"Mining Bibliographic Data" – Using Author's Publication History for a Brighter Reviewing Future within Conference Management Systems

Christian Caldera, René Berndt, Eva Eggeling
Fraunhofer Austria Research GmbH, Graz, Austria
Email: {christian.caldera, rene.berndt, eva.eggeling}
@fraunhofer.at

Martin Schröttner
Institute of Computer Graphics and Knowledge Visualization
Graz University of Technology, Graz, Austria
Email: martin.schroettner@cgv.tugraz.at

Dieter W. Fellner
Institute of Computer Graphics and Knowledge Visualization (CGV), TU Graz, Austria
GRIS, TU Darmstadt & Fraunhofer IGD, Darmstadt, Germany
Email: d.fellner@igd.fraunhofer.de

Abstract—Organizing and managing a conference is a cumbersome and time consuming task. Electronic conference management systems support reviewers, conference chairs & the International Programme Committee members (IPC) in managing the huge amount of submissions. These systems implement the complete workflow of scientific conferences. One of the most time consuming tasks within a conference is the assignment of IPC members to the submissions. Finding the best-suited person for reviewing a paper strongly depends on the expertise of the IPC member. There are already various approaches like "bidding" or "topic matching". However, these approaches allocate a considerable amount of resources on the IPC member side. This article introduces how the workflow of a conference looks like and what the challenges for an electronic conference management are. It will take a close look on the latest version of the Eurographics Submission and Review Management system (SRMv2). Finally, it will introduce an extension of SRMv2 called the Paper Rating and IPC Matching Tool (PRIMA), which reduces the workload for both – IPC members and chairs – to support and improve the assignment process.

Keywords–conference management, conference tools, paper assignment, matching algorithms, TF-IDF, information retrieval.

I. INTRODUCTION

Conferences and journals play an important role in the scientific world. Both are important channels for the exchange of information between researchers. The publication list of a researcher defines his standing within the scientific community. In order to ensure quality standards for these publications, submitted work go through the so called peer-review process.

An approach of finding suitable reviewers for this process has been presented at the *International Conference on Creative Content Technologies* (CONTENT), where the foundation of this article has been discussed [1].

This process is used to maintain standards, improve performance and provide credibility [2]. Today almost every conference or journal uses an electronic conference management system in order to organize this process.

In-a-nutshell the peer review process for a conference undergoes the following steps (see Figure 1):

- **Submission Phase** In the submission phase authors need to specify a certain amount of descriptive meta-data, which is required for organizing the process.
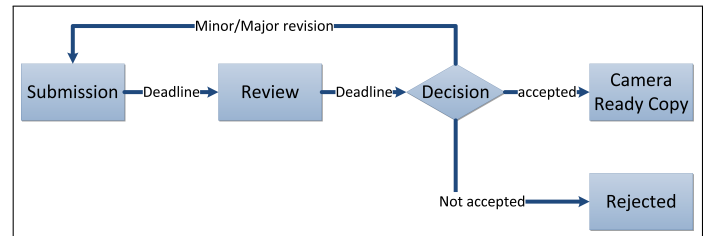

Figure 1. The typical phases of a conference and how submissions pass through this conference.

Besides title and co-authors these can include abstract, keywords and conference specific categorizations (e.g., poster or tutorial tracks). Finally, they need to provide their paper and optionally additional material (e.g., videos).

- **Review Phase** When the submission deadline has passed, the submitted papers are distributed to the reviewers. The conference reviewers are usually members of the International Programme Committee and – depending on the size of the conference – a pool of external experts. Each reviewer receives a certain amount of submitted papers depending on his expertise and workload. Further he must not be in any kind related to the author to prevent a conflict of interests. Assigning the submitted papers to the IPC members is a crucial task in the peer review process because their reviews decide, if the paper is accepted or not.

- **CRC Phase** In case of acceptance the author is allowed to upload a camera-ready-copy (CRC) version of the paper, which is then published in the proceedings of the conference.

- **Rejected** If the paper is not accepted, the submissions enter a special rejected phase. Rejected submissions are not further considered in the review process and are only used for statistical purposes (e.g., total amount of submissions in the conference).

This is the essence of the peer review process. Within the peer review process there exist variations, which mostly differ in what information is revealed to whom. The most commonly used are the single and double blinded peer reviewing process:

- **Single Blind** In the single blinded peer review the identity of the reviewer is unknown to the author. But, the reviewer knows the identity of the author. In this setting, the reviewer can give a critical review without the fear that the person itself will be targeted by the author.

- **Double Blind** In the double blinded peer review, the identity of the reviewer and author is unknown to each other. This process guarantees the same chances for unknown and famous scientist and universities by not putting the name on the paper.

There are further versions of peer reviewing like open peer reviewing or additions like post-publication peer reviewing, but they are rarely been applied [3][4].

Although there are many criticisms about peer review [5], the following quote from Mayur Amin from Elsevier at the APE (Academic Publishing in Europe) Conference, Berlin in January 2011 describes the current situation:

> Peer review is not perfect, but it's the best we have.

One particular point of criticism is the poor referee selection [6]. Especially conferences with a large number of submitted papers experience an enormous time pressure for finding suitable reviewers.

How can a conference managing system support the conference chair during the reviewer assignment phase? One idea is to utilize information available from different sources about the particular persons in order to find a suitable reviewer or identify conflicts of interest. Especially bibliographic data can be a valuable source of information for finding suitable reviewers. Figure 2 shows the image section of data sources related to publications within the Linking Open Data cloud. An example of such a data source is the DBLP [7], which provides bibliographic information on computer science.

Another bibliographic service is Microsoft Academic Search [8]. Figure 3 shows the visualization of a co-author graph from Microsoft Academic Search. If a person is a direct neighbour in the co-author graph of an author, this person does most likely have a conflict of interest with the author and cannot review the paper.

II. CONFERENCE EXAMPLE: EUROGRAPHICS ANNUAL CONFERENCE

Since the year 2000 the Eurographics (EG) uses the MCP system (Managing Conference Proceedings) [9] and the successor SRM (Submission and Review Management) as their conference management system. These systems have been especially tailored to support the needs of the EG.

In order to get an insight of the work of a conference chair we take a detailed look at the Eurographics Annual Conference, which is organized by the Eurographics. Figure 4 shows the number of submitted/accepted papers over the last 14 years. The requirement of at least four reviews for each paper leads to approximately more than 1000 review assignments. Assuming that the average workload of a reviewer should not exceed five reviews means that at least 200 suitable (and willing) persons have to be found.

How does the review process work in detail? The submitted papers are distributed to the members of the IPC. Each paper is assigned one primary and one secondary reviewer. These act as
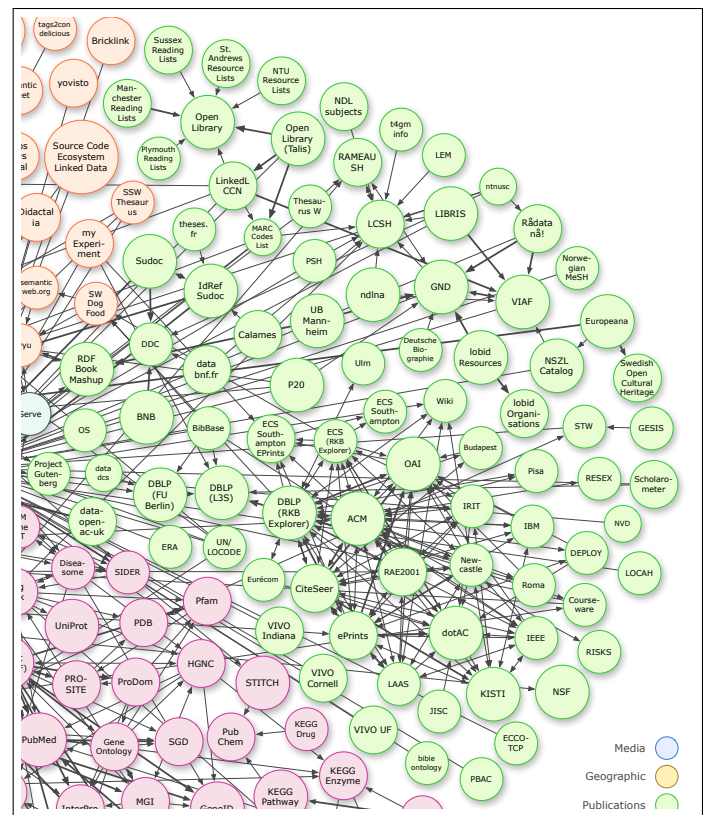

Figure 2. An excerpt of the Linking Open Data cloud diagram, by Richard Cyganiak and Anja Jentzsch. Image Source: http://lod-cloud.net

editors for this particular paper meaning they are responsible for finding at least three additional reviewers. Distributing the available submissions to the IPC members has turned out to be the most time-consuming task for the conference chairs in the last years. In order to support the distribution process the so-called "Bidding-Phase" has been introduced with SRM. IPC members are presented a list of all submitted papers (title, abstract, keywords). For each of these papers the IPC member could specify one of the following options: "want review", "could review", "not competent", "conflict of interest". Based on this classifications the system creates an automatic suggestion how to distribute the IPC members as primary/secondary reviewers. It is further possible at the start of the conference to define a list of categories. To this available categories the IPC members could specify the degree of expertise. ("expert", "passing", "not competent"). These values were matched with the author-selected categories for each paper. The weighted sum of both values indicate then the appropriateness of an IPC member for that specific paper.

Although the process of peer reviewing is unquestioned within Eurographics, over the years valuable input from the chairs in order to improve the process have been made. One of the most discussed issues was the selection of suitable reviewers. Although this weighted sum works well for the distribution, the bidding values have to be entered by each IPC manually. Going through a list of more then 200 titles and abstracts is cumbersome.
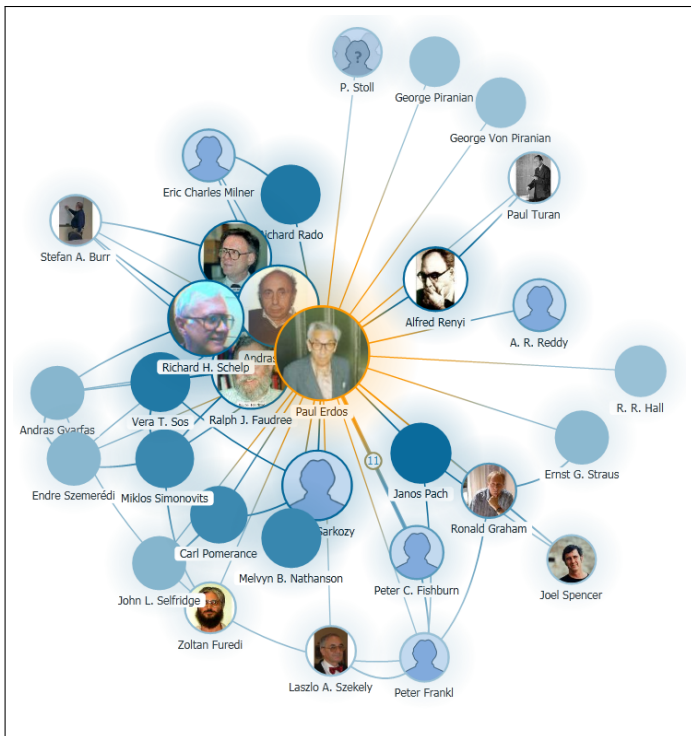
Therefore, the next version of SRM should use a new

Figure 3. The co-author graph of Paul Erdös in the Microsoft Academic Search. Image Source: Screenshot in the Microsoft Academic Search http://academic.research.microsoft.com
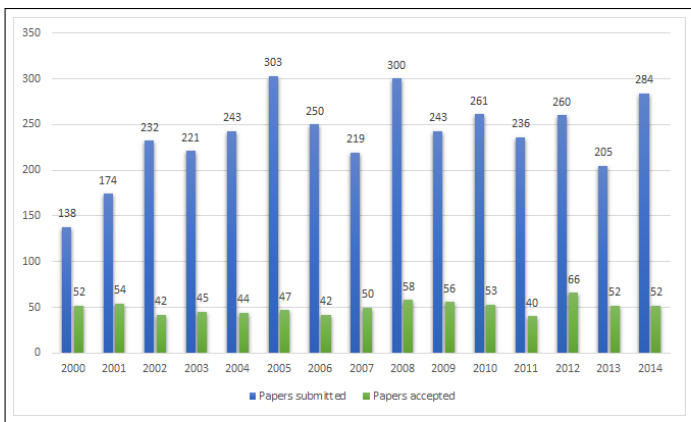

Figure 4. The submitted and accepted papers of the Eurographics Annual Conference over the last 14 years.

approach to use information available by Linked Open Data, especially bibliographic information. With the new system it should be easy to interact with 3rd party applications or data sources. It should be easy to harvest the data to create statistics and further usage of this data. Additionally, this solution should handle in a similar fashion like the current system.

III. RELATED WORK

This section will give a small overview of some of other reviewed conference management systems:

EasyChair is a free service for managing a conference.

It provides the basic features like uploading a submission and a reviewing system. It has multiple conference models a chair can choose from to customize the conference. Beside the models it is not possible to further modify the conference [10]. The review assignment process in EasyChair works manually or automatically. When using the automatic mode the International Programme Committee members define the conflicts of interests and then these members specify, which papers they are interested to review. After this is done EasyChair tries to create a good matching between the Committee members and the papers [11].

COMS Conference Management System has a one time set up fee for creating a website to satisfy the needs of the conference chair. This website will be the frontend for the chair's conference management system. Once the homepage is created the chair may define nine different review fields. The review assignment works again either manual or automatic. The automatic mode takes the reviewers biddings like in Easychair and creates a matching between the reviewers and the submissions [12].

OpenConf is a php based conference management tool, which has again the standard functionality for managing a conference. OpenConf provides the basic conference management tools. There are additional modules to add functionality to the program. One of these modules called the bidding module adds the functionality for the International Programme Committee members to define, which papers they want to review. After this bidding OpenConf provides some different algorithms to create a matching between the reviewers and the papers [13].

Confious has also the standard features for managing a conference. Confious has like the other systems an automated and a manual reviewer assignment system. But unlike the other systems Confious takes the paper topics into consideration. Authors define, which topics their paper is in and the Committee members set their experience in these topics. Then, it tries to create a good matching. Confious also tries to generate automated conflicts based on the Email and the institute of the IPC member and the author [14].

Conftool is a tool, which provides many different languages to manage a conference. Like Confious its automated review assignment takes the IPCs bidding and the paper topics into consideration when creating a review assignment. It also tries to create conflicts like Confious according to the Email, the organization and the surname of the reviewer [15].

IV. THE EUROGRAPHICS CONFERENCE MANAGEMENT SYSTEM

The Eurographics Association has already a long tradition in maintaining its own conference managing system. The start of the activities date back to 1997, when Prof. Fellner was the chairman of the Eurographics annual conference, during which he experienced the amount of complex work by himself. At that time, conference support systems was just coming into existence. The first Eurographics conference management system was the *Managing Conference Proceedings* system (MCP), which has been developed by Marco Zens as a part of his PhD thesis [9]. Based on this prototype an improved version called SRM (Submission and Review Management) was developed. Since then SRM received several updates, managed 25 conferences and had over 11.000 members. Over the years valuable input from conference chairs were gathered.

These suggestions and the knowledge of the current SRM system has led to a new SRM concept, which is introduced in the next section.
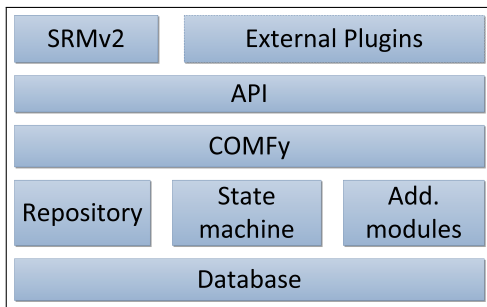

Figure 5. The concept of COMFy.

A. System Architecture

Based on the conference example in Section II the concept how the new system works is now presented. The new system provides an application programming interface (API) for managing a conference. The new SRMv2 system communicates with the core layer through a well defined API. The core layer of the new conference management framework (COMFy) provides the application logic, while the SRMv2 purely consists of the user interface. Additionally, external programs are also able to communicate with that API. COMFy itself maintains the conference data and uses the repository pattern in order to separate the business logic from the database layer. The states of a paper are represented by a state-machine. The core layer can be extended by additional modules.

The new system is divided into five different layers. The lowest layer is the database and each of the upper layers uses the functionality of the lower layers and adds additional features to the system (see Figure 5).

- **Database** The bottom layer is a relational database. Microsoft SQL Server [16] was chosen because of its filestream feature. Usually, files are stored either in the database or in a directory [17]. When stored in the database the performance to access the files is decreased drastically. On the other hand, when the file is stored in the directory the transactional consistency is lost. The filestream feature of the Microsoft SQL Server combines these features by managing the files in database management system, but stores it in a directory, which is managed by the database itself. With filestreams the transactional consistency is guaranteed and the file can be accessed fast via the directory.

- **Repositories** On the second layer there are repositories. The task of the repositories is to provide the upper layers an easy way to access the data in the database. When COMFy queries one of the repositories, this repository maps the request to an SQL statement. When the query is executed, it returns the data to the upper layer. It also works in the other direction, so the upper layer can insert new data or update existing entries.

- **State-machine** The second part of the second layer is the state machine. This is the core of the framework.

It manages the phases of every submission. When a submission changes its phase it also changes the access rights of different users. For example, an author may not submit a new paper once the reviewing process starts. Another integral part of the design with the state-machine is that it is easily extensible. For example, another phase like a rebuttal phase where authors may object to the decision of the conference chair, can easily be added to the system. The current state-machine can be seen in Figure 1.

- **COMFy** This layer contains the business logic of the conference management system. It exposes these functionalities through a well defined API. It queries the repositories, parses the data and creates the response. It is also responsible for applying the different user roles, e.g., an author does not have access to the reviewer names, etc. It is designed as a model view controller pattern. So the controller takes care of the request, queries the repository, fills the model with the data from the repository and returns it to the view. Depending on the client the requested data can be delivered as Extensible Markup Language (XML), JavaScript Object Notation (JSON) or Hyper-Text Markup Language (HTML).

B. COMFy API

The API on COMFy is based on the representational state transfer (REST) paradigm [18], which utilizes the well-known, well-defined methods of the Hypertext Transfer Protocol (HTTP) protocol (GET, POST, PUT, DELETE). This paradigm is based on clean, human readable, hierarchical uniform resource locators (URL) for accessing the resources. COMFy uses clean, structured URLs to access the requested data. The API calls can be divided in four different categories:

- **UserHome:** The "UserHome" API calls are used for retrieving information of conferences and submissions, which are tied to the user. This way the user can quickly access his own submissions or conferences.

- **Account:** The "Account" API calls are used for managing user accounts, e.g., logging into the system, registering or changing the profile information.

- **Conference:** The "Conference" API calls are for managing and viewing a conference. These calls are primarily by the chair when setting up the conference.

- **Submission:** The "Submission" calls are used for managing and viewing a particular submission. They need the conference identifier because the paper identifier is only unique within one conference. This way it is easy to identify in the URL the conference a submission is in. The calls are mainly used by authors and reviewers.

The COMFy API encapsulates the core elements of a conference system. However, an API cannot provide a clear use-case model what API calls a user needs to do for a certain task. Therefore, SRMv2 is implemented on top of COMFy as one sample application. Next to SRMv2 it is also possible to use external plugins to extend the functionality of COMFy. The *Paper Rating And IPC Matching Tool* (PRIMA) in Section V is an example, how to use this API and add extra functionality to the system.

C. Conference Setup

With the old SRM system it became clear that different chairs require different conference settings. One of the design goals of COMFy is that the conference setup should dynamic and adapt to the needs of the conference chair. The best way to show the dynamic nature of COMFy are its dynamic fields. In the two most important phases of the conference, the submission and review phase (see Figure 1) the fields can be dynamically adjusted to the conference. The following section will show how these dynamical fields work.

1) Custom Submission Fields: Figure 6 shows the submission page, where the authors enter all required meta-data for their submission. While title and abstract are common fields for all conferences, the other form fields correspond to the configured custom fields as shown in Figure 7. It is also optionally possible to upload a representative image, which identifies the submission.

Figure 7. A possible setup for a conference. These fields have to be entered by the author when he submits a paper to the conference.

2) Custom Review Fields: Experiences from the past conferences have shown the need for customizable review forms. In order to address these requirements, SRMv2 supports four different types of custom review fields:

- **TextArea** defines a simple free-text field for the reviewer.
- **ComboBox** allows the reviewer to select a value from a predefined vocabulary.
- **ScoreBox** works like the ComboBox. But the text is matched against an integer value. This can be used, for example, like school grades to calculate the average score of the reviews and their deviation. The first value of the scorebox can be defined with the start value. Every following entry matches to the incremental integer value.
- **CheckBox** A simple checkbox, which can be ticked.

For all four types it is possible to add an optional description to the field and define the order how they appear to the reviewer. It is also possible to add a comment field for each field. This way the reviewer is able to state the reason behind his review entry. Each review has a dynamic overall recommendation and dynamic evaluation confidence. These are fixed fields, as they appear in every review. It is also possible to define, which dynamic fields should appear behind and before these two fields.

An example of a possible review setup can be seen in Figure 8. This setup generates a form for the reviewer, which can be seen in Figure 9. This review form can be downloaded as XML or HTML form. The reviewer can fill out the form offline and upload it later. In the end of the reviewing process the primary reviewer can use the overview (see Figure 10) to give the chair a senior recommendation. Or the chair can check himself/herself at the state of the reviews.

Figure 6. A dynamic generated submission page. The fields in the submission page are defined in Figure 7.

D. Review Assignment

One major challenge in the Eurographics conference is the assignment of reviewers. Within SRMv2 two different approaches exist to accomplish this task. The automated and the manual assignment. With the manual approach the chairs assign reviewers to submissions from an user-pool. In the automatic approach the program tries to create a distribution between the reviewers and the submissions.

The manual assignment can be seen in Figure 11. After choosing the submission the person who will be assigned to

After the form has been filled out, the authors can upload their paper in the portable document format (PDF) along with additional multimedia material. Authors can modify their submission data, until the submission deadline has passed. Only the last uploaded version will be go into the reviewing process. The corresponding author receives an email for each successful upload of his paper.

Figure 8. A possible setup for a conference. These fields have to be entered by the reviewer in order to complete his review.

Figure 9. The example review fields setup in Figure 8 generates this review form. This review form has to be filled up by the reviewer to complete his review.

Figure 10. The review overview allows comparing the various reviews for one paper.

Figure 11. A warning when assigning Paul Erdös to a paper where there might be a conflict of interest with the authors of the paper.

review the paper, SRMv2 checks Linked Open Data sources like the Digital Bibliography & Library Project (DBLP) if there are conflicts of interest between the reviewer and the authors. A strong link is found, when the full names of the author and the reviewer who was selected appear in the co-author list. A weak link is found, when the domain name of the e-mail, the organization or the surname of the author and the assigned reviewer matches.

The assigning person can ignore the warning if he knows that there is no conflict of interest between the assigned person and the author. Then the assigning person has to select the user and set his reviewing role. Currently, there are three different roles: primary, secondary and tertiaries. After this task it is possible to modify the standard e-mail to create a more personalized e-mail. At last the assigning person has to confirm the assignment, so the email will be sent and the person gets

Figure 12. Harvested co-authors entries retrieved from the DBLP.

his assignment, which he can accept or decline.

A new feature implemented in SRMv2 is that the chair can now access information from the DBLP to indicate, if the person might have a conflict of interest because of a co-author relationship. In the current version the bibliographic data from DBLP is used to help identifying these conflicts. The DBLP provides an API. On this API users can query for authors. Every author in the DBLP system has an unique author identifier. After querying an author for the author identifier, it can be used to get the co-authors of that particular person. They also provide information about the amount of publications the two authors wrote together. It is also possible with the DBLP API to receive bibtex files of papers. These papers can also be found with the mentioned author pointer. Figure 12 shows the information within SRMv2 collected from the DBLP. Access to other sources like Mendeley [19] or Microsoft Academic Search are already under development.

While the first version of the Eurographics conference management system Manage Conference Proceedings (MCP) completely relied on the expertise and experience of the conference chairs, the successor SRM implemented two new features: Reviewing preferences and Bidding. With the release of SRMv2 the Paper Rating and IPC Matching Tool (PRIMA-Tool) was introduced. These approaches are described in detail in the following paragraphs.

For the automatic assignment it is necessary for IPC members to complete three steps. At first they are presented with a list of all authors. On this author list they can set their conflicts of interest with them. Then they set in, which area they are experts in. In their last step they are presented with every paper. There they set, which paper they would like to review and in, which they are not knowledgeable enough to review it. Once this is done for every IPC member COMFy tries to create the best matching of reviewer to the submission.

Before such a matching is created COMFy currently cross checks the DBLP, if there are some coauthor links, which are not defined by an IPC. If some links are found the chair is notified in the suggested matching. Currently, this system is redundant as IPC members are checking their conflicts by hand. In the future this automated assignment process will be improved and the cross checks against the DBLP should replace the manual conflict settings.

1) Reviewing Preferences: One new feature of SRM was the area of expertise list (AoE). Based on this AoE list authors could select up to five topics, where their paper fits best. In the first version of SRM all conferences shared the same AoE list, which was based on the 1998 version the ACM classification scheme[1]. Especially for some workshop series this scheme was too broadly defined, so it was decided that each conference could specify their own AoE list. The (obvious) pitfall of this decision was that now for every conference the IPC members would have to newly specify their reviewing preferences. Even for the same workshop series the AoE lists did not stay stable. Figure 13 shows an example, where the user can specify whether he is an **expert**, **knowledgeable**, **passing** or has **no knowledge** about the given topics.

Figure 13. An IPC member can specify his knowledge in the areas of the conference.

2) Bidding: In order to support the distribution process the so-called "Bidding-Phase" has been introduced with SRM. IPC members are presented a list of all submitted papers (title, abstract, keywords). For each abstract the IPC members can specify whether they **want review (0)**, **could review (1)**, are **not competent (2)** or have a **conflict (3)**. Figure 14 shows the resulting paper/IPC matrix.

SRM uses a weighted sum combining the *Reviewing Preferences* and the *Bidding*:

$$IPC suggestion = aoeRating + biddingRating \quad (1)$$

The first term in the sum of Equation (1) calculates the score from the IPC reviewing preferences and the categories entered by the author:

$$aoeRat. = 25 \cdot \frac{4 \cdot \#\{e\} + 2 \cdot \#\{k\} + \#\{p\}}{\#\{Area \quad of \quad expertises \quad of \quad Paper\}} \quad (2)$$

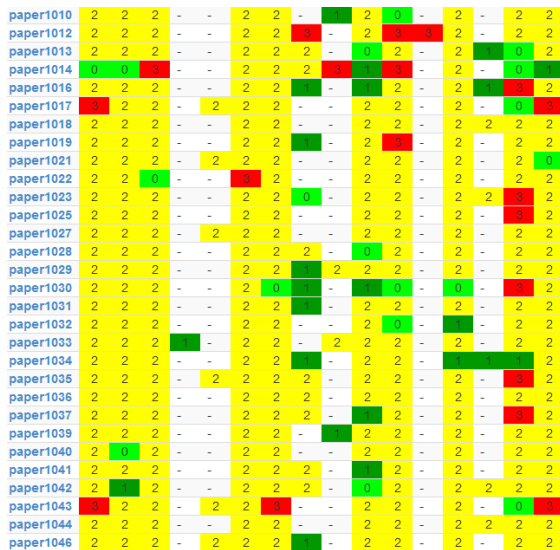[1]http://www.acm.org/about/class/1998

Figure 14. Bidding matrix showing for each submission the bidding value entered by the IPC members. 0 means they want to review this paper, 1 that they could review the paper, 2 that they are not competent and 3 that there is a conflict of interest on one of the authors. If there is no value on the bidding matrix, then there is no information about this author on this paper.

The intersection of categories where the IPC member is expert in and the paper categories is defined as *e*. *k* is the intersection, where the IPC member is knowledgeable and *p*, where the IPC has only passing knowledge. The resulting *aoeRating* will be a value between 0 and 100.

The second term comes from bidding and maps the user input to a value between 0 and 100 (see Equation (3)):

$$biddingRating = \begin{cases} 100 & Want\ Review \\ 80 & Could\ Review \\ 0 & no\ Expertise \end{cases} \quad (3)$$

V. PAPER RATING AND IPC MATCHING TOOL - PRIMA

Paper Rating and IPC Matching Tool (PRIMA) is the third option for IPC members to define their reviewing preferences. It is a standalone extension to SRMv2. PRIMA uses the *Term Frequency Inverse Document Frequency* [20] (TF/IDF) algorithm in order to calculate the similarity between all submitted papers and previous papers of the IPC members for extracting a matching value allowing an automatic distribution of submissions to IPC members.

Figure 15 shows the workflow of the automatic score generation with PRIMA. In the first step, the PRIMA tool is initialized with the required data for the IF/IDF calculation:

- The submitted paper along with their meta-data
- Information about the IPC members of the selected event.

PRIMA uses the API of the SRMv2 framework [21] in order to fetch the required information. After the initialization, the IPC members are invited by e-mail to upload their publications fitting the scope of the conference. The more papers a user uploads into PRIMA the better the algorithm can find different matchings to the submissions of the conference. After all initial data is available (submitted papers of the conference and the uploaded publications of the IPC members), the paper scores are calculated. These scores are then transmitted to SRMv2 in

order to support the pre-ordering for the bidding process and to support the automatic assignment proposal.

Before the calculation itself starts some preprocessing steps are necessary to improve the TF/IDF result:

- **Text extraction:** For all uploaded publications, the raw text is extracted from the PDF documents. The extracted text still contains a large number of unnecessary information, which do not have an impact on the paper classification, for example, numbers, special characters, code, URLs, e-mail addresses, punctuation, authors, addresses, IDs, etc. Future work on PRIMA concentrates on further separating the text, which is useful for the TF/IDF score generation from the overhead part, which interferes with the generation [22].

- **Removal of stop words:** Stop words are words, which occur often in a text but do not add any informational value to the text. Some examples of this stop words are: *and*, *or*, *the*, *an*, *important*, *however*, *just* and so on. All these words are necessary for the creation of sentences. But two texts do not relate strongly to each other just because they have a lot of "and" together [23].

- **Stemming:** Stemming reduces words to their common root. For example, "overview" and "overviews" are not the same words in a computational matching, so the word *overviews* is reduced to *overview*. These two words will then match in the algorithm. [24].

After the preprocessing steps PRIMA starts the TF/IDF algorithm.

A. Term Frequency Inverse Document Frequency

The TF/IDF algorithm can be separated into two parts. The first part is the *Term Frequency* part. It uses the frequency of terms in a document to classify the document. The second part of the algorithm is the *Inverse Document Frequency*. It weights the terms according to the occurrence in all other documents. The more a term is used in different documents the less information it provides for classifying a document [25]. The algorithm itself is already a quite understood and researched topic in different areas like text categorization, text analysis, mining and information retrieval techniques [20].

The function $f(t, d)$ in Equation (4) counts every term t in the document d. After it has been counted every term is normalized with the logarithm.

$$tf(t, d) = log(1 + f(t, d)) \quad (4)$$

The inverse document frequency (see Equation (5)) counts the occurrences of a term across all documents in a given document corpus. This is done by taking the logarithm of the quotient between the total number of documents $|D|$ and the amount of documents d containing the term t. A term, which occurs in every document is not useful for categorizing, so it has to be penalized for being not important in the current global text corpus. Terms, which occur in fewer documents receive a higher value with this formula.

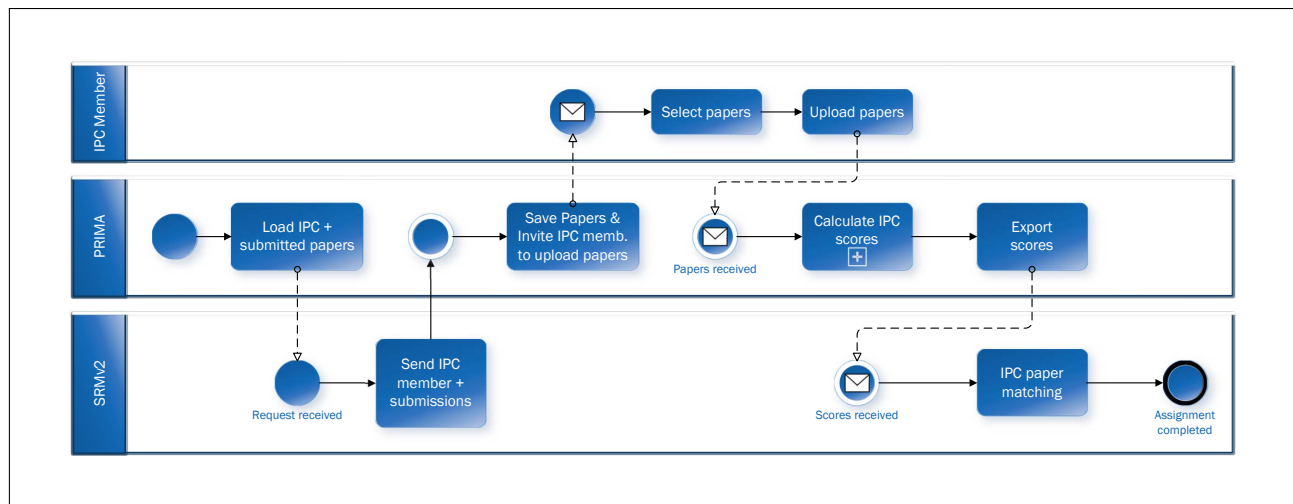$$idf(t, D) = log \frac{|D|}{|\{d \in D : t \in d\}|} \quad (5)$$

Figure 15. The complete workflow of the Paper Rating And IPC Matching Tool (PRIMA).

By multiplying the term frequency with the inverse document frequency the TF/IDF is received (see Equation (6)). This value classifies a term in a document and its classification significance across all documents [26].

$$tfidf(t, d, D) = tf(t, d) \cdot idf(t, D) \qquad (6)$$

All TF/IDF values of a document form a vector, which classifies the document. By calculating the angle (see Equation (7)) between two documents, it is possible to extract a similarity value [27].

$$\cos(a, b) = \frac{a \cdot b}{\parallel a \parallel \parallel b \parallel} \qquad (7)$$

After the TF/IDF values are calculated, each submission is compared against all papers of the IPC members with the cosine similarity. If an IPC member has provided multiple publications, all of them are checked against a single submission paper. Currently, the average of the five best matching papers is saved. This is done to more robust against statistical outliers. Furthermore, not all papers are taken into consideration as a person might upload a lot of papers belonging to different areas.

VI. Performance and results of PRIMA

PRIMA was first tested for the Eurographics 2014. The papers and the reviewers are anonymized and randomly re-ordered. The International Programme Committee consisted of 70 members and a total of 290 submissions were received. Every IPC member had entered their conflicts, defined areas of expertise to create a pre-filtering for the submissions and finally bidded on the paper. This final bidding matrix consists of 290 x 70 = 20300 entries. Figure 16(a) shows a small excerpt of the bidding matrix. The rows represent five reviewers (a to e), the columns represent 16 submissions (1 - 16). The cells are formatted with the following color scheme:

- **Light green** (0) The IPC member wants to review the paper.

- **Dark green** (1) The IPC member could review the paper.

- **Yellow** 2) The IPC member considered himself as not competent enough to review this paper.

- **Red** (3) The IPC member has a conflict with the authors of the submitted paper.

- **White** (−1) No data has been provided by the IPC member.

Most reviewers take the default *not competent* or did not submit any values at all. In the first prototype the default value for the bidding was *could review*, but due to requests from the majority of IPC members over several events this was changed to *not competent*. Therefore only a few papers contain information on the suitability of the IPC member for reviewing this paper [1].

About 300 randomly selected papers of these IPC members were uploaded and together with the 290 submissions analyses through the TF/IDF algorithm. Figure 16(b) shows the same excerpt for the TF/IDF algorithm. For each paper a value between 0 and 1 is calculated by PRIMA, where 0 means no word overlap in both documents and 1 means every word in both papers appear at the same amount. In order to archive a similar appearance like the initial bidding matrix, the following thresholds have been applied:

- **Light green** (1 − 0.1) High correlation between the paper and the uploaded papers of a IPC member

- **Dark green** (0.1 − 0.05) Medium correlation between the paper and the uploaded papers of a IPC member

- **Yellow** (0.05 − 0.0) Low correlation between the paper and the uploaded papers of a IPC member.

- **Red** The conflicts of the original bidding.

Figure 16(c) shows the transposed bidding and calculated matrix of reviewer C for easier comparison. The first row shows the values of the calculation, the second the bidding result of the reviewer. An important observation is, that the left bidding matrix consists of a large number of not entered information. Possible explanations are, that an IPC only checked
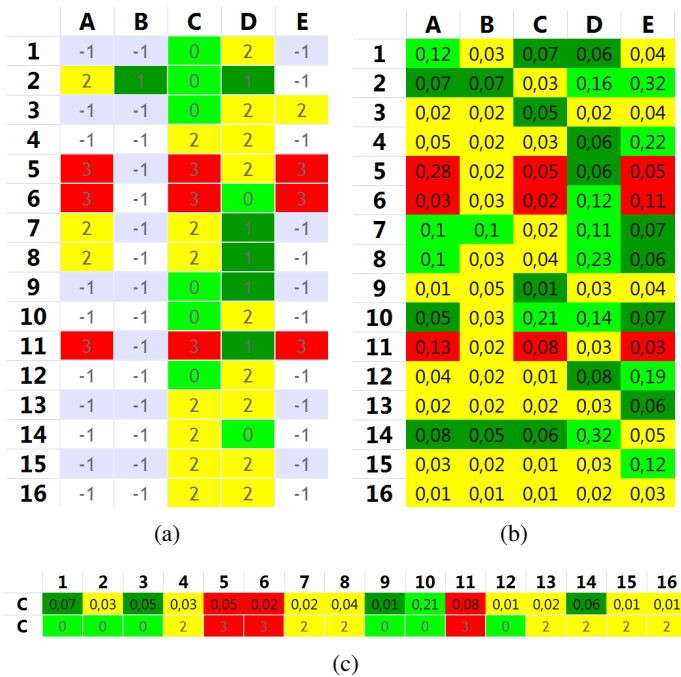
Figure 16. Figure (a) shows a small excerpt of the bidding matrix. Most reviewers set the values to the default not competent or did not submit any values at all. The values and colour scheme is the same like in Figure 14. Figure (b) shows an excerpt of the PRIMA matrix with the same color encoding like the bidding matrix and thresholds at 0.05 and 0.1. Figure (c) shows the transposed bidding and calculated matrix of reviewer C for easier comparison.

the papers in his own area of expertise or was not able to read all 290 submission abstracts, because of a lack of time.

Another important observation is that some good matchings are conflicts (e.g., cell $C11$ with the value 0.08 is a conflict in the left figure). It can be expected that a person who is an expert in an area also might have a project cooperation with other experts in this field and therefore has a conflict of interests with these persons.

Furthermore, it can be observed that most of the bidding values match the calculated values of PRIMA $C1$, $C3$, $C9$, $C10$ (Figure 16(c)). In addition, also the not competent column matches with the biddings $C4$, $C7$, $C8$, $C13$, $C16$. For test phase only publications from previous Eurographics events were uploaded as input for the IPC members and the amount of uploaded data also differs. For example, IPC member D had 18 uploaded papers, but person B only five. For this reason person D is much better classified by the TF/IDF and therefore has a in general better matching than person B.

Strong differences between the bidding and the calculated classification, e.g., for person C the cells $C2$, $C12$, and $C14$, can have multiple reasons. According to the TF/IDF the IPC member would be well suited as a reviewer, but he considered himself as not competent. This can have different reasons:

- The TF/IDF has analysed an older paper of the person, but the expertise focus of the person has changed.
- The title and abstract from the bidding might have been misleading.
- The submission was overlooked by the IPC member

and this submissions stayed on the default value, which is *not competent*.

The first item will be addressed in further research in order to analyze if penalty value for older paper will improve the results. But also cases where the rating from the TF/IDF shows a low score, but the persons claimed that he *wants to review* occur, for example, in $C2$ and $C12$:

- Most likely the system does not have a current paper of the IPC member on this topic.
- The reviewer is interested in a paper and "wants to review" it, but does not have the necessary knowledge to review it.

The calculated values provided by PRIMA have a huge advantage: They provide indications whether an IPC member is a suitable reviewer for a given paper even if the IPC member provided no bidding information. As stated before a large portion of the bidding matrix is not filled up. In these cases it is possible to create a better reviewer-to-paper assignment instead of randomly distributing the submitted papers to the reviewers. For example, in submission 4 the best matches are person D and E, for submitted paper 13 person E would be a good choice.
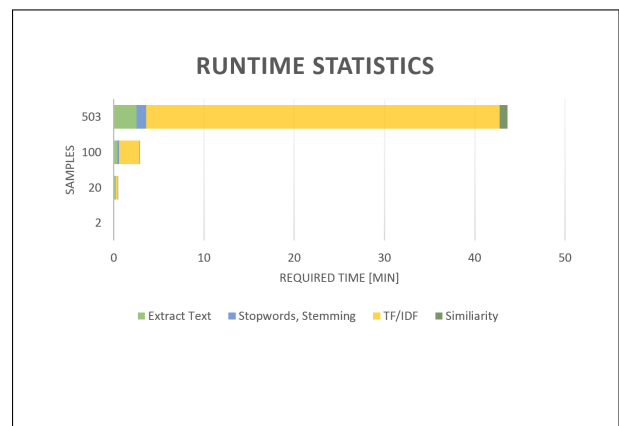

Figure 17. The amount of time each of the tasks take. It can be seen that the algorithm has an exponential growth.
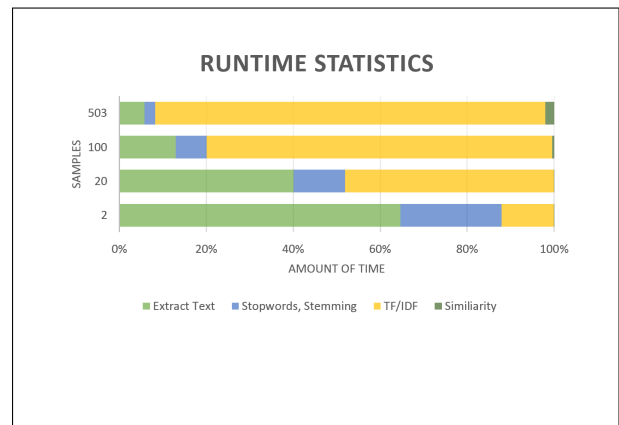

Figure 18. The amount of time each of the tasks take, split by the tasks and scaled to 100%.

Figure 17 shows the runtime statistics of the PRIMA tool split into the five steps *extract text*, *stopwords, stemming*, *TF/IDF*, and *similiarity*. It can be seen that for a small number of papers the extraction of the text and the stopwords removal and stemming takes up most of the time. If the number of papers increases, the more time the TF/IDF algorithm itself takes. The text extraction and stopwords removal and stemming can be precalculated and stored. However, this step takes less than 10% of the time during the full calculation using more than 500 papers. The TF/IDF itself cannot be precalculated as every further submission changes the weighting of each word in the calculation process. This means that the calculation can only start once all papers of all IPC members are available.

VII. Conclusion & Future Work

In this paper, we presented the COMFy conference management framework and its interaction with the PRIMA tool, which automatically calculates a ranking between submitted papers and the available reviewers. By using the TF/IDF for categorizing the submitted papers along the reviewers expertise, the workload of the reviewers and the conference chairs is reduced dramatically.

Currently, the selection and upload of publications is done manually by the reviewers. Using citation portals like DBLP [28], Citeseer [29] and other sources, the selection and retrieval of the full-text version (e.g., when available through the Open Access [30] initiative) can be automated as well.

Another important point, which might be improved is the text extraction itself. At the moment, the whole paper is used for the TF/IDF calculation. And although the numbers, special characters, URLs, stopwords, etc., are removed there are still words, which slip through, which should not be used for the analysis. For example, words like the author, the institution, figure explanations, headings, formulas and so on.

For the upcoming Eurograhics conference it is planned to use SRMv2 alongside with PRIMA and to evaluate the scores by presenting submissions to the authors in descending order. Then, the IPC member can concentrate on the title/abstracts, which fit best to the topics of his own publications. The values that the PRIMA tool generates can also be used as suggestions for the reviewer during the bidding process. This way the member can skim over the values and check if they fit. This will save the IPC members valuable time, which can be used for more important research [1].

References

[1] C. Caldera, R. Berndt, M. Schröttner, E. Eggeling, and D. Fellner, "PRIMA - towards an automatic review/paper matching score calculation," in Proceedings of The Sixth International Conference on Creative Content Technologies. IARIA, 2014, pp. 70–75.

[2] Academia Publishing, "What is Peer Review?" 2014, URL: http://academiapublishing.org/ [accessed: 2014-12-09].

[3] R. M. Blank, "The effects of double-blind versus single-blind reviewing: Experimental evidence from the american economic review," American Economic Review, vol. 81, no. 5, December 1991, pp. 1041–67, [retrieved: 03, 2014]. [Online]. Available: http://ideas.repec.org/a/aea/aecrev/v81y1991i5p1041-67.html

[4] M. W. Consulting, "Peer review in scholarly journals: perspective of the scholarly community–an international study," Author, Bristol, UK, 2008.

[5] R. Smith, "Peer review: a flawed process at the heart of science and journals," JRSM, vol. 99, 2006.

[6] D. Shatz, Peer Review: A Critical Inquiry (Issues in Academic Ethics (Paper)). Rowman & Littlefield Publishers, Inc., Nov. 2004. [Online]. Available: http://www.amazon.com/exec/obidos/redirect?tag=citeulike07-20\&path=ASIN/0742514358

[7] M. Ley, "Dblp: some lessons learned," Proc. VLDB Endow., vol. 2, no. 2, Aug. 2009, pp. 1493–1500. [Online]. Available: http://dl.acm.org/citation.cfm?id=1687553.1687577

[8] Microsoft Corporation, "Microsoft academic search," Dec. 2012, URL: http://academic.research.microsoft.com/ [accessed: 2014-12-09].

[9] M. Zens, "Creation, management and publication of digital documents using standard components on the internet," Ph.D. dissertation, Technische Universität Braunschweig, 2004.

[10] A. Voronkov, "Easy chair conference system," Dec. 2012, URL: http://www.easychair.org/ [accessed: 2014-12-09].

[11] N. Garg, T. Kavitha, A. Kumar, K. Mehlhorn, and J. Mestre, "Assigning papers to referees," Algorithmica, vol. 58, no. 1, 2010, pp. 119–136.

[12] M. Mandl, "Conference management system (coms)," Dec. 2012, URL: http://www.conference-service.com/ [accessed: 2014-12-09].

[13] Zakon Group, "Openconf," Dec. 2012, URL: http://www.openconf.com/ [accessed: 2014-12-09].

[14] M. Papagelis and D. Plexousakis, "Confious," Dec. 2012, URL: http://www.confious.com/ [accessed: 2014-12-09].

[15] H. Weinreich, "conftool - conference management tool," Mar. 2013, URL: http://www.conftool.net/ [accessed: 2014-12-09].

[16] G. Fritchey and S. Dam, SQL Server 2008 Query Performance Tuning Distilled, 1st ed. Berkely, CA, USA: Apress, 2009.

[17] R. Sears, C. van Ingen, and J. Gray, "To blob or not to blob: Large object storage in a database or a filesystem?" CoRR, vol. abs/cs/0701168, 2007.

[18] R. T. Fielding, "Architectural styles and the design of network-based software architectures," Ph.D. dissertation, University of California, Irvine, 2000, aAI9980887.

[19] Mendeley Ltd., "Mendeley," Jan. 2013, URL: http://www.mendeley.com/ [accessed: 2014-12-09].

[20] R. Baeza-Yates and B. Ribeiro-Neto, Modern Information Retrieval, 1st ed. Addison Wesley, May 1999.

[21] C. Caldera, R. Berndt, and D. W. Fellner, "Comfy - A Conference Management Framework," Information Services and Use, vol. 33, no. 2, 2013, pp. 119–128, [retrieved: 03, 2014]. [Online]. Available: http://dx.doi.org/10.3233/ISU-130697

[22] E. Rahm and H. H. Do, "Data cleaning: Problems and current approaches," IEEE Data Eng. Bull., vol. 23, no. 4, 2000, pp. 3–13.

[23] C. Silva and B. Ribeiro, "The importance of stop word removal on recall values in text categorization," in Neural Networks, 2003. Proceedings of the International Joint Conference on, vol. 3, 2003, pp. 1661–1666 vol.3.

[24] M. Porter, "An algorithm for suffix stripping," Program: electronic library and information systems, vol. 14, no. 3, 1980, pp. 130–137.

[25] J. Ramos, "Using TF-IDF to Determine Word Relevance in Document Queries," Department of Computer Science, Rutgers University, 23515 BPO Way, Piscataway, NJ, 08855e, Tech. Rep., 2003.

[26] C. D. Manning, P. Raghavan, and H. Schtze. Cambridge University Press, 2008, [retrieved: 03, 2014]. [Online]. Available: http://dx.doi.org/10.1017/CBO9780511809071.007

[27] A. Huang, "Similarity Measures for Text Document Clustering," in New Zealand Computer Science Research Student Conference, J. Holland, A. Nicholas, and D. Brignoli, Eds., Apr. 2008, pp. 49–56. [Online]. Available: http://nzcsrsc08.canterbury.ac.nz/site/digital-proceedings

[28] M. Ley et al., "DBLP Computer Science Bibliography," 2013, URL: http://www.informatik.uni-trier.de/~ley/db/ [accessed: 2014-12-09].

[29] The Pennsylvania State University, "CiteSeer," 2014, URL: http://citeseerx.ist.psu.edu/ [accessed: 2014-12-09].

[30] Georg-August-Universität Göttingen Niedersächsische Staats- und Universitätsbibliothek Göttingen, "Open Access," 2013, URL: http://open-access.net/ [accessed: 2014-12-09].