: 10.1109/ACCESS.2018.2815030