



HAL
open science

Expert suggestion for conference program committees

Hong Diep Tran, Guillaume Cabanac, Gilles Hubert

► **To cite this version:**

Hong Diep Tran, Guillaume Cabanac, Gilles Hubert. Expert suggestion for conference program committees. 11th International Conference on Research Challenge in Information Science (RCIS 2017), May 2017, Brighton, United Kingdom. pp. 221-232. hal-01873803

HAL Id: hal-01873803

<https://hal.science/hal-01873803v1>

Submitted on 13 Sep 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Open Archive TOULOUSE Archive Ouverte (OATAO)

OATAO is an open access repository that collects the work of Toulouse researchers and makes it freely available over the web where possible.

This is an author-deposited version published in : <http://oatao.univ-toulouse.fr/>
Eprints ID : 19048

The contribution was presented at RCIS 2017 :

<http://rcis-conf.com/rcis2017/index.html>

To link to this article URL :

<http://doi.org/10.1109/RCIS.2017.7956540>

To cite this version : Tran, Hong Diep and Cabanac, Guillaume and Hubert, Gilles *Expert suggestion for conference program committees*. (2017) In: 11th International Conference on Research Challenge in Information Science (RCIS 2017), 10 May 2017 - 12 May 2017 (Brighton, United Kingdom).

Any correspondence concerning this service should be sent to the repository administrator: staff-oatao@listes-diff.inp-toulouse.fr

Expert suggestion for conference program committees

Hong Diep Tran, Guillaume Cabanac, Gilles Hubert
Institut de Recherche en Informatique de Toulouse – IRIT
Université Toulouse 3 Paul Sabatier, 118 route de Narbonne, F-31062 Toulouse, France
Email: {Hong-Diep.Tran, Guillaume.Cabanac, Gilles.Hubert}@irit.fr

Abstract—With the rapid development of science, the academic community requires higher and higher quality of the published articles. This great responsibility is placed on editorial boards of journals, on program committees of conferences and their members. In addition, with a large number of scientific conferences held each year, searching for experts that would be invited to join the program committees is an increasingly hard task. In this paper, we propose an expert modelling method on different indicators of expertise, in particular on citation networks in order to build suggestions of members for program committees of conferences. Contrary to most of the works about expert search in the literature which are based on textual sources of evidence to define the expertise of an expert candidate, our model leverages non-textual indicators only. We evaluate the proximity of an expert candidate according to the weights of his/her various relations with the conference for which the candidate is proposed to be member of the next program committee. We report various experiments in the context of the SIGIR conference editions. These experiments give hints on the composition of program committees and on the effectiveness of our approach.

Keywords—Expertise, academia, program committee.

I. INTRODUCTION

The peer review process in science is the cornerstone supporting the communication of scientific progress [1]. Manuscripts submitted to a scientific journal are handled by the journal’s editorial board composed of gatekeepers who serve as referees [2]. Journal editorial boards are rather stable in time: only a fraction of the board is usually renewed each year. For some scientific fields, however, journals are not the only way to communicate research results. In computer science, for instance, international conferences play a key role in this respect. Paper submissions — usually full papers, not just abstracts as in other fields — are refereed by three or more members of the conference Program Committee (PC). Unlike for journals, PCs are renewed each year upon suggestion of the PC chairs. Constituting an efficient PC is a big challenge, as one needs to identify and invite scholars both active in all the topics addressed by the conference and diverse in many aspects (location, age, gender, and so on). It is an even more daunting task for large conferences such as CIKM, SIGIR, or WWW that attract hundreds of submissions [3]. Various works showed the interest of automating different tasks related to conference organisation [4], [5], [6].

The issue tackled in this paper relates to the expertise retrieval question raised some fifteen years ago [7]. Diverse contexts find an interest in expert search, such as determining

peers to review submitted manuscripts [4], suggesting researchers matching scientific interests [8], finding supervisors for PhD students [9], or recommending helping experts for technical support of medical software [10]. However, we failed to find much work about suggesting expert candidates to invite as members of a renewed conference PC.

The literature features work about finding experts that could respond to an expertise need [4], [11], [12]. Such works take advantage of information retrieval (IR) models by processing documents related to experts to represent experts through textual profiles. Those expert profiles are then matched with the expertise need to find the most similar ones.

Most works build expert profiles from documents related to expert activities in enterprise or academic domains. However, some studies showed an interest in exploiting citations [13], [14] and their combination with other scientific activities for researchers such as participations to conferences [8]. Citations are one of the main subject of Sociology of Science [15], [16], [17]. However, identifying automatically the exact meaning of a citation remains an open issue. To this day, few approaches proposed to use citation networks to find experts, most of them focusing on searching articles or scientific documents.

This paper introduces a model for suggesting PC candidates for a given conference. It exploits scholars’ bibliographies and citation networks to suggest candidates active in the topics of the conference. These suggestions are deemed to inform PC chairs about active albeit perhaps overlooked scholars who would serve the scientific community by joining the conference program committees.

This paper aims at suggesting expert candidates to help renewing the program committee of a given conference. This context differs from those aforementioned by the limited textual content describing a conference. Indeed, the description of a conference is usually limited to a set of topics outlined by some keywords. Searching in a bibliographic database with these keywords would inevitably return a highly “noisy” result. Thus, the approach proposed in this paper is based on non textual elements only.

The contribution of this paper is twofold, by proposing:

- 1) a modelling through a graph of the elements involved in the context of conferences (such as the conference editions, the articles, and the expert-candidates) and their different types of relations (such as being author of an

article, citing an article, and participating to a previous PC),

- 2) a method for evaluating the proximity of an expert candidate with a conference.

A set of experiments in the context of the SIGIR conference editions was carried out to evaluate the effectiveness of our approach.

This paper is organized as follows. The next section introduces the representative prior literature related to PC recommendation and expertise retrieval. Our approach is described in section III and section IV. Section V presents the experiments carried out and the results obtained applying our approach on various collected data about the SIGIR conference editions. Finally, conclusions and future work are given in last section.

II. RELATED WORK

This paper can be related to what is called “expertise retrieval” in the literature [7], [18]. Expertise retrieval gathers two main tasks: expert profiling and expert finding. Expert profiling aims at answering the question: *What topics a person is expert on?*, whereas expert finding aims at answering the question: *Who are the experts on this topic?*

Expert profiling usually intends to construct automatically expert profiles, and in particular topical profiles [19], [20]. Such topical profiles are keyword-based, and built from sources of expertise evidence, which are mainly documents such as enterprise documents [21], curricula vitae (CVs) [22], or research publications [8].

Expert finding intends to identify the persons with a high expertise for a given topic. The majority of the published work focuses on this task thanks to the datasets proposed by the TREC Enterprise track organised during four years [23]. Expert finding can be seen as a ranking task and many approaches took inspiration from information retrieval models [11]. Expert profiles are usually represented as vectors of terms as well as the topics and classical measures, such as the cosine measure, are used to evaluate the similarity between the user profiles and the topics [24].

A key step to address these tasks, is to find suitable sources of expertise evidence. Most approaches consider documents directly linked to expert candidates, such as the documents they authored or the documents mentioning their professional activities. Some works proposed to weight the importance of the documents used benefiting from bibliographic networks [25].

In addition, some approaches that could be related to expertise retrieval showed that other sources of evidence such as citations [13], [14] or co-authorship and participations to conferences [8] are of interest to link researchers to scientific domains. Such approaches take advantage of the works on the building of bibliographic databases, providing resources to build co-authorship networks [26] and citation graphs [27], [28]. However, one problem is that citations convey different meanings, and each citation meaning corresponds to an implicit level of link between the citing article and the cited article [17]. [28] proposed to classify the citations according

to their extracted contexts, while [29] focused on mining the semantics of citation relations applying various NLP (Natural Language Processing) methods.

A citation graph (or citation network) is a directed graph representing documents and their citation relations, i.e., the links between the citing articles and the cited articles. On one hand, such a graph enables one to determine direct and indirect relations between documents. On the other hand, some properties related to the citations, such as the number of citations within a document and the time between the date of the citing article and the date of the cited article, can be used to differentiate the weights of the links between the documents and estimate some proximities between two documents.

Two types of approaches are distinguished in order to estimate the proximities between the nodes of a citation graph. The first type focuses on the direct citations to define the proximity. [15] introduced the concept of bibliographic coupling as two documents citing the same third one and as representing a proximity between the two citing documents. Later, [16] considered co-citation patterns as supporting relationships between documents. Two documents frequently cited together by other ones convey a topical similarity. More recently, [30] introduced a new approach for measuring the similarity between documents based on the proximity of co-citations within an article’s full-text. A comparison of various citation-based similarity measures is presented in [31]. However, these measures are interesting for document retrieval tasks, which do not correspond to the issue tackled in this paper.

The second type of approaches intends to leverage the structure of the graph. In this case, the proximity between two nodes of the graph is determined by the paths connecting them. In expertise retrieval, most of the approaches use the graph measure which was proposed in [32]. The Katz measure is based on the sum of the path weights linking two nodes, giving more weight to short paths. [33] studied several node proximity measures for inferring new interactions among the members of a social networks. The experiments showed good performance of the Katz measure for this purpose. For paper recommendation, [14] considered three types of citation relations and proposed a method using a graph distance based on the Katz distance and weighted links to measure the relevance between two papers. This work is close to our approach in the way the relations between papers are considered. However, the issue we tackle involves other elements and relations (e.g., co-authorship).

Various approaches of expertise retrieval were proposed with different goals, for instance, for finding reviewers for journal submissions [4], [12], for recommending researchers matching some scientific interests [8], finding PhD supervisors for students [9], or finding helping experts for technical support of medical software [10]. Helping to elaborate or renew the program committees of conferences has recently aroused interest. [6] tackles the issue of selecting experts meeting some criteria, such as covering a given set of competences for staffing an organization’s board or for covering the topics of a

conference. This approach builds a collection of textual candidate profiles and propose an IR-based approach to perform the expert group selection. Despite, experimenting the approach in the context of conference PCs, the suggested PCs were not compared with official PCs but evaluated with questionable measures of researcher performance, such as the H-index. In addition, this approach is based on textual descriptions only. [5] studied the possibility of recommending PC members by combining a content-based approach for expert finding and additional indicators related to the publication history of the candidates, their social closeness and their authority. The experiment results showed that the publication history and the social closeness are good indicators while the authority is not. This work gives hints on some indicators suitable for suggestions of PC members, which our approach is also based on.

Since the descriptions of the conference topics are limited, usually quite general, and sometimes in other languages than English, searching for expert candidates in a bibliographic database would inevitably return a highly “noisy” result. Thus, we propose an approach based on non textual elements only, modelled through networks.

Our approach gathers two aspects described in the following sections:

- The modelling of data related to the conferences, the publications, and the researchers;
- The evaluation of the proximity between the expert candidates (i.e., researchers) and the given conference edition concerned by the suggestion of PC members.

III. BIBLIOGRAPHIC GRAPH-BASED DATA MODEL

Considering a given conference for which to suggest candidates to join the PC of the next edition, we distinguish two subsets of data related to the given conference : the data related to the expert candidates (who are directly involved in an edition of the given conference) and the data related to the expert candidates who are involved in the conferences close to the given conference. These two subsets of data are represented by two sub-models that form a final bibliographic model, which are detailed in the next sections.

A. Bibliographic model of the given conference

To suggest committee members, a first issue is to get and to exploit appropriate pieces of evidence about the expert candidates w.r.t. the conference. Firstly, we base our approach on the publications related to the given conference and their authors. Secondly, our approach relies on the past program committees of the given conference, and more particularly their members.

Our model differs from the existing work with regard to different aspects :

- The pieces of evidence are modelled through a 3-mode network;
- Various relationships are modelled between the conference and the expert candidates and are considered of different weights;

- The graph is considered as non-oriented to identify the paths linking the conference to the expert candidates, including the co-citation as well as the co-coupling relations.

We model the bibliographic data as a 3-mode network, in which the nodes represent the concerned conference, the articles related to the conference editions (i.e., the articles published in the editions of the conference, the articles cited by or citing some articles published in the conference), and the expert candidates (i.e., authors and/or past committee members). Four types of edges are distinguished, connecting the different types of nodes: 1) the citation links between two articles (citing-cited), 2) the publication links between the conference editions and the articles, 3) the authoring links between the articles and the researchers, 4) the membership links between the conference editions and the researchers who participated to a program committee of a conference edition. A network example is shown in Figure 1. The relations are represented as oriented to facilitate their comprehension but the orientation is not considered in our approach.

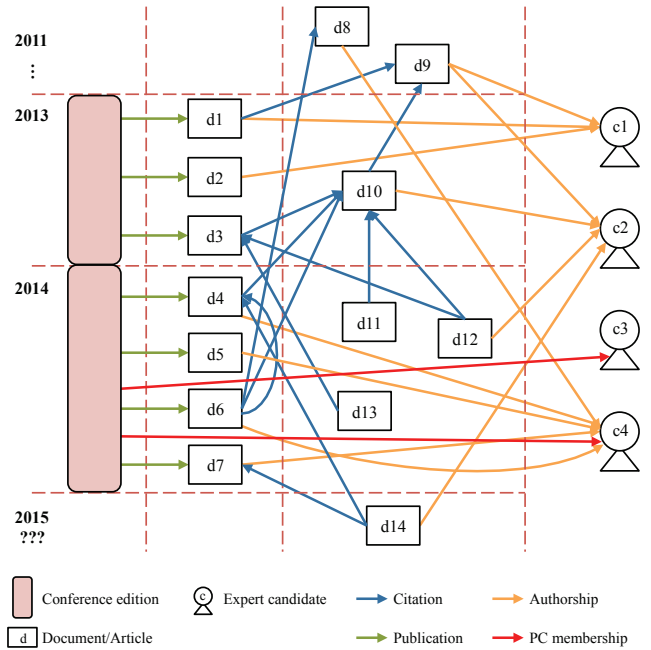


Fig. 1. Example of a 3-mode network modelling bibliographic pieces of evidence about a set of expert candidates and their relations with the given conference.

We consider four types of relations between the expert candidates and the given conference (as illustrated in Figure 2):

- The expert was PC member of an edition of the given conference;
- The expert was author of an article published in an edition of the given conference;
- The expert was author of an article that cites an article published in an edition of the given conference;

- The expert was author of an article that is cited by an article published in an edition of the given conference.

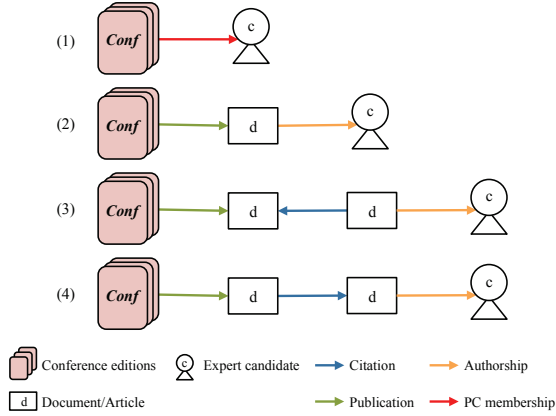


Fig. 2. The relation types between the expert candidates and the given conference.

These types of nodes and edges constitute the base of our model.

B. Bibliographic model of a close conference

Each conference relates to a particular scientific domain. A scientific domain gathers several conferences, which take interest of some common topics. The researchers interested in a particular scientific domain submit papers to the various conferences of this domain, and thus tend to be authors of papers published in these conferences. We call “close conferences” the conferences of a scientific domain that share common topics. Consequently, for a given conference the researchers involved in the close conferences represent expert candidates that could be invited to join the PC.

We follow a similar approach for modelling data about the close conferences, considering the papers published in these conferences, and the researchers concerned by authoring relations as well as PC membership relations. Two types of relations are thus considered: 1) the expert candidate authored a paper of the close conferences, 2) the expert candidate was member of one PC of the close conference.

The pieces of evidence about the given conference (see section III-A) and those about the close conferences are consequently connected through three types of relations as illustrated in Figure 3: 1) an article of the given conference cites or is cited by an article of the close conference, 2) an author of one conference (the given one or the close one) was member of a program committee of the other, and 3) an expert was both member of one edition of the given and of the close conference.

To reduce the complexity introduced by the various types of relations between the given conference and the close conferences, we combine them into a single one as illustrated in Figure 4. We first calculate the weight of the relations between the given conference and the close conferences, then a link between the given conference and each close conference is added to the graph.

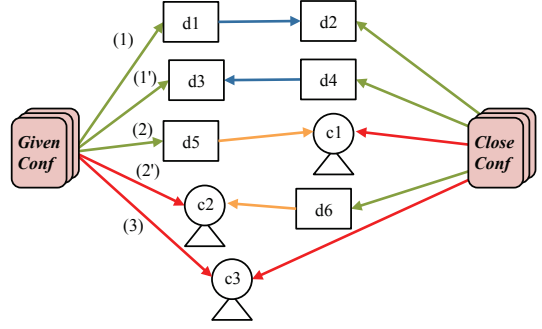


Fig. 3. The types of relations between the given conference and a close conference.

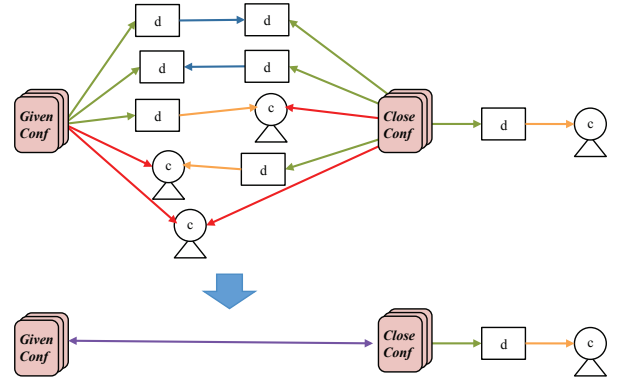


Fig. 4. Relation reduction between close conferences.

1) *Final bibliographic data model*: Eventually, our model corresponds to a 3-mode network with three types of nodes and five types of edges. Figure 5 shows a network example w.r.t. our model. We do not consider the articles that cite or are cited by the articles of the close conferences, and the authors of such articles. This could be an extension of our model.

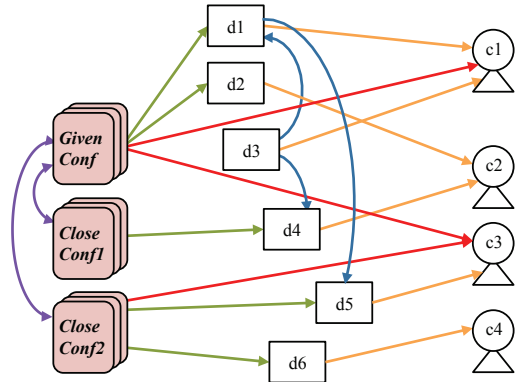


Fig. 5. Example of a 3-mode network modelling the bibliographic pieces of evidence about the expert candidates and their relations with the conference.

IV. EXPERT CANDIDATE PROXIMITY

To evaluate the proximity between an expert candidate and an edition of the given conference, we define different weights

w.r.t the different relations represented in our model, and then deduce the weights of the paths linking the expert candidate to the conference. Finally, the proximity is based on all the paths between an expert candidate and the given conference.

A. Weights of the direct links in the network of the given conference

The weights of the direct links existing in a given conference network defined from the considered data (see section III) as follows:

- $l_{C,d}$ corresponds to the weight of a link between the edition of the conference C at the year t_x and an article d published in the previous conference edition of the year t_d :

$$l_{C,d} = \frac{Q_d}{Q_{max}} \cdot e^{-\frac{t_x - t_d}{\Delta t1_{max}}}$$

Where Q_d is the number of articles that cite the article d , Q_{max} is the maximum number of articles that cite another article, and $\Delta t1_{max}$ is the maximum difference of time between t_x and the publication year of an article. This definition promotes the articles that are very cited and more particularly the most recent ones.

- $v_{citing,cited}$ corresponds to the weight of a link between a *citing* article and a *cited* article:

$$v_{citing,cited} = \frac{Q_{cited}}{Q_{max}} \cdot e^{-\frac{t_{citing} - t_{cited}}{\Delta t2_{max}}}$$

Where Q_{cited} is the number of articles that cite the *cited* article, Q_{max} is the maximum number of articles that cite another article, t_{citing} is the publication year of the *citing* article, t_{cited} is the publication year of the *cited* article, and $\Delta t2_{max}$ is the maximum difference between the publication years of a citing article and a cited article. This definition promotes the citation links with the most cited articles and more particularly the oldest ones.

- $w_{d,c}$ corresponds to the weight of a link between the article d and the author c :

$$w_{d,c} = \frac{1}{A_d}$$

Where A_d is the set of authors of the article d .

This definition considers equal contributions of the co-authors of the article. A variant could also consider the position of each co-author in the list of authors.

- $u_{C,c}$ corresponds to the weight of a link between the edition of the conference C at the year t_x and a member c of the PC of the year t_p :

$$u_{C,c} = e^{-\frac{t_x - t_p}{\Delta t3_{max}}}$$

Where $\Delta t3_{max}$ is the maximum difference between t_x and the PC membership year of an expert candidate.

This definition promotes the expert candidates who participated to the recent editions of the conference.

B. Weights of the links with the close conferences

As aforementioned in section III-B, various relations between the given conference and a close conference are considered. For simplification, these relations are combined into a single link between the given conference and a close conference, and then the links between the expert candidates and this close conference are considered. The weight of the link between the two conferences C_1 and C_2 is based on all the paths between them. The weights of the three path types (see section III-B) are defined as follows:

$$Sth_1 = \frac{l_{C_1,citing} + v_{citing,cited} + l_{C_2,cited}}{l_{max} + v_{max} + l_{max}}$$

$$Sth_2 = \frac{l_{C_1,d} + w_{d,c} + u_{C_2,c}}{l_{max} + w_{max} + u_{max}}$$

$$Sth_3 = \frac{u_{C_1,c} + u_{C_2,c}}{u_{max}}$$

The sum of the weights according to each path type is then computed. For the relation between a close conference and the given conference, we calculate the sum of the weights on the same type. For instance, Sth_1^* corresponds to the sum of path weights in type (1) between the given conference and the close conference.

$$Sth_i^* = \sum_{\theta \in Path_{C_1,C_2}} Sth_i$$

Where θ is a path of type i , $Path_{C_1,C_2}$ is the set of all the paths between the two conferences.

Finally, the weight of the relation between two conferences is defined as:

$$m_{C_1,C_2} = \frac{Sth_1^*}{Sth_{1max}^*} + \frac{Sth_2^*}{Sth_{2max}^*} + \frac{Sth_3^*}{Sth_{3max}^*}$$

C. Paths between the given conference and the expert candidates

To suggest committee members, we compute a proximity score between the new conference edition and each expert candidate. We assume that this proximity depends on:

- the number of paths between the conference node and the expert node. The proximity increases according to this number;
- the weight of the paths between the conference node and the expert node. The weight of a path depends on the weight of each segment composing the path.

The weight of a path between the given conference node and an expert node depends on the six factors corresponding to the possible segments that can be involved in a path. A path comprises from one to three segments depending on the roles played by the expert candidate.

- Role 1: the expert candidate c was member of a past program committee of the given conference C . The path

comprises one segment corresponding to the membership link. Therefore, the weight of the path is defined as:

$$Weight_1 = \frac{u_{C,c}}{u_{max}}$$

- Role 2: the expert candidate c authored the article d published in the given conference C . The path comprises two segments corresponding to the publication link and the authoring link.

$$Weight_2 = \frac{l_{C,d}}{l_{max}} + \frac{w_{d,c}}{w_{max}}$$

- Role 3: the expert candidate c authored the article d' cited by the article d published in the given conference C . The path contains three segments corresponding to the publication link, the citation link, and the authoring link.

$$Weight_3 = \frac{l_{C,d}}{l_{max}} + \frac{v_{d,d'}}{v_{max}} + \frac{w_{d',c}}{w_{max}}$$

- Role 4: the expert candidate c authored the article d' that cites the article d published in the given conference C . The path contains three segments corresponding to the publication link, the citation link, and the authoring link.

$$Weight_4 = \frac{l_{C,d}}{l_{max}} + \frac{v_{d',d}}{v_{max}} + \frac{w_{d',c}}{w_{max}}$$

- Role 5: the expert candidate c was member of a program committee of the close conference C' . The path contains two segments corresponding to the inter-conference link and the membership link.

$$Weight_5 = \frac{m_{C,C'}}{m_{max}} + \frac{u_{C',c}}{u_{max}}$$

- Role 6: the expert candidate c authored the article d' published in the close conference C' . The path contains three segments corresponding to the inter-conference link, the publication link, and the authoring link.

$$Weight_6 = \frac{m_{C,C'}}{m_{max}} + \frac{l_{C',d'}}{l_{max}} + \frac{w_{d',c}}{w_{max}}$$

D. Proximity between the given conference and an expert candidate

For an expert candidate, the sum of the path weights for each path type is computed:

$$Weight_i^* = \sum_{\theta \in Path_{C,c}} Weight_i$$

Where θ is a path of type i and $Path_{C,c}$ is the set of all the paths between the given conference and the expert candidate c .

Finally, the proximity between the given conference C and an expert candidate c is defined as :

$$Proximity_{C,c} = \sum_{i=1}^6 \eta_i \frac{Weight_i^*}{Weight_{i,max}^*}$$

Where η_i are parameters that modulate the importance of the expert roles in the evaluation of the proximity between the given conference and the expert candidate.

V. EXPERIMENTS

A. Dataset

We experimented our approach using data about all the editions of the international conference on research and development in information retrieval (*SIGIR*). This conference is of interest because it is one of the main conferences in the field of information retrieval. It exists since 1971 and comprises a lot of editions, and thus a lot of program committees. Furthermore, it gathers a lot of publications, citing other important publications, as well as being cited by other important publications. Finally, a lot of researchers participated to the different editions as authors of papers and/or as members of the program committees. An additional interest of this conference is the availability of the data related to the conference through the bibliographic databases ACM and DBLP.

The preliminary work consisted in collecting all the papers published in the conference editions, i.e., 3,554 papers of the 40 editions of the *SIGIR* conference from 1971 to 2015. This is the initial set of papers, it is denoted B1 in the remainder of the paper.

The second work consisted in collecting the publications citing the *SIGIR* papers as well as the publications cited by the *SIGIR* papers. This set of 29,907 additional papers is denoted B2.

With regard to the close conferences, we considered eight prominent conferences that share some topics of interest with *SIGIR*: *CIKM*, *WWW*, *ECIR*, *RecSys*, *IJCAI*, *KDD*, *ACL*, and *WSDM*. The set of 35181 additional articles that were published in these conferences is denoted B3.

Finally, all the participations of experts to the various program committees were collected in the available editions, from 2004 for *SIGIR*, *CIKM*, *WWW*, *ECIR*, *KDD*, from 2007 for *RecSys*, from 2003 for *IJCAI*, from 2005 for *ACL*, and from 2010 for *WSDM*. The number of collected editions for each conference depends on the available editions.

B. Results

The objectives of the experiments were to answer two questions:

Q1. To what extent our approach can suggest the members of the official PCs?

Q2. To what extent our approach can suggest relevant new members for the next official PCs?

Q1 is related to the capacity of our approach to identify known PC members. To answer Q1 we compared the PCs suggested by our approach with the official PCs. Q2 is related to the capacity of our approach to identify some potential new members. Since there are no sets of experts identified as potential new *SIGIR* PC members available to serve as ground truth, we compared the new members of the PCs suggested by our approach with the new members of the official PCs to answer Q2. Since our approach aims at helping to renew PCs Q2, answering Q2 is decisive. We evaluated the suggestions according to the usual measures: precision (number of correct

suggestions/number of suggestions) and recall (number of correct suggestions/number of members to suggest).

To present the results of our experiments, we use the notations listed in Table I.

TABLE I
THE NOTATIONS OF THE SETS USED IN THE EXPERIMENTS

Notation	Description
M_y	Set of members of the official program committee of the suggested year y
M_y^-	Set of members of all official program committees of the years before the suggested year y
N_y	Set of <i>new</i> members of the official program committee of the suggested year y . $N_y = M_y \setminus M_y^-$
M_y^{+i}	Set of members of the official program committee of the i -th year after the suggested year y
S_y	Set of experts who are suggested for the year y
N_{S_y}	Set of <i>new</i> experts who are suggested for the year y . $N_{S_y} = S_y \setminus M_y^-$
C_i	Intersection between S_y and a combination of some official committees (or sub-committees) M .

We tested the seven configurations presented in Table II of the parameters involved in the proximity definition defined in section IV.

TABLE II
THE SEVEN CONFIGURATIONS TESTED WITH THE VALUES OF THE PARAMETERS

	η_1	η_2	η_3	η_4	η_5	η_6
Configuration 1	1.0	1.0	1.0	1.0	1.0	1.0
Configuration 2	1.0	0.5	0.5	0.5	0.5	0.5
Configuration 3	0.5	1.0	0.5	0.5	0.5	0.5
Configuration 4	0.5	0.5	1.0	0.5	0.5	0.5
Configuration 5	0.5	0.5	0.5	1.0	0.5	0.5
Configuration 6	0.5	0.5	0.5	0.5	1.0	0.5
Configuration 7	0.5	0.5	0.5	0.5	0.5	1.0

1) *Suggested PCs compared with official PCs*: The first series of experiments consisted in suggesting committees for the last 5 editions, from 2011 to 2015. For each of these considered editions, the sets of papers B2 and B3 were built using the data about the previous SIGIR editions (e.g., bibliographic data from 2004 to 2013 were used to construct the suggestions for the 2014 PC) and the related close conferences.

We performed two types of comparisons for the suggested lists (S_y) of PC members for the year y :

- the comparison with the official program committees (M_y) of the same year y to determine the number of correct suggestions, i.e., $C1 = S_y \cap M_y$;
- the comparison with the set of experts who participated at least to one past program committee (M_y^-) before the year y , i.e., $C2 = S_y \cap M_y^-$.

For each experiment, we suggested three different lists: a suggested list comprising the same number of expert candidates as the corresponding official PC (*Nb of experts* = $|M_y|$); a list with 50% of expert candidates more than the corresponding official PC (*Nb of experts* = 150% $|M_y|$) and a list with 100% of expert candidates more than the corresponding official PC (*Nb of experts* = 200% $|M_y|$).

Table III reports the comparison results (according to precision and recall) when applying the seven configurations of our approach, compared with the official PC of the corresponding year y ($C1$).

The results show that our approach suggests a PC that shares from about 40% to 53% of official PC members. Up to more than 67% can be suggested when doubling the suggestions, but to the detriment of precision. The best results were reached by the configuration 2 of our approach, which promotes the participations of expert candidates to the past PCs of the conference. The lowest results are obtained with configurations 5 and 6, which promote the participations to PC of close conferences and the publications in close conferences, respectively. The results shown in in Table IV, which compare the suggestions with the whole set of past PCs (M_y^-). A recall up to 96% can be reached by the configuration 2 when doubling the size of the suggested list (Table IV, $|S_y| = 200\%|M_y^-|$). Since the configurations 5 and 6 are those identifying less existing PC members, they are likely to suggest more candidates that could be relevant new PC members. This is the concern of the experiments reported in the next section.

2) *New members suggested compared with the new members of the official PCs*: There are annually some new members who join the PC of a conference for the first time. Table V shows that there is a not negligible proportion of new members each year in the PC of the conferences related to our experiments. This observation supports the interest of our approach, helping to find expert candidates to invite for joining a new PC. To evaluate the capacity of suggesting relevant expert candidates applying our approach, we should compare our suggested new candidates with a set of identified relevant candidates (representing a ground truth) or submit our suggestions to expert assessors. Since there are no sets of experts identified as potential new SIGIR PC members for a given year y available to serve as ground truth, we compared the new members of the PCs suggested by our approach with the new members of the official PCs of the years after y to answer Q2. The precision and recall measures were still used. These experiments have thus some limitations because the new SIGIR PC members appearing in the next conference editions represent an incomplete ground truth. There is a need for a more complete ground truth, which should not depend on the current practice of constituting PCs. However, finding experts to suggest expert candidates or to assess suggestions is a difficult task. This constitutes an important future work.

For these experiments, as the first series of experiments run, we suggested PCs for the editions from 2011 to 2015. For each suggested PC for a given year y we considered the new suggested members N_{S_y} , i.e., the members who were not present in one of the official PCs before the year y . We considered three different cases:

- 1) The set of new suggested experts who are actually new members in the corresponding official PC, i.e.,

$$C3 = N_{S_y} \cap (M_y \setminus M_y^-)$$

TABLE III
COMPARISON OF SUGGESTED PCs WITH THE CORRESPONDING OFFICIAL PC.

Year	My	Sy = My			Sy = 150% My			Sy = 200% My		
		C1	Recall	Precision	C1	Recall	Precision	C1	Recall	Precision
<i>Configuration1</i> : $\eta_1 = \eta_2 = \eta_3 = \eta_4 = \eta_5 = \eta_6 = 1$										
2011	426	197	46.24 %	46.24 %	240	56.34 %	37.56 %	261	61.27 %	30.63 %
2012	486	218	44.86 %	44.86 %	255	52.47 %	34.98 %	281	57.82 %	28.91 %
2013	431	201	46.64 %	46.64 %	237	54.99 %	36.69 %	266	61.72 %	30.86 %
2014	448	200	44.64 %	44.64 %	236	52.68 %	35.12 %	266	59.38 %	29.69 %
2015	432	192	44.44 %	44.44 %	233	53.94 %	35.96 %	260	60.19 %	30.09 %
<i>Configuration2</i> : $\eta_1 = 1; \eta_2 = 0.5; \eta_3 = 0.5; \eta_4 = 0.5; \eta_5 = 0.5; \eta_6 = 0.5$										
2011	426	226	53.05 %	53.05 %	266	62.44 %	41.63 %	288	67.61 %	33.80 %
2012	486	243	50.00 %	50.00 %	273	56.17 %	37.45 %	300	61.73 %	30.86 %
2013	431	223	51.74 %	51.74 %	259	60.09 %	40.06 %	288	66.82 %	33.41 %
2014	448	217	48.44 %	48.44 %	260	58.04 %	38.69 %	300	66.96 %	33.48 %
2015	432	214	49.54 %	49.54 %	253	58.56 %	39.04 %	278	64.35 %	32.18 %
<i>Configuration3</i> : $\eta_1 = 0.5; \eta_2 = 1; \eta_3 = 0.5; \eta_4 = 0.5; \eta_5 = 0.5; \eta_6 = 0.5$										
2011	426	202	47.42 %	47.42 %	242	56.81 %	37.87 %	268	62.91 %	31.46 %
2012	486	222	45.68 %	45.68 %	259	53.29 %	35.53 %	286	58.85 %	29.42 %
2013	431	202	46.87 %	46.87 %	244	56.61 %	37.77 %	269	62.41 %	31.21 %
2014	448	203	45.31 %	45.31 %	240	53.57 %	35.71 %	270	60.27 %	30.13 %
2015	432	195	45.14 %	45.14 %	237	54.86 %	36.57 %	265	61.34 %	30.67 %
<i>Configuration4</i> : $\eta_1 = 0.5; \eta_2 = 0.5; \eta_3 = 1; \eta_4 = 0.5; \eta_5 = 0.5; \eta_6 = 0.5$										
2011	426	196	46.01 %	46.01 %	239	56.10 %	37.40 %	262	61.50 %	30.75 %
2012	486	218	44.86 %	44.86 %	255	52.47 %	34.98 %	281	57.82 %	28.91 %
2013	431	201	46.64 %	46.64 %	239	55.45 %	37.00 %	266	61.72 %	30.86 %
2014	448	200	44.64 %	44.64 %	235	52.46 %	34.97 %	264	58.93 %	29.46 %
2015	432	191	44.21 %	44.21 %	232	53.70 %	35.80 %	259	59.95 %	29.98 %
<i>Configuration5</i> : $\eta_1 = 0.5; \eta_2 = 0.5; \eta_3 = 0.5; \eta_4 = 1; \eta_5 = 0.5; \eta_6 = 0.5$										
2011	426	197	46.24 %	46.24 %	238	55.87 %	37.25 %	262	61.50 %	30.75 %
2012	486	219	45.06 %	45.06 %	256	52.67 %	35.12 %	283	58.23 %	29.12 %
2013	431	204	47.33 %	47.33 %	239	55.45 %	37.00 %	264	61.25 %	30.63 %
2014	448	199	44.42 %	44.42 %	235	52.46 %	34.97 %	265	59.15 %	29.58 %
2015	432	192	44.44 %	44.44 %	232	53.70 %	35.80 %	258	59.72 %	29.86 %
<i>Configuration6</i> : $\eta_1 = 0.5; \eta_2 = 0.5; \eta_3 = 0.5; \eta_4 = 0.5; \eta_5 = 1; \eta_6 = 0.5$										
2011	426	182	42.72 %	42.72 %	218	51.17 %	34.12 %	240	56.34 %	28.17 %
2012	486	195	40.12 %	40.12 %	236	48.56 %	32.37 %	258	53.09 %	26.54 %
2013	431	186	43.16 %	43.16 %	221	51.28 %	34.21 %	246	57.08 %	28.54 %
2014	448	189	42.19 %	42.19 %	219	48.88 %	32.59 %	247	55.13 %	27.57 %
2015	432	183	42.36 %	42.36 %	218	50.46 %	33.64 %	244	56.48 %	28.24 %
<i>Configuration7</i> : $\eta_1 = 0.5; \eta_2 = 0.5; \eta_3 = 0.5; \eta_4 = 0.5; \eta_5 = 0.5; \eta_6 = 1$										
2011	426	173	40.61 %	40.61 %	215	50.47 %	33.65 %	243	57.04 %	28.52 %
2012	486	194	39.92 %	39.92 %	236	48.56 %	32.37 %	262	53.91 %	26.95 %
2013	431	187	43.39 %	43.39 %	221	51.28 %	34.21 %	243	56.38 %	28.19 %
2014	448	182	40.63 %	40.63 %	216	48.21 %	32.14 %	247	55.13 %	27.57 %
2015	432	182	42.13 %	42.13 %	216	50.00 %	33.33 %	243	56.25 %	28.13 %

For instance, an expert who is suggested for 2011 is actually participating to the official PC in 2011 as a new member.

- 2) The set of new suggested experts who are actually new members in the corresponding official PC or the next one, i.e.,

$$C4 = N_{Sy} \cap ((My^{+1} \cup My) \setminus My^{-})$$

For instance, an expert who is suggested for 2011, is actually joining the official PC as new member in 2011 or in 2012.

- 3) The set of new suggested experts who are actually new members of the corresponding official PC or the next two ones, i.e.,

$$C5 = N_{Sy} \cap ((My^{+1} \cup My^{+2} \cup My) \setminus My^{-})$$

For instance, an expert who is suggested for 2011 is

actually joining the official PC as new member in 2011, in 2012, or in 2013.

The results for the first case, i.e., the comparison with the new members of the corresponding PC (C3), reported in Table VI show that a low proportion of suggested members corresponds to official new members of the corresponding year. However, the results show that the configuration of our approach that promotes the participations to PCs of close conferences (Configuration 6) yields the best results on average w.r.t. the other configurations. Doubling the suggestions can reach a higher recall but lowers precision.

The cases 2) and 3) were intended to verify if a delay may exist for an expert candidate to be invited to join the PC of a conference. The results reported in Table VII for the case 2) and in Table VIII for the case 3) show that more new suggested expert candidates become relevant since they join the official PCs in the next years. These results confirm the

TABLE IV
COMPARISON OF EXPERTS RETURNED BY OUR SYSTEM WITH THE PAST PROGRAM COMMITTEES.

Year	$ M_y^- $	$ S_y = M_y^- $			$ S_y = 150\% M_y^- $			$ S_y = 200\% M_y^- $		
		C2	Recall	Precision	C2	Recall	Precision	C2	Recall	Precision
<i>Configuration1</i> : $\eta_1 = \eta_2 = \eta_3 = \eta_4 = \eta_5 = \eta_6 = 1$										
2011	1386	683	49.28 %	49.28 %	832	60.03 %	40.02 %	1077	77.71 %	25.90 %
2012	1482	720	48.58 %	48.58 %	879	59.31 %	39.54 %	1143	77.13 %	25.71 %
2013	1537	785	51.07 %	51.07 %	948	61.68 %	41.13 %	1238	80.55 %	26.85 %
2014	1624	806	49.63 %	49.63 %	982	60.47 %	40.31 %	1263	77.77 %	25.92 %
2015	1706	855	50.12 %	50.12 %	1038	60.84 %	40.56 %	1340	78.55 %	26.18 %
<i>Configuration2</i> : $\eta_1 = 1; \eta_2 = 0.5; \eta_3 = 0.5; \eta_4 = 0.5; \eta_5 = 0.5; \eta_6 = 0.5$										
2011	1386	812	58.59 %	58.59 %	992	71.57 %	47.72 %	1294	93.36 %	31.12 %
2012	1482	851	57.42 %	57.42 %	1052	70.99 %	47.32 %	1386	93.52 %	31.17 %
2013	1537	927	60.31 %	60.31 %	1134	73.78 %	49.20 %	1479	96.23 %	32.08 %
2014	1624	962	59.24 %	59.24 %	1154	71.06 %	47.37 %	1465	90.21 %	30.07 %
2015	1706	1010	59.20 %	59.20 %	1220	71.51 %	47.67 %	1569	91.97 %	30.66 %
<i>Configuration3</i> : $\eta_1 = 0.5; \eta_2 = 1; \eta_3 = 0.5; \eta_4 = 0.5; \eta_5 = 0.5; \eta_6 = 0.5$										
2011	1386	694	50.07 %	50.07 %	835	60.25 %	40.16 %	1078	77.78 %	25.93 %
2012	1482	733	49.46 %	49.46 %	893	60.26 %	40.17 %	1158	78.14 %	26.05 %
2013	1537	801	52.11 %	52.11 %	964	62.72 %	41.82 %	1245	81.00 %	27.00 %
2014	1624	830	51.11 %	51.11 %	997	61.39 %	40.93 %	1266	77.96 %	25.99 %
2015	1706	878	51.47 %	51.47 %	1051	61.61 %	41.07 %	1350	79.13 %	26.38 %
<i>Configuration4</i> : $\eta_1 = 0.5; \eta_2 = 0.5; \eta_3 = 1; \eta_4 = 0.5; \eta_5 = 0.5; \eta_6 = 0.5$										
2011	1386	684	49.35 %	49.35 %	832	60.03 %	40.02 %	1072	77.34 %	25.78 %
2012	1482	724	48.85 %	48.85 %	877	59.18 %	39.45 %	1145	77.26 %	25.75 %
2013	1537	790	51.40 %	51.40 %	948	61.68 %	41.13 %	1236	80.42 %	26.81 %
2014	1624	815	50.18 %	50.18 %	985	60.65 %	40.44 %	1256	77.34 %	25.78 %
2015	1706	858	50.29 %	50.29 %	1039	60.90 %	40.60 %	1338	78.43 %	26.14 %
<i>Configuration5</i> : $\eta_1 = 0.5; \eta_2 = 0.5; \eta_3 = 0.5; \eta_4 = 1; \eta_5 = 0.5; \eta_6 = 0.5$										
2011	1386	681	49.13 %	49.13 %	833	60.10 %	40.07 %	1074	77.49 %	25.83 %
2012	1482	723	48.79 %	48.79 %	882	59.51 %	39.68 %	1145	77.26 %	25.75 %
2013	1537	792	51.53 %	51.53 %	949	61.74 %	41.17 %	1236	80.42 %	26.81 %
2014	1624	812	50.00 %	50.00 %	984	60.59 %	40.39 %	1261	77.65 %	25.88 %
2015	1706	858	50.29 %	50.29 %	1040	60.96 %	40.64 %	1339	78.49 %	26.16 %
<i>Configuration6</i> : $\eta_1 = 0.5; \eta_2 = 0.5; \eta_3 = 0.5; \eta_4 = 0.5; \eta_5 = 1; \eta_6 = 0.5$										
2011	1386	611	44.08 %	44.08 %	749	54.04 %	36.03 %	983	70.92 %	23.64 %
2012	1482	646	43.59 %	43.59 %	795	53.64 %	35.76 %	1041	70.24 %	23.41 %
2013	1537	714	46.45 %	46.45 %	850	55.30 %	36.88 %	1123	73.06 %	24.35 %
2014	1624	726	44.70 %	44.70 %	882	54.31 %	36.21 %	1141	70.26 %	23.42 %
2015	1706	780	45.72 %	45.72 %	927	54.34 %	36.23 %	1221	71.57 %	23.86 %
<i>Configuration7</i> : $\eta_1 = 0.5; \eta_2 = 0.5; \eta_3 = 0.5; \eta_4 = 0.5; \eta_5 = 0.5; \eta_6 = 1$										
2011	1386	602	43.43 %	43.43 %	741	53.46 %	35.64 %	977	70.49 %	23.50 %
2012	1482	638	43.05 %	43.05 %	787	53.10 %	35.40 %	1057	71.32 %	23.77 %
2013	1537	699	45.48 %	45.48 %	863	56.15 %	37.44 %	1130	73.52 %	24.51 %
2014	1624	735	45.26 %	45.26 %	896	55.17 %	36.78 %	1154	71.06 %	23.69 %
2015	1706	782	45.84 %	45.84 %	946	55.45 %	36.97 %	1234	72.33 %	24.11 %

TABLE V
PROPORTION OF MEMBERS WHO PARTICIPATED FOR THE FIRST TIME TO A GIVEN YEAR PC

Year	Conf.	$ M_y $	$ N_y $	Ratio	Conf.	$ M_y $	$ N_y $	Ratio
2011	SIGIR	426	92	21.60 %	WWW	570	320	56.14 %
2012		486	96	19.75 %		547	241	44.06 %
2013		431	55	12.76 %		955	521	54.55 %
2014		448	87	19.42 %		941	455	48.35 %
2015		432	82	18.98 %		502	173	34.46 %
2011	CIKM	697	295	42.32 %	KDD	418	117	27.99 %
2012		716	407	56.84 %		677	341	50.37 %
2013		515	131	25.44 %		696	306	43.97 %
2014		629	353	56.12 %		784	281	35.84 %
2015		700	226	32.29 %		794	274	34.51 %
2011	ECIR	229	68	29.69 %	ACL	789	291	36.88 %
2012		250	94	37.60 %		649	154	23.73 %
2013		233	22	9.44 %		748	232	31.02 %
2014		268	57	21.27 %		874	269	30.78 %
2015		311	88	28.30 %		885	228	25.76 %

influence of the participations to PCs of the close conference

in the invitation of new members. However, these results do not inform us about the relevance of the other suggested new experts and motivate us to carry out experiments involving the assessment of such suggestions by SIGIR experts to get more accurate evaluations. Our future work will be devoted to this.

VI. CONCLUSION AND FUTURE WORK

We introduced an approach for helping program chairs of conferences to find expert candidates who could be invited to join the next conference PC. The proposed approach exploits bibliographic data to build a 3-mode network modelling the pieces of evidence of expert candidates w.r.t a given conference. Such pieces of evidence are related to the various relations between the experts and the given conference, such as authoring, PC membership, and citation, as well as the close conferences. A measure to evaluate the proximity between a conference edition and an expert candidate is proposed. It is

TABLE VI
NEW MEMBERS OF PROGRAM COMMITTEE IN THE CORRESPONDING YEAR WHO ARE SUGGESTED

Year	N _y	S _y = M _y				S _y = 150% M _y				S _y = 200% M _y			
		C3	Recall	N _{S_y}	Precision	C3	Recall	N _{S_y}	Precision	C3	Recall	N _{S_y}	Precision
<i>Configuration1</i> : $\eta_1 = \eta_2 = \eta_3 = \eta_4 = \eta_5 = \eta_6 = 1$													
2011	92	0	0.00 %	102	0.00 %	3	3.26 %	199	1.51 %	6	6.52 %	337	1.78 %
2012	96	5	5.21 %	114	4.39 %	7	7.29 %	239	2.93 %	10	10.42 %	391	2.56 %
2013	55	2	3.64 %	82	2.44 %	2	3.64 %	182	1.10 %	3	5.45 %	292	1.03 %
2014	87	0	0.00 %	88	0.00 %	2	2.30 %	198	1.01 %	4	4.60 %	312	1.28 %
2015	82	0	0.00 %	75	0.00 %	1	1.22 %	165	0.61 %	3	3.66 %	274	1.09 %
<i>Configuration2</i> : $\eta_1 = 1; \eta_2 = 0.5; \eta_3 = 0.5; \eta_4 = 0.5; \eta_5 = 0.5; \eta_6 = 0.5$													
2011	92	0	0.00 %	49	0.00 %	1	1.09 %	122	0.82 %	3	3.26 %	233	1.29 %
2012	96	3	3.13 %	64	4.69 %	5	5.21 %	156	3.21 %	8	8.33 %	286	2.80 %
2013	55	0	0.00 %	33	0.00 %	2	3.64 %	107	1.87 %	2	3.64 %	202	0.99 %
2014	87	0	0.00 %	37	0.00 %	0	0.00 %	113	0.00 %	2	2.30 %	218	0.92 %
2015	82	0	0.00 %	28	0.00 %	0	0.00 %	90	0.00 %	2	2.44 %	183	1.09 %
<i>Configuration3</i> : $\eta_1 = 0.5; \eta_2 = 1; \eta_3 = 0.5; \eta_4 = 0.5; \eta_5 = 0.5; \eta_6 = 0.5$													
2011	92	1	1.09 %	98	1.02 %	4	4.35 %	194	2.06 %	6	6.52 %	321	1.87 %
2012	96	6	6.25 %	113	5.31 %	8	8.33 %	234	3.42 %	10	10.42 %	378	2.65 %
2013	55	2	3.64 %	78	2.56 %	3	5.45 %	171	1.75 %	3	5.45 %	280	1.07 %
2014	87	0	0.00 %	82	0.00 %	2	2.30 %	184	1.09 %	5	5.75 %	305	1.64 %
2015	82	0	0.00 %	70	0.00 %	1	1.22 %	155	0.65 %	5	6.10 %	264	1.89 %
<i>Configuration4</i> : $\eta_1 = 0.5; \eta_2 = 0.5; \eta_3 = 1; \eta_4 = 0.5; \eta_5 = 0.5; \eta_6 = 0.5$													
2011	92	0	0.00 %	103	0.00 %	3	3.26 %	200	1.50 %	6	6.52 %	333	1.80 %
2012	96	5	5.21 %	117	4.27 %	7	7.29 %	239	2.93 %	10	10.42 %	385	2.60 %
2013	55	2	3.64 %	82	2.44 %	2	3.64 %	181	1.10 %	3	5.45 %	290	1.03 %
2014	87	0	0.00 %	87	0.00 %	2	2.30 %	195	1.03 %	4	4.60 %	308	1.30 %
2015	82	0	0.00 %	74	0.00 %	1	1.22 %	164	0.61 %	3	3.66 %	271	1.11 %
<i>Configuration5</i> : $\eta_1 = 0.5; \eta_2 = 0.5; \eta_3 = 0.5; \eta_4 = 1; \eta_5 = 0.5; \eta_6 = 0.5$													
2011	92	1	1.09 %	103	0.97 %	3	3.26 %	199	1.51 %	6	6.52 %	331	1.81 %
2012	96	5	5.21 %	116	4.31 %	8	8.33 %	236	3.39 %	10	10.42 %	387	2.58 %
2013	55	2	3.64 %	79	2.53 %	2	3.64 %	179	1.12 %	3	5.45 %	291	1.03 %
2014	87	0	0.00 %	86	0.00 %	2	2.30 %	194	1.03 %	4	4.60 %	310	1.29 %
2015	82	0	0.00 %	72	0.00 %	1	1.22 %	164	0.61 %	3	3.66 %	272	1.10 %
<i>Configuration6</i> : $\eta_1 = 0.5; \eta_2 = 0.5; \eta_3 = 0.5; \eta_4 = 0.5; \eta_5 = 1; \eta_6 = 0.5$													
2011	92	3	3.26 %	133	2.26 %	4	4.35 %	252	1.59 %	5	5.43 %	399	1.25 %
2012	96	6	6.25 %	147	4.08 %	9	9.38 %	291	3.09 %	10	10.42 %	467	2.14 %
2013	55	2	3.64 %	111	1.80 %	3	5.45 %	223	1.35 %	4	7.27 %	355	1.13 %
2014	87	1	1.15 %	109	0.92 %	3	3.45 %	230	1.30 %	3	3.45 %	367	0.82 %
2015	82	1	1.22 %	96	1.04 %	2	2.44 %	198	1.01 %	5	6.10 %	322	1.55 %
<i>Configuration7</i> : $\eta_1 = 0.5; \eta_2 = 0.5; \eta_3 = 0.5; \eta_4 = 0.5; \eta_5 = 0.5; \eta_6 = 1$													
2011	92	0	0.00 %	146	0.00 %	2	2.17 %	255	0.78 %	5	5.43 %	393	1.27 %
2012	96	5	5.21 %	158	3.16 %	7	7.29 %	299	2.34 %	8	8.33 %	454	1.76 %
2013	55	2	3.64 %	115	1.74 %	2	3.64 %	234	0.85 %	2	3.64 %	360	0.56 %
2014	87	1	1.15 %	124	0.81 %	2	2.30 %	247	0.81 %	4	4.60 %	380	1.05 %
2015	82	0	0.00 %	105	0.00 %	2	2.44 %	217	0.92 %	3	3.66 %	331	0.91 %

based on the definition of weights for the various types of relations that connect the expert candidates to the conference editions.

Different configurations of our approach can be obtained by modulating the importance assigned to the different types of relations. These configurations yield PC suggestions composed quasi-entirely of members from past PCs, or PC suggestions mainly composed of “new” candidates, passing by more balanced suggestions.

Future work will be devoted to:

- enhance the evaluation protocol to get more accurate evaluations, by involving some assessments of the suggestions, for example asking some experts involved in the studied conferences, such as the previous PC chairs;
- test different configurations of our approach to evaluate the impact of each considered relation;
- experiment our approach on other conferences, to detect

the differences which could exist in the PC compositions between some close conferences, and between the different PC chairs;

- refine our approach by considering different types of citations or incorporating information about the conference topics;
- propose a multi-criteria approach to build a program committee and then apply the appropriate configuration of our approach to fit the expected criteria.

REFERENCES

- [1] H. Zuckerman and R. K. Merton, “Patterns of evaluation in science: Institutionalisation, structure and functions of the referee system,” *Minerva*, vol. 9, no. 1, pp. 66–100, 1971.
- [2] D. Crane, “The gatekeepers of science: Some factors affecting the selection of articles for scientific journals,” *The American Sociologist*, vol. 2, no. 4, pp. 195–201, 1967.

TABLE VII
THE SUGGESTED EXPERTS WHO ARE NEW MEMBERS OF PROGRAM COMMITTEE IN THE NEXT YEAR

Year	$ N_{y \cup y+1} $	$ S_y = M_y \cup M_{y+1} $				$ S_y = 150\% M_y \cup M_{y+1} $				$ S_y = 200\% M_y \cup M_{y+1} $			
		C4	Recall	N_{S_y}	Precision	C4	Recall	N_{S_y}	Precision	C4	Recall	N_{S_y}	Precision
<i>Configuration 1</i> : $\eta_1 = \eta_2 = \eta_3 = \eta_4 = \eta_5 = \eta_6 = 1$													
2011	188	8	4.26 %	176	4.55 %	13	6.91 %	363	3.58 %	17	9.04 %	566	3.00 %
2012	151	8	5.30 %	164	4.88 %	12	7.95 %	335	3.58 %	17	11.26 %	542	3.14 %
2013	142	2	1.41 %	152	1.32 %	6	4.23 %	298	2.01 %	10	7.04 %	499	2.00 %
2014	169	3	1.78 %	146	2.05 %	7	4.14 %	296	2.36 %	11	6.51 %	467	2.36 %
<i>Configuration 2</i> : $\eta_1 = 1; \eta_2 = 0.5; \eta_3 = 0.5; \eta_4 = 0.5; \eta_5 = 0.5; \eta_6 = 0.5$													
2011	188	4	2.13 %	109	3.67 %	12	6.38 %	257	4.67 %	15	7.98 %	439	3.42 %
2012	151	6	3.97 %	99	6.06 %	9	5.96 %	236	3.81 %	14	9.27 %	420	3.33 %
2013	142	2	1.41 %	82	2.44 %	3	2.11 %	209	1.44 %	9	6.34 %	368	2.45 %
2014	169	0	0.00 %	80	0.00 %	4	2.37 %	200	2.00 %	8	4.73 %	346	2.31 %
<i>Configuration 3</i> : $\eta_1 = 0.5; \eta_2 = 1; \eta_3 = 0.5; \eta_4 = 0.5; \eta_5 = 0.5; \eta_6 = 0.5$													
2011	188	9	4.79 %	171	5.26 %	14	7.45 %	350	4.00 %	19	10.11 %	550	3.45 %
2012	151	8	5.30 %	159	5.03 %	13	8.61 %	317	4.10 %	19	12.58 %	527	3.61 %
2013	142	3	2.11 %	140	2.14 %	8	5.63 %	288	2.78 %	13	9.15 %	485	2.68 %
2014	169	1	0.59 %	133	0.75 %	8	4.73 %	285	2.81 %	12	7.10 %	455	2.64 %
<i>Configuration 4</i> : $\eta_1 = 0.5; \eta_2 = 0.5; \eta_3 = 1; \eta_4 = 0.5; \eta_5 = 0.5; \eta_6 = 0.5$													
2011	188	8	4.26 %	177	4.52 %	13	6.91 %	360	3.61 %	17	9.04 %	562	3.02 %
2012	151	8	5.30 %	162	4.94 %	12	7.95 %	335	3.58 %	17	11.26 %	536	3.17 %
2013	142	2	1.41 %	149	1.34 %	7	4.93 %	299	2.34 %	11	7.75 %	497	2.21 %
2014	169	3	1.78 %	141	2.13 %	7	4.14 %	294	2.38 %	11	6.51 %	464	2.37 %
<i>Configuration 5</i> : $\eta_1 = 0.5; \eta_2 = 0.5; \eta_3 = 0.5; \eta_4 = 1; \eta_5 = 0.5; \eta_6 = 0.5$													
2011	188	8	4.26 %	178	4.49 %	14	7.45 %	361	3.88 %	17	9.04 %	563	3.02 %
2012	151	7	4.64 %	165	4.24 %	12	7.95 %	333	3.60 %	17	11.26 %	537	3.17 %
2013	142	2	1.41 %	150	1.33 %	7	4.93 %	297	2.36 %	10	7.04 %	499	2.00 %
2014	169	3	1.78 %	143	2.10 %	7	4.14 %	296	2.36 %	11	6.51 %	464	2.37 %
<i>Configuration 6</i> : $\eta_1 = 0.5; \eta_2 = 0.5; \eta_3 = 0.5; \eta_4 = 0.5; \eta_5 = 1; \eta_6 = 0.5$													
2011	188	10	5.32 %	222	4.50 %	14	7.45 %	431	3.25 %	19	10.11 %	643	2.95 %
2012	151	11	7.28 %	204	5.39 %	14	9.27 %	399	3.51 %	18	11.92 %	609	2.96 %
2013	142	3	2.11 %	179	1.68 %	8	5.63 %	364	2.20 %	11	7.75 %	576	1.91 %
2014	169	3	1.78 %	175	1.71 %	8	4.73 %	346	2.31 %	13	7.69 %	538	2.42 %
<i>Configuration 7</i> : $\eta_1 = 0.5; \eta_2 = 0.5; \eta_3 = 0.5; \eta_4 = 0.5; \eta_5 = 0.5; \eta_6 = 1$													
2011	188	6	3.19 %	231	2.60 %	13	6.91 %	419	3.10 %	16	8.51 %	630	2.54 %
2012	151	8	5.30 %	213	3.76 %	10	6.62 %	393	2.54 %	14	9.27 %	601	2.33 %
2013	142	3	2.11 %	202	1.49 %	4	2.82 %	368	1.09 %	9	6.34 %	571	1.58 %
2014	169	3	1.78 %	199	1.51 %	7	4.14 %	360	1.94 %	8	4.73 %	548	1.46 %

- [3] G. Cabanac and T. Preuss, "Capitalizing on order effects in the bids of peer-reviewed conferences to secure reviews by expert referees," *Journal of the American Society for Information Science and Technology*, vol. 64, no. 2, pp. 405–415, 2013.
- [4] M. Rodriguez and J. Bollen, "An algorithm to determine peer-reviewers," in *17th ACM conference on Information and knowledge management (CIKM'08)*, 2008, pp. 319–328.
- [5] S. Han, J. Jiang, Z. Yue, and D. He, "Recommending program committee candidates for academic conferences," in *Proceedings of the 2013 Workshop on Computational Scientometrics: Theory and Applications (CompSci'13)*. New York, NY, USA: ACM, 2013, pp. 1–6.
- [6] G. A. Sfyris, N. Fragkos, and C. Doukeridis, "Profile-based selection of expert groups," in *Proceedings of the 20th International Conference on Theory and Practice of Digital Libraries, (TPDL 2016)*. Springer International Publishing, 2016, pp. 81–93.
- [7] K. Balog, Y. Fang, M. de Rijke, P. Serdyukov, and L. Si, "Expertise retrieval," *Foundations and Trends in Information Retrieval*, vol. 6, no. 2–3, pp. 127–256, 2012.
- [8] G. Cabanac, "Accuracy of inter-researcher similarity measures based on topical and social clues," *Scientometrics*, vol. 87, no. 3, pp. 597–620, 2011.
- [9] P. Liu and P. Dew, "Using semantic web technologies to improve expertise matching within academia," in *International conference I-KNOW'04*, 2004, pp. 370–378.
- [10] D. W. McDonald and M. S. Ackerman, "Expertise recommender: A flexible recommendation system and architecture," in *ACM conference on Computer supported cooperative work (CSCW'00)*, 2000, pp. 231–240.
- [11] H. Fang and C. Zhai, "Probabilistic Models for Expert Finding," in *29th European conference on IR research (ECIR'07)*, 2007, pp. 418–430.
- [12] M. Afzal and H. Maurer, "Expertise Recommender System for Scientific Community," *Journal of Universal Computer Science*, vol. 17, no. 11, pp. 1529–1549, 2011.
- [13] T. Strohman, W. Croft, and D. Jensen, "Recommending citations for academic papers," in *30th annual international ACM SIGIR conference on Research and development in information retrieval (SIGIR '07)*, 2007, pp. 705–706.
- [14] Y. Liang, Q. Li, and T. Qian, "Finding relevant papers based on citation relations," *Web-Age Information Management*, vol. 6897, pp. 403–414, 2011.
- [15] M. Kessler, "Bibliographic coupling between scientific papers," *American documentation*, vol. 14, no. 1, pp. 10–25, 1963.
- [16] H. Small, "Co-citation in the scientific literature: A new measure of the relationship between two documents," *Journal of the American society for information science*, vol. 24, pp. 265–269, 1973.
- [17] B. Cronin, *The Citation Process: The Role and Significance of Citations in Scientific Communication*. London: Taylor Graham, 1984.
- [18] V. Mangaravite, R. L. Santos, I. S. Ribeiro, M. A. Gonçalves, and A. H. Laender, "The LEXR Collection for Expertise Retrieval in Academia," in *Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'16)*. New York, NY, USA: ACM, 2016, pp. 721–724.
- [19] K. Balog and M. de Rijke, "Determining expert profiles (with an application to expert finding)," in *Proceedings of the 20th International Joint Conference on Artificial Intelligence (IJCAI'07)*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2007, pp. 2657–2662.
- [20] J. Rybak, K. Balog, and K. Nørsvåg, "Temporal expertise profiling," in *Proceedings of the 36th European Conference on IR Research (ECIR 2014)*. Springer International Publishing, 2014, pp. 540–546.

TABLE VIII
THE SUGGESTED EXPERTS WHO ARE NEW MEMBERS OF PROGRAM COMMITTEES FOR THE NEXT TWO YEARS

Year	$ N_y \cup N_{y+1} \cup N_{y+2} $	$ S_y =$ $ M_y \cup M_{y+1} \cup M_{y+2} $				$ S_y = 150\%$ $ M_y \cup M_{y+1} \cup M_{y+2} $				$ S_y = 200\%$ $ M_y \cup M_{y+1} \cup M_{y+2} $			
		C5	Recall	N_{S_y}	Precision	C5	Recall	N_{S_y}	Precision	C5	Recall	N_{S_y}	Precision
<i>Configuration1</i> : $\eta_1 = \eta_2 = \eta_3 = \eta_4 = \eta_5 = \eta_6 = 1$													
2011	243	11	4.53 %	220	5.00 %	19	7.82 %	442	4.30 %	23	9.47 %	684	3.36 %
2012	238	9	3.78 %	229	3.93 %	19	7.98 %	460	4.13 %	28	11.76 %	716	3.91 %
2013	224	5	2.23 %	201	2.49 %	13	5.80 %	403	3.23 %	18	8.04 %	636	2.83 %
<i>Configuration2</i> : $\eta_1 = 1; \eta_2 = 0.5; \eta_3 = 0.5; \eta_4 = 0.5; \eta_5 = 0.5; \eta_6 = 0.5$													
2011	243	7	2.88 %	141	4.96 %	15	6.17 %	330	4.55 %	19	7.82 %	551	3.45 %
2012	238	7	2.94 %	148	4.73 %	15	6.30 %	348	4.31 %	24	10.08 %	585	4.10 %
2013	224	3	1.34 %	126	2.38 %	7	3.13 %	284	2.46 %	15	6.70 %	501	2.99 %
<i>Configuration3</i> : $\eta_1 = 0.5; \eta_2 = 1; \eta_3 = 0.5; \eta_4 = 0.5; \eta_5 = 0.5; \eta_6 = 0.5$													
2011	243	13	5.35 %	210	6.19 %	17	7.00 %	424	4.01 %	25	10.29 %	671	3.73 %
2012	238	11	4.62 %	223	4.93 %	23	9.66 %	446	5.16 %	32	13.45 %	703	4.55 %
2013	224	6	2.68 %	191	3.14 %	17	7.59 %	389	4.37 %	21	9.38 %	618	3.40 %
<i>Configuration4</i> : $\eta_1 = 0.5; \eta_2 = 0.5; \eta_3 = 1; \eta_4 = 0.5; \eta_5 = 0.5; \eta_6 = 0.5$													
2011	243	11	4.53 %	218	5.05 %	18	7.41 %	437	4.12 %	23	9.47 %	682	3.37 %
2012	238	9	3.78 %	230	3.91 %	20	8.40 %	459	4.36 %	29	12.18 %	712	4.07 %
2013	224	5	2.23 %	201	2.49 %	12	5.36 %	394	3.05 %	19	8.48 %	631	3.01 %
<i>Configuration5</i> : $\eta_1 = 0.5; \eta_2 = 0.5; \eta_3 = 0.5; \eta_4 = 1; \eta_5 = 0.5; \eta_6 = 0.5$													
2011	243	12	4.94 %	220	5.45 %	18	7.41 %	440	4.09 %	23	9.47 %	683	3.37 %
2012	238	9	3.78 %	226	3.98 %	20	8.40 %	457	4.38 %	29	12.18 %	712	4.07 %
2013	224	5	2.23 %	199	2.51 %	12	5.36 %	395	3.04 %	19	8.48 %	631	3.01 %
<i>Configuration6</i> : $\eta_1 = 0.5; \eta_2 = 0.5; \eta_3 = 0.5; \eta_4 = 0.5; \eta_5 = 1; \eta_6 = 0.5$													
2011	243	16	6.58 %	283	5.65 %	19	7.82 %	510	3.73 %	25	10.29 %	758	3.30 %
2012	238	15	6.30 %	280	5.36 %	22	9.24 %	537	4.10 %	29	12.18 %	795	3.65 %
2013	224	7	3.13 %	244	2.87 %	15	6.70 %	475	3.16 %	18	8.04 %	703	2.56 %
<i>Configuration7</i> : $\eta_1 = 0.5; \eta_2 = 0.5; \eta_3 = 0.5; \eta_4 = 0.5; \eta_5 = 0.5; \eta_6 = 1$													
2011	243	10	4.12 %	278	3.60 %	17	7.00 %	508	3.35 %	20	8.23 %	759	2.64 %
2012	238	8	3.36 %	288	2.78 %	17	7.14 %	525	3.24 %	24	10.08 %	793	3.03 %
2013	224	5	2.23 %	258	1.94 %	10	4.46 %	471	2.12 %	16	7.14 %	713	2.24 %

- [21] P. Serdyukov, M. Taylor, V. Vinay, M. Richardson, and R. W. White, "Automatic people tagging for expertise profiling in the enterprise," in *Proceedings of the 33rd European Conference on Advances in Information Retrieval (ECIR'11)*. Springer-Verlag, 2011, pp. 399–410.
- [22] I. S. Ribeiro, R. L. Santos, M. A. Gonçalves, and A. H. Laender, "On tag recommendation for expertise profiling: A case study in the scientific domain," in *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining (WSDM'15)*. New York, NY, USA: ACM, 2015, pp. 189–198.
- [23] K. Balog, I. Soboroff, P. Thomas, P. Bailey, N. Craswell, and A. P. de Vries, "Overview of the TREC 2008 enterprise track," in *Proceedings of The Seventeenth Text REtrieval Conference (TREC 2008)*, 2008.
- [24] G. Demartini, J. Gaugaz, and W. Nejdl, "A vector space model for ranking entities and its application to expert search," in *Proceedings of the 31th European Conference on IR Research on Advances in Information Retrieval (ECIR'09)*. Springer-Verlag, 2009, pp. 189–201.
- [25] S. H. Hashemi, M. Neshati, and H. Beigy, "Expertise retrieval in bibliographic network: a topic dominance learning approach," in *Proceedings of the 22nd ACM international conference on Conference on information and knowledge management (CIKM'13)*. New York, NY, USA: ACM, 2013, pp. 1117–1126.
- [26] M. Ley, "The DBLP Computer Science Bibliography: Evolution, Research Issues, Perspectives," in *Proceedings of the 9th International Symposium on String Processing and Information Retrieval (SPIRE 2002)*. London, UK, UK: Springer-Verlag, 2002, pp. 1–10.
- [27] E. Garfield, "Citation Indexes for Science: A New Dimension in Documentation through Association of Ideas," *Science*, vol. 122, no. 3159, pp. 108–111, 1955.
- [28] J. Tang, J. Zhang, L. Yao, J. Li, L. Zhang, and Z. Su, "Armetminer: extraction and mining of academic social networks," in *14th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD '08)*, 2008, pp. 990–998.
- [29] Z. Huang and Y. Qiu, "A multiple-perspective approach to constructing and aggregating citation semantic link network," *Future Generation Computer Systems*, vol. 26 Issue 3, pp. 400–407, 2010.
- [30] B. Gipp and J. Beel, "Citation Proximity Analysis (CPA) - A New Approach for Identifying Related Work Based on Co-Citation Analysis," in *Proceedings of the 12th International Conference on Scientometrics and Informetrics (ISSI'09)*, B. Larsen and J. Leta, Eds., vol. 2. Rio de Janeiro, Brazil: International Society for Scientometrics and Informetrics, Jul. 2009.
- [31] B. Gipp, N. Meuschke, and M. Lipinski, "CITREC: An Evaluation Framework for Citation-Based Similarity Measures based on TREC Genomics and PubMed Central," in *Proceedings of the iConference 2015*, Newport Beach, California, 2015, pp. 24–27.
- [32] L. Katz, "A new status index derived from sociometric analysis," *Psychometrika*, vol. 18, no 1, pp. 39–43, 1953.
- [33] D. Liben-Nowell and J. Kleinberg, "The link-prediction problem for social networks," in *Twelfth international conference on Information and knowledge management (CIKM'03)*, 2003, pp. 556–559.