Comparing and Recommending Conferences

Grettel Monteagudo García[1], Bernardo Pereira Nunes[1,2],
Giseli Rabello Lopes[3], Marco Antonio Casanova[1], Luiz André P. Paes Leme[4]

[1] Pontifical Catholic University of Rio de Janeiro, RJ, Brazil

[2] Federal University of the State of Rio de Janeiro, Rio de Janeiro, RJ, Brazil

[3] Federal University of Rio de Janeiro, Rio de Janeiro, RJ, Brazil

[4] Federal Fluminense University, Niterói, RJ, Brazil

{ggarcia,bnunes,casanova}@inf.puc-rio.br, giseli@dcc.ufrj.br,
lapaesleme@ic.uff.br

Abstract. *This paper first introduces techniques for comparing conferences that use familiar similarity measures and a new measure based on co-authorship communities. Then, it focuses on two families of techniques for conference recommendation, the first one based on the similarity measures and the second on the idea of finding the most related authors in the co-authorship network. The experiments suggest that the best performing techniques are: the technique for comparing conferences that uses the new similarity measure based on co-authorship communities; and the conference recommendation technique that explores the most related authors in the co-authorship network.*

Resumo. *Este trabalho introduz técnicas para comparar conferências a partir de medidas de similaridade clássicas e uma nova medida baseada em comunidades de coautores. Em seguida, o trabalho foca em dois grupos de técnicas para recomendação de conferências: o primeiro se baseia nas medidas de similaridade entre conferências e o segundo na estratégia de encontrar os autores mais relacionados na rede de coautores. Os experimentos sugerem que as melhores técnicas são: a técnica de comparação de conferências que utiliza a nova medida de similaridade baseada em comunidades de coautores; e a técnica para recomendação de conferências que explora os autores mais relacionados na rede de coautores.*

1. Introduction

In this work, we propose, implement and evaluate techniques to automatically compare and recommend conferences. The techniques for comparing conferences adopt familiar similarity measures, such as the Jaccard similarity coefficient, the Pearson correlation similarity and the cosine similarity, and a new similarity measure, the *co-authorship network community similarity index*.

We proceed to define two families of conference recommendation techniques. The first family of techniques adopts collaborative filtering based on the conference similarity measures investigated in the first part of the paper. The second family includes two techniques based on the idea of finding, for a given author, the strongest related authors in the co-authorship network and recommending the conferences that his co-authors usually publish in. The first recommendation technique uses the *weighted semantic connectivity score – WSCS* [Nunes et al. 2013], an index for measuring relatedness of actors. However, since this index proved to be accurate, but quite costly for large co-authorship networks, we define a second recommendation technique that adopts an approximation of the *WSCS*, but which is much faster to compute and as accurate as the *WSCS*.

The paper concludes with a description of experiments to test and compare the techniques with data extracted from a triplified version of the DBLP repository, which stores Computer Science bibliographic data for more than 4,500 conferences and 1,500 journals (as of early 2016). The experiments were performed using a Web-based application that enables users to interactively analyze and compare a set of conferences. The experiments indicate that the best performing techniques are: (1) the technique for comparing conferences that uses the new co-authorship network community similarity index; and (2) the conference recommendation technique that explores the co-authorship network and adopts an approximation of the WSCS. These two techniques are therefore the major contributions of this paper.

This paper is structured as follows. Section 2 summarizes related work. Section 3 presents the techniques for comparing conferences. Section 4 covers the conference recommendation techniques. Section 5 briefly presents an application that implements the techniques and covers the evaluation of the techniques. Finally, Section 6 contains the conclusions and proposes future work.

2. Related Work

Henry et al. (2007) analyzed a group of the four major conferences in the field of Human-Computer Interaction (HCI). The authors discovered many global and local patterns using only article metadata, such as authors, keywords and year. Blanchard (2012) presented a ten-year analysis of the paper production in Intelligent Tutoring Systems (ITS) and Artificial Intelligence in Education (AIED) conferences and shows that Western, Educated, Industrialized, Rich, and Democratic bias observed in psychology may be influencing AIED research. Chen, Zhang and Vogeley (2009) proposed an extension of the contemporary co-citation network analysis to identify co-citation clusters of cited references. Intuitively, the authors synthesize thematic contexts in which these clusters are cited and trace how the research focus evolves over time. Gasparini, Kimura and Pimenta (2013) present a visual exploration of the field of Human Computer Interaction in Brazil from a fifteen-year analysis of paper production in the Brazilian Symposium on Human Factors in Computing Systems (IHC). Chen, Song and Zhu (2007) opened a wide range of opportunities for research agendas and trends in ER conferences.

Zervas et al. (2014) applied social network analysis (SNA) metrics to analyzing the co-authorship network of the Educational Technology & Society (ETS) Journal. Procópio, Laender and Moro (2011) did a similar analysis for the Databases field. Cheong and Corbitt (2009a, 2009b) analyzed the Pacific Asia Conference on Information Systems and the Australasian Conference on Information Systems.

Recently, Lopes et al. (2015, 2016) carried out an extensive analysis of the WEBIST conferences, involving authors, publications, conference impact, topics coverage, community analysis and other aspects. Linked Data principles to publish conference data were also used by [Batista and Lóscio 2013; Lopes et al. 2015, 2016].

Contrasting with the above references, in this paper we propose, implement and evaluate several techniques to compare conferences in general, and not a specific conference series. The current implementation works with the triplified version of the DBLP repository, which covers the vast majority of Computer Science conferences.

Turning to conference recommendation, Luong et al. (2012) proposed and compared three recommendation methods for conferences. To define the methods, they recursively collected the co-authors of the co-authors, until a network of 3 levels deep was created, in a set of the more important co-authors. The best performing recommendation method, which we will refer simply as the *most frequent conference,* weights the contributions of each co-author by the number of papers they have co-authored with the main author. It is defined as:

$$coauthor\_CONF_i = \sum_{m=1}^{N} coauthors\_w_{i,m} \quad (1)$$

where N is main author(s) of the test paper and $coauthors\_w_{i,m}$ is the co-authors' conference weight between the main author m and her co-authors in the network and is defined as:

$$coauthors\_w_{i,m} = \sum_{k=1}^{CoA} \left(nfreq\_CONF_{i,m} + nfreq\_CONF_{i,k}\right) * w\_CoA_{k,m} \quad (2)$$

where CoA is a co-author(s) of the main author m who have published respectively at conference i, $w\_CoA_{k,m}$ is the number of times a main author m co-authored papers with another member k in the network, and $nfreq\_CONF_{i,m}$ is the probability of the author m to publish in conference i.

In this paper, we propose two conference recommendation techniques based on the social network analysis of the co-authorship network, but we adopt a measure of the strength of the connections between the authors in the network which is computed differently from Luong's method. We first propose to estimate the relatedness of actors in a social network by using a semantic connectivity score [Nunes et al. 2013], denoted SCS, which is in turn based on the Katz index [Katz 1953]. This score takes into account the number of paths between two nodes of the network and the accumulated weights of these paths. Then, we propose a second score that approximates the SCS score and that uses the shortest path between two nodes. In addition to these two strategies, we also propose to construct a utility matrix and to implement recommendation techniques based on collaborative filtering using a utility matrix.

3. Comparing Conferences

In what follows, we use the following notation:

- C is a set of conferences
- A is a set of authors
- P is a set of papers

- $p: A \rightarrow 2^P$ is a function that assigns to each author $i \in A$ the set of papers $p(i) \subseteq P$ that author *i* published (in any conference)
- $pc: A \times C \rightarrow 2^P$ is a function that assigns to each author $i \in A$ and each conference $x \in C$ the set of papers $pc(i,x) \subseteq P$ that author *i* published in conference *x*
- A_x and A_y are the set of authors that published in conferences x and y, that is, $A_x = \{i \in A \ / \ |pc(i,x)| > 0\}$ and, likewise, $A_y = \{i \in A \ / \ |pc(i,y)| > 0\}$.
- $A_{x,y}$ is the set of authors that published in both conferences *x* and *y*, that is, $A_{x,y} = \{i \in A \ / \ |pc(i,x)| > 0 \ \wedge \ |pc(i,y)| > 0\}$
- $G_x = (N_x, E_x)$, the *co-authorship network* of conference *x*, is an undirected and unweighted graph where $i \in N_x$ indicates that author *i* published in conference *x* and $\{i,j\} \in E_x$ represents that authors *i* and *j* co-authored one or more papers published in conference *x*

In what follows, we adapt familiar similarity measures to conferences or authors and introduce a new measure called *communities similarity*.

The *Jaccard similarity coefficient* for conferences x and y is defined as

$$jaccard\_sim(x,y) = \left(A_x \cap A_y\right) / \left(A_x \cup A_y\right) \quad (3)$$

The utility matrix expresses the preferences of an author for a conference to publish his research. More formally, the utility matrix $[r_{x,i}]$ is such that the lines represent conferences and the columns represent authors and is defined as:

$$r_{x,i} = \frac{|pc(i,x)|}{|p(i)|} \quad (4)$$

Based on the utility matrix $[r_{x,i}]$, we define the Pearson's correlation coefficient similarity between conferences *x* and *y* as follows:

$$pearson\_sim(x,y) = \frac{\sum_{i \in A_{x,y}} \left(r_{x,i} - \overline{r_x}\right)\left(r_{y,i} - \overline{r_y}\right)}{\sqrt{\sum_{i \in A_{x,y}} \left(r_{x,i} - \overline{r_x}\right)^2 \sum_{i \in A_{x,y}} \left(r_{y,i} - \overline{r_y}\right)^2}} \quad (5)$$

where $\overline{r_x}$ is the average of the elements of line x of the utility matrix (and likewise for $\overline{r_y}$).

Again based on the utility matrix $[r_{x,i}]$, we define the cosine similarity between conferences *x* and *y* as follows:

$$cos\_sim(x,y) = \frac{\sum_{i \in A_{x,y}} \left(r_{x,i} r_{y,i}\right)}{\sqrt{\sum_{i \in A_{x,y}} \left(r_{x,i}\right)^2 \sum_{i \in A_{x,y}} \left(r_{y,i}\right)^2}} \quad (6)$$

We introduce a new similarity measure between conferences based on communities defined over the co-authorship network of the conferences.

Given the co-authorship network $G_x = (N_x, E_x)$ of conference *x*, we define an *author community* c_x of *x* as the net of nodes of a connected component of G_x. Let c_x and c_y be author communities in the co-authorship networks of conferences x and y, respectively. We say that c_x and c_y are *equivalent* w.r.t. a similarity measure *sim* and a threshold level α iff $sim(c_x, c_y) \geq \alpha$. For example, *sim* may be defined using Jaccard similarity coefficient between pairs of conferences introduced above.

Let C_x and C_y be the sets of communities of conferences x and y, respectively. Let $EQ[sim, \alpha](x, y)$ be the set of communities in the co-authorship network of conference x that have an equivalent community in the co-authorship network of conference y (and symmetrically $EQ[sim, \alpha](y, x)$).

The *co-authorship network communities similarity* (based on a similarity measure *sim* and a threshold level α) between conferences x and y is then defined as:

$$c\_sim[sim, \alpha](x, y) = \frac{|EQ[sim, \alpha](x, y)|}{\min\{|C_x|, |C_y|\}} \quad (7)$$

Note that $C_x > 0$ and $C_y > 0$ since G_x and G_y must have at least one node each and therefore at least one connected component each.

4. Recommending Conferences

4.1. Conference Recommendation Techniques based on Classical Similarity Measures

As defined in [Leskovec, Rajaraman and Ullman 2014], in a recommendation system, there are two classes of entities – *users* and *items*. Users have preferences for certain items, which must be extracted from the data. The data itself is represented as a utility matrix giving, for each user-item pair, a value that represents what is known about the degree of preference or *rating* of that user for that item. An unknown rating implies that there is no explicit information about the user's preference for the item. The goal of a recommendation system is to predict the unknown ratings in the utility matrix.

In our context, we recall from Section 3 that the utility matrix $[r_{x,i}]$ is such that $r_{x,i}$ expresses the preference (i.e., rating) of an author *i* for a conference *x* to publish his research. To predict an unknown rating, we compute the similarity between conferences and detect their nearest neighbors or most similar conferences. With this information, the rating of conference *x* for author *i* is defined as follows:

$$CF(x, i) = \frac{\sum_{y \in S_x} (r_{y,i}) sim(x, y)}{\sum_{y \in S_x} sim(x, y)} \quad (8)$$

where S_x is the set of conferences most similar to x and $r_{y,i}$ is the rating of conference y for author i.

Therefore, we may immediately define a family of conference recommendation techniques based on the utility matrix and the classical similarity measures introduced in Section 3, that we call *CF-Jaccard*, *CF-Pearson*, *CF-Cosine* and *CF-Communities*, according to the similarity measure adopted. Section 5 discusses how they perform in detail.

4.2. Conference Recommendation Techniques based on the Weighted Authorship Network

Recall from Section 3 that $p: A \rightarrow 2^P$ is the function that assigns to each author $i \in A$ the set of papers $p(i) \subseteq P$ that author i published (in any conference). The *weighted co-authorship network* based on p is the edge-weighted undirected graph $G = (N, E, w)$, where $i \in N$ represents an author, $\{i, j\} \in E$ indicates that i and j are co-authors, that is, $\{i, j\} \in E$ iff $p(i) \cap p(j) \neq \emptyset$, and $w(\{i, j\})$ assigns a weight to the co-authorship relationship between i and j and is defined as:

$$w(\{i,j\}) = \frac{|p(i) \cap p(j)|}{|p(i) \cup p(j)|} \quad (9)$$

Hence, the larger $w(\{i, j\})$ is, the stronger the co-authorship relationship will be: if authors i and j co-authored all papers they published, then $w(\{i, j\}) = 1$; and if they have not co-authored any papers, then the edge $\{i, j\}$ does not exist.

The second family of conference recommendation techniques explores the weighted co-authorship network and adopts two scores: the *weighted semantic connectivity score – WSCS* and the *modified weighted semantic connectivity score – MWSCS*. Hence, these techniques are called *WSCS-based* and *MWSCS-based recommendation techniques*.

The *weighted semantic connectivity score*, $WSCS_e$, is defined by modifying the semantic connectivity score SCS_e to take into account the weight of the paths between two authors i and j, computed as the sum of the weights of the edges in the path:

$$WSCS_e(i,j) = \sum_{w=1}^{T} \beta^w \cdot \left|paths_{<i,j>}^{<w>}\right| \quad (10)$$

where $|paths_{<i,j>}^{<w>}|$ is the number of paths of weight equal to w between i and j and T is the maximum weight of the paths and $0 < \beta \leq 1$ is a positive damping factor.

The conference recommendation technique based on $WSCS_e$ works as follows. Given an author i, it starts by computing $WSCS_e(i, j)$, the score between i and any other author j in the weighted co-authorship network. Then, it sorts authors in decreasing order of $WSCS_e$, since authors that are better related to author i will have a higher $WSCS_e(i, j)$ value. For better performance, the technique considers only the first n authors in the list ordered by $WSCS_e$. Call this set F_i. For each author j in F_i, the technique selects the conference with the highest rank, denoted $MaxC_j$. The rank of conference x for author i is defined as follows:

$$rank(x,i) = \sum_{j \in F_i \text{ and } MaxC_j = x} WSCS_e(i,j) \quad (11)$$

Since computing the $WSCS_e$ score can be very slow for large graphs, we propose to compute only the shortest paths from author i to other authors using Dijkstra's algorithm. We then redefine the score as follows:

$$MWSCS_e(i,j) = \beta^w \quad (12)$$

where w is a length of the shortest path from author i to author j. The recommendation technique remains basically the same, except that it uses the *MWSCS* score.

5. Evaluation and Results

5.1. Application Architecture

Figure 1 summarizes the architecture of the application developed to run the experiments. The *Conferences Data Service* handles queries to the triple store with conference data. The *Co-authorship Network Service* receives data from the *Conferences Data Service* and handles queries to the *Neo4j* database. When an analysis is executed, the system stores the results for future reuse; the *Previous Calculation Service* manages these functions. All experiments that follow were executed in an Intel Core Quad 3.00GHz, with 6 GB RAM, running Windows 7.

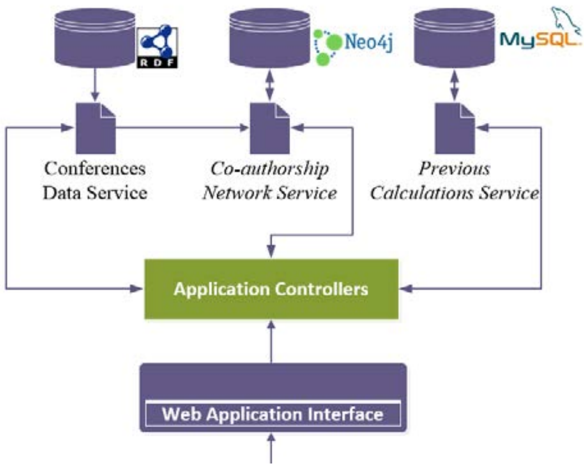

Figure 1. Web Application Architecture.

5.2. Experiments with the Conference Similarity Techniques

We evaluated the conference similarity techniques assuming that the most similar conferences should fall into the same category. We selected as benchmark the List of Computer Science Conferences defined in Wikipedia[1], which contains 248 academic computer science conferences, classified into 13 categories. That is, the categories define a partition ***P*** of the set of conferences. Then, we applied the same clustering algorithm to this set of conferences, but using each of the conference similarity measures. Finally, we compared the clusters thus obtained with ***P***. The best conference similarity measure would therefore be that which results in conference clusters that best match ***P***.

We adopted the hierarchical agglomerative clustering algorithm, which treats each conference as a singleton cluster at the outset and then successively merges (or agglomerates) pairs of clusters, using similarity measures, until achieving the desired number of clusters. To determine how similar clusters are, and agglomerate them, a linkage criterion was used. The shortest value of these links that remains at each step causes the fusion of the two clusters whose elements are involved.

Let $d(a, b)$ denote the distance between two elements a and b. Familiar linkage criteria between two sets of elements A and B are:

[1] https://en.wikipedia.org/wiki/List_of_computer_science_conferences

- Complete-linkage: the distance $D(A,B)$ between two clusters A and B equals the distance between the two elements (one in each cluster) that are farthest away from each other:

$$D(A,B) = \max\{d(a,b) \;/\; a \in A, b \in B\} \quad (13)$$

- Single-linkage clustering: the distance $D(A,B)$ between two clusters A and B equals the distance between the two elements (one in each cluster) that are closest to each other:

$$D(A,B) = \min\{d(a,b) \;/\; a \in A, b \in B\} \quad (14)$$

- Average linkage clustering: the distance $D(A,B)$ between two clusters A and B is taken as the average of the distances between all pairs of objects:

$$D(A,B) = \frac{\sum_{a\in A}\sum_{b\in B} d(a,b)}{|A||B|} \quad (15)$$

Before explaining the measures used to compare how well different data clustering algorithms perform on a set of data, we need the following definitions. Given a set of *n* elements *S* and two partitions *X* and *Y* of *S*, where *X* is the correct partition and *Y* is the computed partition, we define:

- TP (True Positive) is the number of pairs of elements in *S* that are in the same set in *X* and in the same set in *Y*
- TN (True Negative) is the number of pairs of elements in *S* that are in different sets in *X* and in different sets in *Y*
- FN (False Negative) is the number of pairs of elements in *S* that are in the same set in *X* and in different sets in *Y*
- FP (False Positive) is the number of pairs of elements in *S* that are in different sets in *X* and in the same set in *Y*

The measures to evaluate the performance of the clustering algorithms using the proposed similarity functions are:

- *Rand Index*: measures the percentage of correct decisions made by the algorithm:

$$RI = \frac{TP + TN}{TP + TN + FP + FN} \quad (16)$$

- *F-measure*: balances the contribution of false negatives by weighting the recall through a parameter $\beta > 0$:

$$F = \frac{(\beta^2 + 1)P.R}{(\beta^2 P) + R} \quad (17)$$

where $P = \frac{TP}{TP+FP}$ and $R = \frac{TP}{TP+FN}$

Figure 2 shows the Rand index obtained by executing the hierarchical agglomerative clustering algorithm with different linkages criteria, using the Jaccard, Pearson, cosine and communities similarity measures. Note that, in general, the algorithm based on communities similarity had the best performance, followed by the Jaccard similarity. In this case, the cosine similarity had the worst behavior.

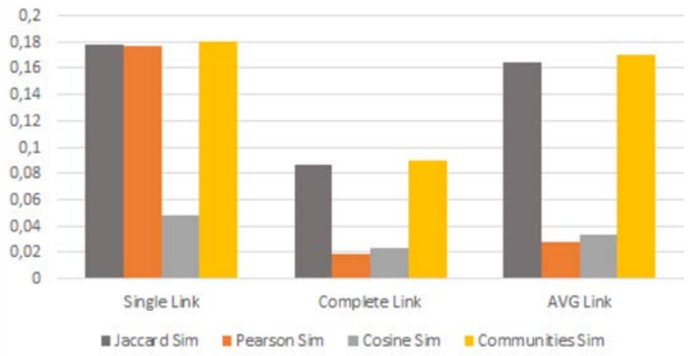

Figure 2. Rand Index of the clustering algorithms.

Figure 3 shows the F-measure obtained by executing the same algorithms. Analyzing the results presented in Figure 3, we observe that the best performances were also obtained using the communities similarity and the Jaccard similarity measures. The worst performance was obtained using the Pearson similarity measure. The algorithm using the cosine similarity measure achieved the worst performance only with the single link linkage criterion.

Therefore, these experiments suggest that the best performing algorithm is that which adopts the communities similarity measure.

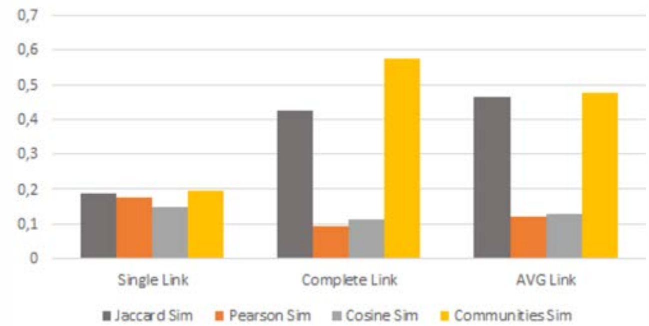

Figure 3. F-measure with β=1 of clustering algorithms.

5.3. Experiments with the Conference Recommendation Techniques

Recall that we proposed two families of recommendation techniques. One family is based on classical similarity measures – Jaccard, Pearson, and cosine similarity – and a new similarity measure, the communities similarity. These techniques are respectively called CF-Jaccard, CF-Pearson, CF-Cosine and CF-Communities. The second family includes two techniques based on the weighted and the modified weighted semantic connectivity, called WSCS-based and MWSCS-based recommendation techniques.

We evaluated the conference recommendation techniques using the same dataset as in Section 5.2, with the 248 academic computer science conferences, and selected 243 random authors to predict their conferences ranking, for that we delete all publications of the author on the conferences that we want to rank. We adopted Luong's most frequent conference technique as the benchmark (see Section 2).

Also recall that the *mean average precision* measures how good a recommendation ranking function is. Intuitively, let a be an author and $\boldsymbol{C}_a$ be a ranked list of conferences recommended for a. Let $\boldsymbol{S}_a$ be a *gold standard* for a, that is, the set of conferences considered to be the best ones to recommend for a. Then, we have:

- $Prec@k(\boldsymbol{C}_a)$, the *precision at position* k of $\boldsymbol{C}_a$, is the number of conferences in $\boldsymbol{S}_a$ that occur in $\boldsymbol{C}_a$ until position k , divided by k
- $AveP(\boldsymbol{C}_a)$, the *average precision of* $\boldsymbol{C}_a$, is defined as the sum of $Prec@k(\boldsymbol{C}_a)$ for each position k in the ranking $\boldsymbol{C}_a$ in which a relevant conference for a occurs, divided by the cardinality of $\boldsymbol{S}_a$:

$$AveP(\boldsymbol{C}_a) = \frac{\sum_k Prec@k(\boldsymbol{C}_a)}{|\boldsymbol{S}_a|} \quad (18)$$

- MAP, the *Mean Average Precision* of a rank score function over all the authors used in these experiments (represented by set $\boldsymbol{A}$) is then defined as follows:

$$MAP = average\ \{AveP(\boldsymbol{C}_a)\ /\ a \in \boldsymbol{A}\} \quad (19)$$

Consider first the two conference recommendation techniques based on the co-authorship network, the WSCS-based and MWSCS-based recommendation techniques. To compare them, we performed experiments that measured their runtime, accuracy and average precision of the Top-10 conferences of an author (thus, in this situation the maximum $|\boldsymbol{S}_a|$ value used in $AveP$ calculation is 10). Figure 4 shows the runtime results of the algorithms that implement these recommendation techniques. Note that the MWSCS-based algorithm is far more efficient than the WSCS-based algorithm.

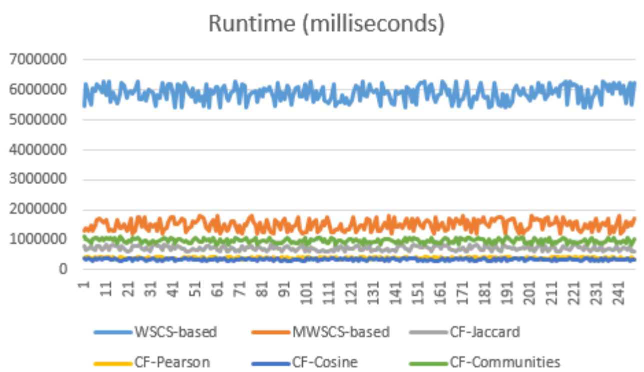

Figure 4. Runtime of the recommendation algorithms.

Table 1 shows the accuracy and *MAP* of the seven conference recommendation techniques. The two proposed techniques (first two rows of Table 1) have very similar accuracy. In fact, of the 243 authors that we tested, the balance of the correct predictions was 201 against 197. Based on these results, we may conclude that the MWSCS-based technique is much more efficient and maintains acceptable accuracy level and *MAP*, when compared with the WSCS-based technique.

Table 1 also indicates that the WSCS-based and the MWSCS-based techniques have better accuracy and *MAP* than the benchmark. The CF-Jaccard and the CF-Communities techniques have very acceptable results and very close to the benchmark, but less that the WSCS-based and the WSCS-based techniques. The CF-Pearson and CF-Cosine techniques have poor accuracy.

Table 1. Comparison of the Accuracy and *MAP* of the recommendation techniques.

Technique	Accuracy	MAP
WSCS-based	**82.72%**	**80.93%**
MWSCS-based	81.07%	80.01%
CF-Jaccard	78.19%	77.73%
CF-Pearson	55.56%	50.21%
CF-Cosine	56.79%	51.89%
CF-Communities	79.02%	77.93%
Benchmark	79.84%	77.88%

6. Conclusions

In this work, we presented techniques to compare and recommend conferences. The techniques to compare conferences are based on some classical similarity measures and on a new similarity measure based on the co-authorship network communities of two conferences. The experiments suggest that the best performance is obtained using the new communities similarity measure.

We introduced two families of conference recommendation techniques, following the collaborative filtering strategy. The first family is based on similarity measures proposed to compare conferences. The second family is based on the relatedness of two authors in the co-authorship network, using the weighted and the modified weighted semantic connectivity score. The experiments suggest that the techniques of the second family perform better than the benchmark and better than the techniques based on similarity measures. Furthermore, the technique based on the new modified weighted semantic connectivity score is much faster that the technique based on the original weighted semantic connectivity score.

As for future work, we plan to make the tool and the test datasets openly available and to expand the scope of the work to other publications datasets.

Acknowledgments

This work was partly funded by CNPq under grant 444976/2014-0, 303332/2013-1, 442338/2014-7 and 248743/2013-9 and by FAPERJ under grant E-26/201.337/2014.

References

Batista, M.G.R., Lóscio, B.F. 2013. "OpenSBBD: Usando Linked Data Para Publicação de Dados Abertos Sobre O SBBD" Proc. 28th Brazilian Symposium on Databases, Recife, Brazil, Sept. 30 – Oct. 03, 2013.

Blanchard, E.G. 2012. “On the WEIRD Nature of ITS/AIED Conferences” Proc. 11th Int’l. Conf. on Intelligent Tutoring Systems, Chania, Greece, June 14-18, pp. 280–285.

Chen, C., Song, Il-Y., Zhu, W. 2007. “Trends in Conceptual Modeling: Citation Analysis of the ER Conference Papers (1979-2005)” Proc. 11th Int’l. Conf. of the International Society for Scientometrics and Informatrics, Madrid, Spain, June 25-27, pp. 189–200.

Chen, C., Zhang, J., Vogeley, M.S. 2009. “Visual Analysis of Scientific Discoveries and Knowledge Diffusion” Proc. 12th Int’l. Conf. on Scientometrics and Informetrics - ISSI 2009, Rio de Janeiro, Brazil, July 14-17, 2009.

Cheong, F., Corbitt, B.J. 2009a. “A Social Network Analysis of the Co-Authorship Network of the Australasian Conference of Information Systems from 1990 to 2006” Proc. 17th European Conf. on Info. Systems, Verona, Italy, June 8-10, 2009.

Cheong, F., Corbitt, B.J. 2009b. “A Social Network Analysis of the Co-Authorship Network of the Pacific Asia Conference on Information Systems from 1993 to 2008” Proc. Pacific Asia Conference on Information Systems 2009, Paper 23.

Gasparini, I., Kimura, M.H., Pimenta, M.S. 2013. “Visualizando 15 Anos de IHC” Proc. 12th Brazilian Symposium on Human Factors in Computing Systems, Manaus, Brazil, pp. 238–47.

Henry, N., Goodell, H., Elmqvist, N., Fekete, J-D. 2007. “20 Years of Four HCI Conferences: A Visual Exploration” *Int’l. J. of Human-Comp. Inter.* 23(3): 239–85.

Katz, L. 1953. “A New Status Index Derived from Sociometric Analysis” *Psychometrika* 18(1): 39–43.

Leskovec, J., Rajaraman, A., Ullman, J.D. 2014. *Mining of Massive Datasets*. Cambridge University Press.

Lopes, G.R. et al. 2015. “Knowing the Past to Plan for the Future - An In-Depth Analysis of the First 10 Editions of the WEBIST Conference” Proc. 11th Int’l. Conf. on Web Information Systems and Technologies, Lisbon, Portugal, May 20-22, 2015, pp. 431–42.

Lopes, G.R. et al. 2016. “A comprehensive analysis of the first ten editions of the WEBIST conference” *Lect. Notes in Business Information Processing* 246:252–274.

Luong, H. et al. 2012. “Publication Venue Recommendation Using Author Network’s Publication History” Proc. 4th Asian Conf. on Intelligent Information and Database Systems - ACIIDS 2012, Kaohsiung, Taiwan, March 19-21, 2012, pp. 426–35.

Nunes, B.P. et al. 2013. “Interlinking Documents Based on Semantic Graphs” *Procedia Computer Science* 22: 231–40.

Procópio, P. S., Laender, A.H.F. and Moro, M.M. 2011. “Análise da Rede de Coautoria do Simpósio Brasileiro de Bancos de Dados” Proc. 26th Brazilian Symposium on Databases, Florianópolis, Brazil, Oct. 3-6, 2011.

Zervas, P. et al. 2014. “Studying Research Collaboration Patterns via Co-Authorship Analysis in the Field of TeL: The Case of Educational Technology & Society Journal” *Journal of Educational Technology & Society* 17(4): 1–16.