Improving Scientific Conferences by enhancing Conference Management System with information mining capabilities

Andreas Pesenhofer*, Rudolf Mayer† and Andreas Rauber†
*iSpaces Group, eCommerce Competence Center – EC3,
Donau-City-Straße 1, A–1220 Wien, Austria
Email: andreas.pesenhofer@ec3.at
†Department of Software Technology and Interactive Systems
Vienna University of Technology,
Favoritenstraße 9–11/188, A–1040 Wien, Austria
Email: {mayer, rauber}@ifs.tuwien.ac.at

***Abstract*— In this paper we identify tasks in the field of conference management where methods from the domain of information retrieval, information management and information organization can assist the organizer, the program committee member and the participants. In particular, we focus on tasks, where (1) the quality of the conference can be increased by assisting in the creation of an improved review process by better matching the reviewers expertise with the paper topics, (2) the conference participants profit by allowing them better access to the wealth of information accumulated throughout a conference series and at the same time (3) reducing the workload of conference organizer by partially automating tedious tasks, such as review assignment and the creation of review plans. We report on case studies from a medium-sized (around 400 participants) as well as a large (more than 700 participants) conference in the computer science as well as in the medical domain.**

I. Introduction

One of the latest news from the International Congress & Convention Association (ICCA)[1] are the statistics of organized meetings for the year 2005. These rankings cover meetings organised by international associations with more than 50 participants and take place on a regular basis and which rotate between a minimum of three countries. For the year 2005 the ICCA Data researchers have identified 5,315 events, a rise of 511 over 2004. These statistics correspond with the list of 455 forthcoming conferences (mostly announced via DBWorld) for the time-frame the May 3rd 2006 to December 19th 2007[2].

The organization of a scientific conference is a challenging endeavor where a small error can have tremendous influence on the event. The IEEE, for example, provides a conferences organization manual[3], where organizers get detailed information about the organization. For the technical (scientific) part of the conference the use of web-based management system (such as CyberChair[1], ConfMan[2], Microsoft CMT[3], and OpenConf[4])[4] is indispensable to handle the huge amount of submissions. These systems fulfill the basic requirements and drastically ease the organization. Yet there are still many tasks where methods from the domain of information management and information visualization can assist to further improve the quality as well as to reduce the workload.

The process of assigning reviewers to submitted papers, for example, is one of the hardest and most time consuming tasks. Usually it is carried out by one single person, the program committee chair, who tries to select three to five reviewers for each paper. For a small conference with less than 50 submissions and up to 20 reviewers the manual assignment can be quite challenging, but for a conference with more than a thousand submissions and some hundreds of reviewers the task is more or less impossible to carry out. Due to many constraints (e.g. conflict of interest, area of expertise, equal amount of reviews per reviewer), this task can be quite difficult and exhausting. Yet, the quality of the reviews and thus, ultimately, the quality of the conference program, is highly dependent on a good assignments. Modern conference management systems [1], [2], [3], [4] have algorithms implemented that can automatically make the assignment. For these to work the program committee members have to select the topics which they are most interested in. In some systems they also have the possibility to bid for specific papers or reject them because of a conflict of interest. Only if all reviewers provide these information the algorithms can work effective otherwise they come to suboptimal solutions. Such solutions can for example be that work load for the reviews is not equally distributed; that reviewers get papers that they are not interested in and the worst case that a reviewer has to write a review for his own paper.

This paper describes the tasks in a conference management system where the use of information mining capabilities provides advanced methods to assist the organizer, the program committee member and the participants. The core goal of the

[1] http://www.iccaworld.com
[2] http://dbms.uni-muenster.de/menu.php3?item=confs
[3] http://www.ieee.org/web/conferences/mom/all_manual.html
[4] c.f. http://www.acm.org/sigs/sgb/summary.html

work presented in this paper are (1) to increase the quality of the conference by assisting in the creation of an improved review process by better matching the reviewers expertise with the paper topics; (2) to increase the benefits of conference participants by allowing them to better access and utilize the wealth of information accumulated throughout a conference series, while at the same time (3) easing the workload of conference organizers by partially automating tedious tasks, such as review assignment and the creation of review plans.

The remainder of this paper is structured as follows. Section II gives an overview of related work. Section III describes the basic functionalities of a conference management system. Three core tasks for further automatization will be tackled in Section IV, Section V and Section VI. Finally we conclude in Section VII.

II. Related Work

Conference management systems (like Confman [5], Confious [6], MyReview [7]) are web-based systems that cover some tasks so that the organization of scientific conferences can be carried out a little bit easier. Such tasks are for example the collection of submissions, the handling of assigned papers that the Program Committee (PC) members have to review, the download of papers, the handling of reviewers preferences and bidding, review progress tracking, web-based PC meeting, notification of acceptance/rejection and sending e-mails for notifications to authors or PCs. Once a bidding process has been performed, the assignment is handled as an as an optimization problem, trying to allocate papers according to reviewer preferences while striving for equal load distribution. A list of other software solutions, some of them are out-of-date, can be found at `http://www.acm.org/sigs/sgb/summary.html`.

Previous work in the area of assigning conference papers to reviewers had approached the problem as one of content-based information retrieval. Dumais and Nielsen used data provided by 15 members of the reviewing committee for the HYPERTEXT'91 conference [8]. These reviewers not only submitted abstracts of their papers and/or interests, but also provided complete relevance assessments for the 117 submitted papers. Information retrieval principles and latent semantic indexing were used to generate the automatic assignments for each reviewer. They achieved an average improvement of 48% with this method compared to the random assignment of articles to reviewers.

Yarowsky and Florian [9] focused on the classification of every paper to exactly one of six conference committees. They used 92 papers which were submitted to the ACL conference in electronic form and additionally requested committee members to provide representative papers so that a reviewer profile could be created. When the number of papers returned by these members was insufficient, they augmented the collection with other papers downloaded from online sources. The main algorithm first computed a centroid for each reviewer and then computed a centroid for each committee as the sum of its reviewer centroids. Then for each paper the cosine similarity was computed and compared with the committee centroids where the highest rank was the selection criteria. They also experimented with a Naïve Bayes classifier where their results outperformed the simple unsupervised model. Furthermore, they compared their systems with the performance of human judges on the same task. They concluded that the automatic methods could be as effective as human judges, especially in the case where the judges may be less experienced.

Paper recommendation is another solution to handle the problem of assigning papers to reviewers. Basu et al. described a content-based system for technical paper recommendation based on different information sources [10]. They treated the problem as one of decomposing reviewer interest and paper content into information sources, and then of combining the information sources using different query formulations. In their experiments they compared two ways of formulating the queries, first the content-based information retrieval and second the collaborative based approach. The algorithm using conjunctive queries outperformed the other approaches. Furthermore they noticed a general performance increase if they used more information.

For automatically establishing semantic similarities among papers and allocating them to common themes a prototype matching system for conference papers was presented in [11]. Furthermore, the system supports the attendees to retrieve the papers from the conference proceedings based on their content similarities The user can take an abstract or a paragraph from an interesting paper and use it as a prototype.

In [12] the assignment of papers is done based on previous collected user ratings. The paper describes a simple method which provides an approximate solution to the problem without requiring each user to rate each item. The method relies on an interactive process where in each step, or ballot, the users have to rate a sample of items. Collaborative filtering is then performed to predict the missing ratings as well as their level of confidence. Performing a new ballot allows to improve the accuracy of prediction. The administrator of the system has to stop the process when a satisfying level is reached. This algorithm tends to lead to a suboptimal solution, if only a sub group of reviewers attends the ballot and if only one ballot round is performed.

In [1] the assignment is made based on the bids for special papers and on the the reviewers' expertise on the conference topics and the willingness to review papers on these topics. The reviewers may bid in several stages and the bids are accumulated. The reviewer are also asked to inform the organizer if a conflict of interest is present for particular papers. The script is tuned to assign as many papers as possible based on the bids. Each paper is assigned to exactly 4 reviewers and all reviewers have the same amount of papers to review. Graph theory is applied to solve this assignment problem.

The latest work in this domain was carried out by Aleman-Meza et al. [13] where they describe a Semantic Web application that detects conflict of interest relationship among potential reviewers and authors of scientific papers. The degree of conflict of interest between the reviewers and authors were calculated based on a populated ontology. As input they integrated entities from two social networks, namely knows, from a FOAF (Friend-of-a-Friend) social network and co-

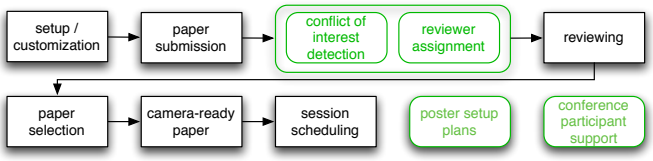

Fig. 1. Phases and tasks in a conference management system, tasks addressed in this paper are highlighted

author, from the underlying co-authorship network of the DBLP bibliography. Combining this two networks let them detect more potential conflict of interest than the simplified method that is implemented in [6].

III. Conference Management

A. User Roles and Tasks

In a conference management system users with different roles have to have access to specific tasks in a predefined time slot. An analysis of this roles and tasks is given in [5] and [6]. The following roles can be identified:

- **Organizers:** all kind of chairs (general, program, publicity, finance, ...), local arrangement group. The Program Committee Chair (PC Chair) is in charge of the coordination and monitoring of the necessary tasks.
- **Program Committee (PC) members** or **Reviewers**: persons who write the reviews and help to set up the scientific program of the conference. Specific PCs are often involved as session chairs. In case of a two stage review process, the Meta-Reviewers also belong to this group.
- **Authors:** persons who contribute papers and make presentations during the conference
- **Participants:** authors and other persons who are interested in this field of research. All persons registered for the conference
- **Visitors:** persons visiting the web page

In Figure 1 the basic tasks of a conference management system for the scientific program are described. Some tasks, such as for example the task ”paper submission”, can be split up in sub-phases e.g. an abstract submission phase followed by a full paper submission phase. Most of the tasks are allocated before the conference, only the participant support starts during the conference and lasts till the next conference. In this paper we will concentrate on the highlighted tasks, which are conflict of interest detection together with the automatic assignment of submissions to reviewer; the poster setup plans and the participant support.

B. Tasks for further automatization

In this section we will focus on tasks were further automatization eases the work of PC members and the PC chairs. We furthermore try to identify means to assist conference participants both during the meeting as well as, particularly, after the conference, to make the most of the wealth of information presented during the meeting, as well accumulating over the years in a conference series.

1) Task: Reviewer Assignment: The submission to reviewer assignment is done either automatically or manually by the PC chair, with an automatic assignment usually being followed by a manual adjustment. For the assignment the following constrains are taken into consideration:

- The submission topics should match with the reviewer interests.
- A reviewer’s bid for specific papers have to be taken into consideration.
- Reviewers should not get their own paper to review. A potential conflict of interest between the PC members and submissions has to be calculated.
- Each PC member should get the same amount of papers to review, so that they have the same work load.

All these tasks rely on the input of the PC members. This can cause troubles if some of the PC members are reluctant or too busy. The PC chair (administrator of the system) has no possible to make decisions for them, being limited to sending reminder mails and asking for their cooperation. The algorithm can not work properly and produces suboptimal solutions that have to be corrected manually by the PC chair.
In Section IV we will focus on an automatic assignment of the submitted papers to the PC members based on their previous publications as a baseline for the manual bidding process. It overcomes these problems, because public available publications of the authors are used to create the profiles for the PCs.

2) Task: Session Creation: Based on the accepted papers the sessions have to be created. Here, similar papers are grouped together. The constraints (e.g. number of parallel sessions, sessions per day) have to be defined by the PC chair. Under this constraints the PC chair has to group papers manually to sessions and assign each session a session chair. Here clustering methods as described in Section V in combination with scheduling algorithms may be used. A final solution to this task is not provided in this paper.

3) Task: Poster Alignment: During most of the conferences posters are presented in a special room or in the lounges of the conference venue. Usually there exists a pre-setup provided from the organizer where authors have to fix their poster. For this case the organizers have to figure out which posters fit best to each other grouping them by topic. PC chair has to align the poster manual. Mnemonic SOMs as described in Section V can be used for this alignment.

4) Task: Participant support: The conference program should be up to date in the web and the proceedings should be searchable either publicity or limited to registered conference participants via dedicated logins. Participants may be interested if they have missed interesting sessions. Mnemonic SOMs and SOMs in combination which the participant’s interests give the participants new insight in the huge amount of information presented during the conference as well as help them to prepare their schedule before attending large events. We will address that in more detail in Section VI.

There are numerous tasks that offer considerable potential for automatization, such as e.g. in the production of consistent printing and on-line material (e.g. web, program broshure,

proceedings, notice boards), accounting, etc. which are not dealt with in this paper.

C. Case Studies

We report on case studies from two conferences, the *9th European Conference on Research and Advanced Technology for Digital Libraries* (ECDL 2005)[5] and the *European Congress of Radiology 2004* (ECR 2004)[6]. The ECDL is the major European conference on digital libraries, and associated technical, practical, and social issues in this field. It can be classified as a mid sized conference with around 100 to 200 submissions, around 80 to 90 program committee members and around 350 to 450 participants. The ECR is a large sized conference species with more than 2,000 scientific paper submissions taking place every year in Vienna. It is the largest radiological meeting in Europe attracting more than 15,000 participants from over 90 countries.

The data has to be transformed into a representation, so that it is understandable by the algorithms. Therefore we indexed the data based on the the well known bag-of-words approach with Lucene[7] using a tfidf weighting scheme [14], which is based on the term frequency (tf) in the given document and the inverse document frequency (idf) of the term in the whole collection. Pre-processing steps in form of removing all numbers, punctuation marks and special characters were applied. The tfidf values were normalized to unit vector length.

1) ECDL corpora: For the ECDL we have to distinguish between three corpera:
ECDL A: Is made out of 723 automatically retrieved publications from PC member's home pages and 125 submissions. Term reduction based on document frequency and term length was applied, resulting in a vector with 8,767 unique terms.
ECDL B: Consists of the accepted poster submissions, of 30 different posters in English language. After applying an English stop word list and other term reductions based on the document frequency, we obtained a vector of 569 different terms.
ECDL C: Is composed the accepted paper and poster submissions, totaling to 71 documents. Applying the same mechanisms as for the ECDL B corpus, we obtained a vector of 5,654 different terms.
In all three cases no stemming was applied.

2) ECR corpus: This corpus consists of the abstracts of the European Congress for Radiology (ECR)[8] from the year 2004. All together these are 943 English documents which were presented during the scientific sessions of the congress and each belongs to one of the 15 different topics (c.f. Table VI). Every abstract is assigned exactly to one topic. Additionally to the settings that were described in the beginning of the section we also applied an english stop word list and only kept those terms that had a df between two and 300. In the end, the corpus consisted of 3,842 unique terms.

Additionally to that we received the radio frequency identification (RFID) logs that were collected during the conference. At the registration every participant received a badge with a unique RFID tag. The entrances to halls of the conference location were guarded with RFID gates, so that the organizer could log if one accesses a session. These are used in the medical domain for the monitoring and issuing of continuous education certificates. They serve to build an anonymized participant profile for our experiments.

[5]http://www.ecdl2005.org
[6]http://www.ecr.org
[7]http://lucene.apache.org
[8]http://www.ecr.org

IV. Profile based Reviewer Assignment

A good paper to reviewer assignment is based on the cooperation of the PC member (reviewer). They have to choose from a list of relevant topics where they are interested in and furthermore they have to bid for special papers by to skimming through the abstracts. Most of the PC member do not bid nor choose their interests so that the algorithms fail in computing a proper assignment. Which is particularly due to the fact that a bidding process for 200 and more papers is a notoriously time consuming task. Our solution overcomes this problem, because the interest of the reviewer is defined based on previous publication that are available on the internet.

A. Profile generation

We used the forename and the surname of the PC members to formulate the search query. We sent the query to two search engines which are providing scientific papers, namely that was CiteSeer.IST[9] and GoogleScholar[10]. From the returned search result pages the URLs linking to the publications was extracted. Using the 87 PC members from the ECDL 2005 conference resulted in 4,369 retrieved URLs. In the next step we downloaded these documents discarding all non PDF documents. As result we obtained the 723 potential publications. Note that for ten PC members no publications have been automatically retrieved.

B. COI detection

The potential conflict of interest detection (COI) was performed based on (1) the occurrence of the last name of a program committee member in the authors line of a submission and (2) the existence of parts from the PC members email domain in the submissions author field (e.g. if the PC members email has the domain tuwien.ac.at and the submission comes from the same domain then we will register a potential conflict of interest). Using these two methods allowed us to identify 46 potential conflict of interest for the PC members for the ECDL 2005 data set.

We compared our results with the COI that the PC members registered during the biding phase of the ECDL conference. Here in only 24 cases a COI was registered. A detailed comparison of the two lists reveals the following:

1) More than the half (57.69%) of the reviewers did not bother to register a COI. This group of persons was additionally identified by our system.
2) Potential COI was detected by the system but not registered by the reviewer, who in principal did register

[9]http://citeseer.ist.psu.edu/
[10]http://scholar.google.com/

paper/reviewer	preferences & bid-based	profile-based
4	40	12
5	9	27
6	2	25
7	5	17
8	31	6
PC member (sum)	87	87

TABLE I
DISTRIBUTION OF THE REVIEW WORKLOAD

50% of the COI. As reasons we identify that the COI was not considered in spite of e.g. the same lab, but no close cooperation and that the paper was overlooked due to the large list of papers. A solution would be to have a system, that detects a potential COI and presents to the reviewer to confirm it.

3) In 7 cases the COI was registered by the reviewer, but not detected by our current system. In these cases co-authorship analysis would have to be included (e.g. DBLP) and for areas that are not covered by a specific digital library of papers web-based search has to be performed.

C. Reviewer Assignment

Before we can calculate the assignment, we have to find out which submission match with the interests of which PC member. Therefore we compute the Euclidian distance between every submission and publication based on the full-text indexed feature vector. A distance of 0 means that the two compared documents are identically and a the higher the value is the more different they are. A PC member has normally more than one publication in his profile, so that we kept only the smallest distance from all his documents to one submission.

As baseline for our evaluation we used the automatic assignment, that was calculated on the ECDL 2005 PC member preferences and their bids. To make our system comparable with the baseline we set up an identical system without the bids and the paper topic interest of the PC members. The aggregated distances were sorted starting with the smallest and ending with the largest. The first ten received a rate level of 4 which correspond to a bid of "eager" to review, the next ten were rated with "interesting" (3) and the remaining received the level 1 ("better not"). For the ten cases where no publications could be found automatically, and therefore no distances to the submission existed we used 2 ("indifferent") as default rating. If a COI in the relation was detected a rate level of 0 ("conflict of interest") was inserted into the data base. These pre-calculated values sere as a basis for the bidding process, that may be optimized by the user.

Table IV-C summarizes the workload distribution of the PC members using the assignment model based on preferences and bids compared to the results that where obtained with the profile-based assignment. In both cases we have 500 reviews that have to be assigned to the 87 PC members, the optimal amount of assigned papers per PC member would have been 5,75. In the first case, the preferences & bid-based model, 40 reviewer get four papers to review and 31 reviewer get the maximum amount of papers (8) assigned. Only 16 reviewer get 5 to 7 papers assigned. In out system, the profile-based one, only 6 reviewer have a workload of 8 papers and 12 PC member have only four papers to review. Most of the PC members (27) got 5 papers, followed by 25 that got 6 and 17 that got 7 papers assigned. In this case much more PC members are allocated around the mean of 5,75 resulting in a more equal distribution than in the first case.

V. POSTER SETUP PLANS

If the set-up of the poster locations is defined by the conference organizers, this can be done in several different ways: the set-up may be organized completely randomly, or sorted alphabetically by author names or submission titles. It may be desirably, though, to organize the submissions by their content - that way, interested conference participants can easily find the areas with posters about topics they are interested in. Organization by content may be done using manually assigned category labels, coming either from the authors themselves during submission, or from the PC. However, such a categorization may in many cases not be available at all, available only for some parts of the submissions, or of poor or varying quality. Then, as an alternative, unsupervised clustering algorithms based only on the submission contents may be utilized to determine a poster set-up.
Independent of the exact set-up, the conference participants should also be provided with a map of the venue, indicating poster locations and topic areas, to assist them in locating the posters they are interested in.

Unsupervised clustering and generating a map of the poster set-up can be achieved using e.g. the Self-Organizing Map (SOM) [15]. The SOM is a neural-network model that provides a mapping from a high-dimensional input space to a lower dimensional output space. In this mapping, the SOM preserves the topology of the input space, i.e. input patterns that are located close to each other in the input space will also be located closely in the output space, while dissimilar patters will be mapped on to opposite map regions. In many applications, this output space is made of a two-dimensional, rectangular map. This representation allows for an easier interpretation of the complex structure of the input patterns by the user.

Another advantage of using the SOM is that it generates a clustering that preservers transitions between clusters - documents that would belong to two different clusters will be mapped on the border in between those clusters.

In our application, the input space will be formed by a vector-space representation of the poster submissions, as described in Section III-C, while the output space will be the map of the poster session area.

As in many cases the area for the poster session may not be of rectangular shape, we use a modifcation to the original SOM algorithm, the *Mnemonic SOM*, as presented in [16]. In the Mnemonic SOM, the output space is two-dimensional, but can take any arbitrary shape. They can be easily generated

Concepts of Digital Libraries, Documents and Metadata		Digital Preservation, Web Archiving	
System Architectures, Open Archives, Integration		Digital Library Applications & Case Studies	
Information Retrieval & Organization, Search & Usage		Multimedia, Audio, Video	
User Studies, Evaluation, Personalization, UI			

TABLE II
CLASS LEGEND FOR THE ECDL 2005 DATA

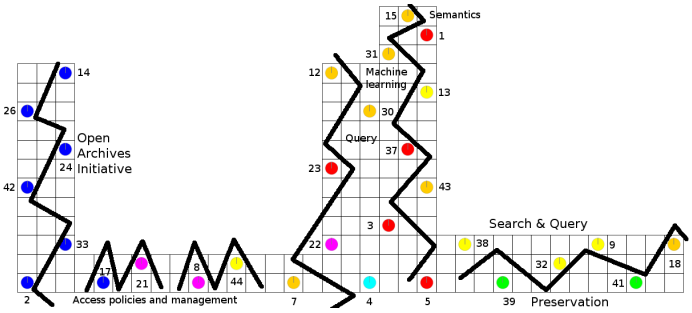

Fig. 2. Poster allignment for the ECDL 2005 conference.

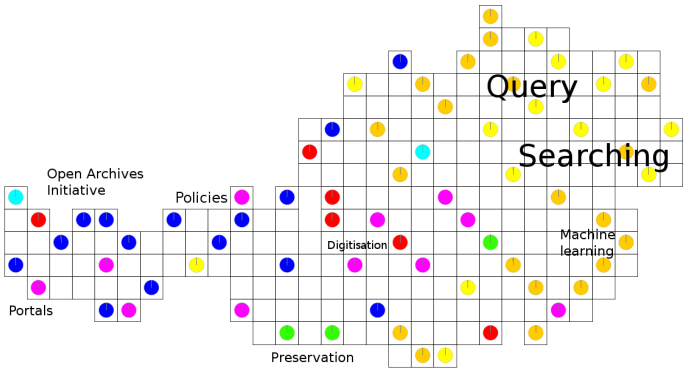

Fig. 3. Map of the submission to ECDL 2005.

Abdominal and Gastrointestinal		Interventional Radiology	
Breast		Musculoskeletal	
Cardiac		Neuro	
Chest		Pediatric	
Computer Applications		Physics in Radiology	
Contrast Media		Radiographers	
Genitourinary		Vascular	
Head and Neck			

TABLE III
CLASS LEGEND FOR THE ECR 2004 DATA.

from a black and white image representing the desired shape, e.g. of the poster presentation area.

We have applied this method for arranging the poster set-up during the *9th European Conference on Research and Advanced Technology for Digital Libraries* (ECDL 2005). Table V gives an overview of the topics of the submitted posters to this conference. The category assignment was given by the authors on submission.

Figure 2 shows the generated mapping, where the output space was made of a grid with the size of 35x15, with 182 units within the map shape. It is based on the layout of the conference poster area. Black lines show the set-up of the poster boards, and numbers indicate the ID assigned to each poster on submission.

We can observe that thematically similar posters get arranged close to each other, e.g. in the top-left we can find posters dealing with the "Open Archives Initiative Protocol". The poster arrangement does not necessarily follow the manual categorization, but arranges them by content.

The given data set contains a lot of different, sometimes rather small clusters. This is due to the small size of the data set (30 accepted posters), and the very heterogenous topics they discuss. However, the quality of the generated mapping is good.

Using the method described above can help the conference organizer both in saving time on the poster set-up, and achieving a better thematically grouped set-up.

VI. PARTICIPANT SUPPORT

The method of the SOM, described in Section V, can also be well utilizied for supporting the participant during and after the conference.

One application is to provide an advanced interface to the proceedings of the conference, in addition to traditional key-word based searching or manually created indices. We again generate representations of all the presentations at the conference via a vector-space represantion of the abstracts, and map the documents on a SOM. Figure 3 gives an example from the ECDL 2005 conference in Vienna (cf. Section V), where we use a map in the shape of Austria as a mnemonic hint for the participants. The submissions, including both papers and posters, have been grouped automatically according to their topic by the SOM algorithm. The colored pie-charts visualize the distribution of the manually assigned categories of the docuements. The labels on the map (e.g. Query, Search, Machine Learning) have been added manually after inspecting the content of the documents grouped together in this region.

The scientific abstracts of the ECR 2004 were also pre-processed as described in Section III-C.2 and mapped onto a SOM, this time following the shape of the logo of the Austria Center Vienna (ACV), the location where the conference takes place every year. The shape of the logo also represents the basic form of the ACV building.

Figure 4 illustrates how this content based mapping on the shape of the ACV is done by the SOM algorithm. In Table VI the category names and colors of the ECR 2004 are provided, so that the evaluation the map with the colored pie-charts, that visualize the distribution of the categories, can be done easily. On the far left corner papers dealing with "Vascular" (magenta; mark 1) are arranged together. The papers dealing with "Computer Applications" (orange; mark 2) have their cluster on the right hand side. Papers dealing with "Interventional Radiology" (grey) are split up into two cluster, where the first one (mark 3) is dealing with embolization and

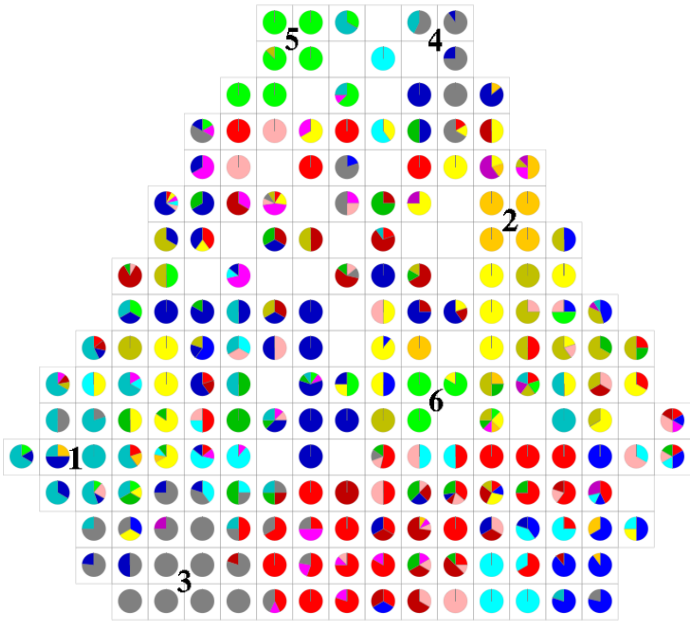

Fig. 4. Scientific submissions to the ECR 2004 mapped on the ACV logo

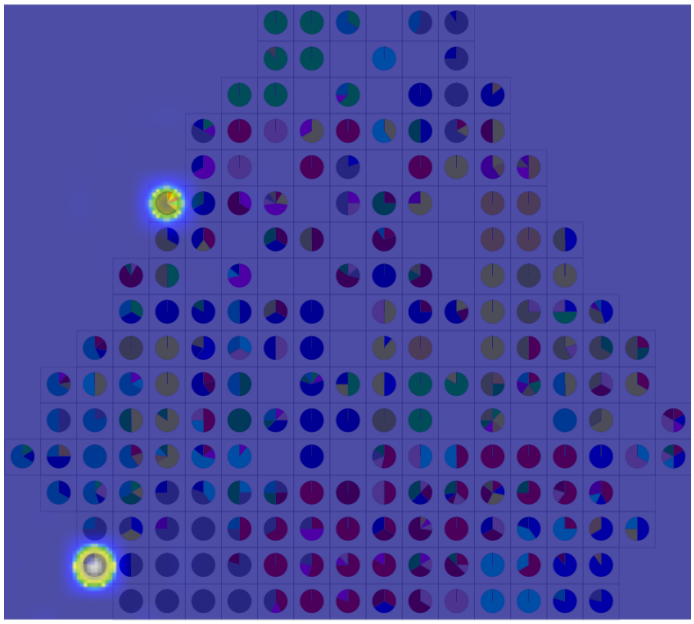

Fig. 5. Fingerprint of participant A

the second one (mark 4) is dealing with different kinds of stents. In the neighboring cluster (mark 5) the papers also deal with stents, in particular with coronary artery stent, chest pain and thrombus detection belonging to the class ”Cardiac” (light green). In the second ”Cardiac” region (mark 6) the documents deal with ventricles, myocardial infarction and myocardial scar.

In Figure 5 and Figure 6 we used the attendance information of a participant (RFID logs) to create personalized fingerprints. We identify the locations of these abstracts that where presented in sessions that the participant attended, and created a hit histogram. The more focused a participant is, the more concentrated the histogram appears on the map. Participants can so immediately see where their interests are located on the map and by looking on surrounding regions of their fingerprints they may find relevant information for them.

In Figure 5 shows a participant interested in ”Interventional Radiology” (lower left side) and ”Neuro” (top left side).

For the second participant (Figure 6) six regions are highlighted. The two sessions entitled ”Myocardial viability and wall-motion” and ”Evaluation of cardiac function” both are part of the ”Cardiac” topic. They are allocated next to each other forming a larger cluster on the right. Two documents from the last session are mapped to the top left of the map, where a second ”Cardiac” cluster can be identified. The session dealing with ”Molecular Imaging” papers can be found on half way down to the ”Neuro” region. Going down and a little bit to the right we come to the ”Musculoskeletal” session and going diagonal to the left we end up in the furthermost left spot, described by ”Interventional Radiology” section. This participant attended sessions with five different topics, what can be seen by the fingerprint.

In case of the ECR the personalized fingerprint can be added to the profile of the user. The accepted scientific papers of the upcoming conference can be trained as a mnemonic SOM in the shape of the ACV. Using the stored fingerprint allows the participants mark their interests on the actual conference map and help them to decide which of the session to visit.

VII. Conclusion

We presented information mining methods that enhance scientific conference management systems. We showed that organizers, reviewers and participants of mid sized and large sized conferences benefit of our proposed methods, by (1) easing the task of paper to reviewer assignment for the organizer; (2) by resulting in more equal paper distribution; (3) by helping in creating poster setup plans; (4) by providing the registered participants better access to the scientific papers and also helping them to decide which of the session they should visit in the future conference.

References

[1] R. van de Stadt, “The CyberChair Software,” Web Site: http://cyberchair.org/, 1997.

[2] P. Halvorsen, K. Lund, V. Goebel, T. Plagemann, T. Preuss, and H. König, “ConfMan (Conference Manager),” Web Site: http://www.ifi.uio.no/confman/ABOUT-ConfMan/, 1999.

[3] S. Chaudhuri and J. Grieves, “The Microsoft Conference Management Toolkit,” Web Site: http://msrcmt.research.microsoft.com/cmt/, 1999.

[4] Zakon Group, “OpenConf Conference Management System,” Web Site: www.openconf.org, 2001.

[5] P. Halvorsen, K. Lund, V. Goebel, T. Plagemann, T. Preuss, and H. König, “Architecture, implementation, and evaluation of ConfMan: Integrated WWW and DBS Support for Conference Organization,” I-1998.016-R, UniK, University of Oslo, Norway, Tech. Rep., December 1998.

[6] M. Papagelis, D. Plexousakis, and P. N. Nikolaou, “CONFIOUS: Managing the Electronic Submission and Reviewing Process of Scientific Conferences,” in *Proceedings of the 6th International Conference on Web Information Systems Engineering (WISE 2005, Industrial Track)*, New York, USA, 2005, pp. 711–720.

[7] P. Rigaux, “The MyReview System,” Internet: http://myreview.lri.fr/, 2004.

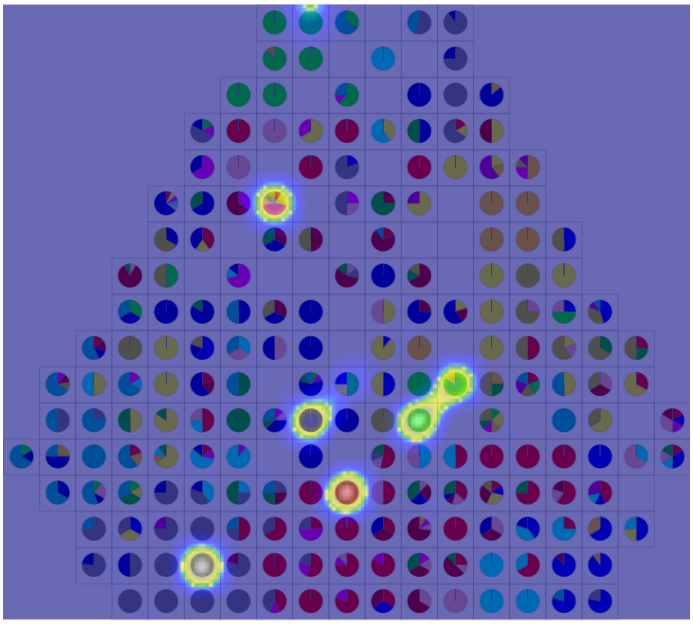

Fig. 6. Fingerprint of participant B

[8] S. T. Dumais and J. Nielsen, "Automating the Assignment of Submitted Manuscripts to Reviewers," in *Research and Development in Information Retrieval*, 1992, pp. 233–244.

[9] D. Yarowsky and R. Florian, "Taking the load off the conference chairs: towards a digital paper-routing assistant," in *Proceedings of the 1999 Joint SIGDAT Conference on Empirical Methods in NLP and Very-Large Corpora.*, 1999.

[10] C. Basu, H. Hirsh, W. Cohen, and C. Nevill-Manning, "Technical paper recommendation: A study in combining multiple information sources," 2001.

[11] K. Antonina, B. Barbro, V. Hannu, T. Jarmo, and V. Ari, "Prototype-matching system for allocating conference papers," in *HICSS '03: Proceedings of the 36th Annual Hawaii International Conference on System Sciences (HICSS'03) - Track 3*. Washington, DC, USA: IEEE Computer Society, 2003, p. 70.

[12] P. Rigaux, "An iterative rating method: application to web-based conference management," in *SAC '04: Proceedings of the 2004 ACM symposium on Applied computing*. New York, NY, USA: ACM Press, 2004, pp. 1682–1687.

[13] B. AlemanMeza, M. Nagarajan, C. Ramakrishnan, L. Ding, P. Kolari, A. P. Sheth, I. B. Arpinar, A. Joshi, and T. Finin, "Semantic Analytics on Social Networks: Experiences in Addressing the Problem of Conflict of Interest Detection," in *Proceedings of the 15th International World Wide Web Conference*, 2006.

[14] G. Salton, *Automatic text processing – The Transformation, Analysis, and Retrieval of Information by Computer*. Addison-Wesley Longman Publishing Co., Inc., 1989.

[15] T. Kohonen, *Self-organizing Maps*, ser. Springer Series in Information Sciences. Berlin, Heidelberg: Springer Verlag, 1995, vol. 30.

[16] R. Mayer, D. Merkl, and A. Rauber, "Mnemonic SOMs: Recognizable Shapes for Self-Organizing Maps," in *Proceedings of the Fifth Workshop on Self-Organizing Maps (WSOM'05)*, M. Cottrell, Ed., Paris, France, September 5–8 2005, pp. 131–138.