STAG: Smart Tools and Applications in Graphics (2019)
M. Agus, M. Corsini and R. Pintus (Editors)

A visualization tool for scholarly data

M. Salinas[1], D. Giorgi[2], F. Ponchio[2] and P. Cignoni[2]

[1]University of Pisa, Italy
[2]National Research Council of Italy, Institute of Information Science and Technologies, Italy

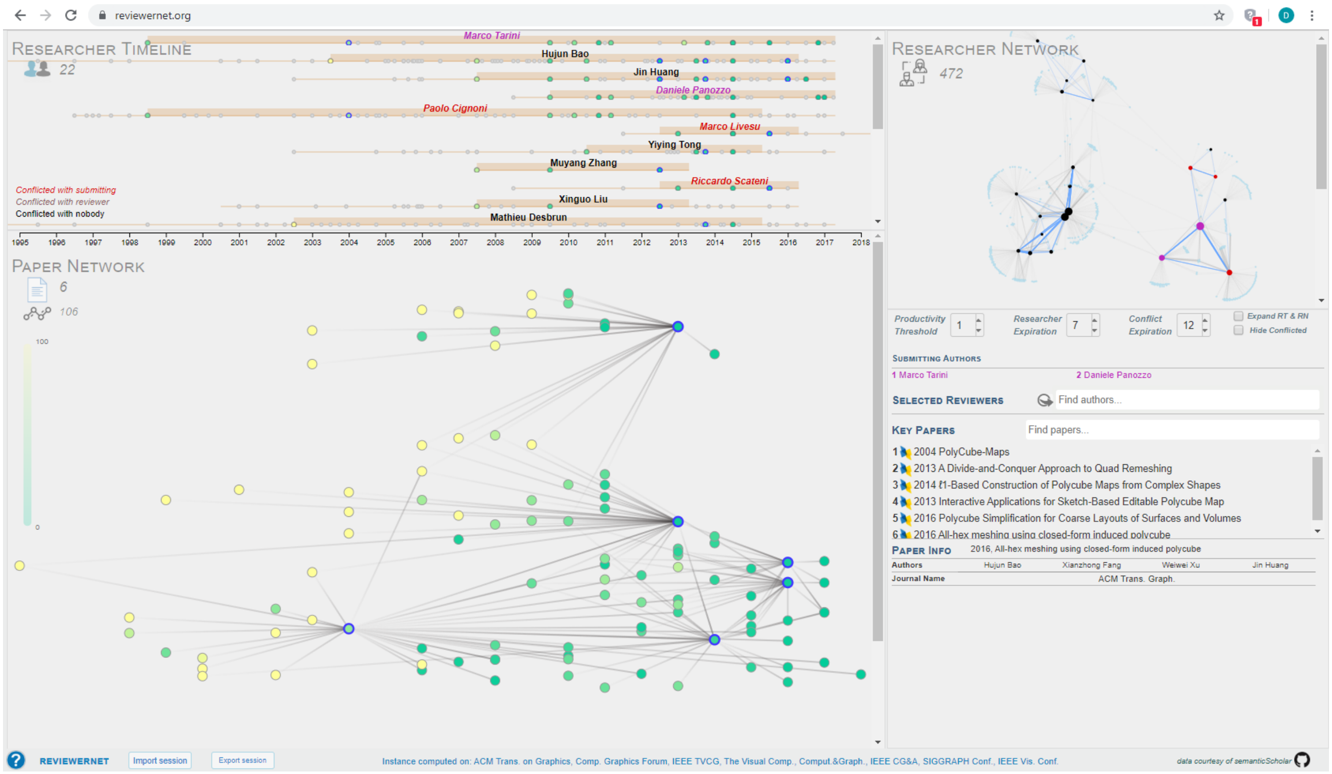


Figure 1: *ReviewerNet is a visualization system that allows one to identify researchers working on a certain topic, to analyse their contributions over time, and to get aware of co-authorship relations and conflicts. Therefore, ReviewerNet can support the reviewer selection process in the academic domain. The interface of ReviewerNet is divided into four main areas: the* RESEARCHER TIMELINE *(top left); the* PAPER NETWORK *(bottom left); the* RESEARCHER NETWORK *(top right); the* CONTROL PANEL *(bottom right).*

Abstract

We propose ReviewerNet, an online, interactive visualization system aimed to improve the reviewer selection process in the academic domain. Given a paper submitted for publication, we assume that good candidate reviewers can be chosen among the authors of a small set of pertinent papers; ReviewerNet supports the construction of such set of papers, by visualizing and exploring a literature citation network. Then, the system helps to select reviewers that are both well distributed in the scientific community and that do not have any conflict-of-interest, by visualising the careers and co-authorship relations of candidate reviewers. The system is publicly available, and is demonstrated in the field of Computer Graphics.

CCS Concepts

• ***Computing methodologies*** → *Graphics systems and interfaces;*

DOI: 10.2312/stag.20191359

1. Introduction

The number of digital academic documents, either newly published papers or documents resulting from digitization efforts, grows at a very fast pace: the Scopus digital repository counts more than 70 million documents and 16 million author profiles [sco18]; the Web of Science platform has more than 155 million records from over 34,000 journals [WoS]; Microsoft Academic collects about 210 million publications [MA]. In 2018, over four thousand new records were added to DBLP [DBL], and bibliometric analysts estimated a doubling of global scientific output roughly every nine years [BM15]. Therefore, the volume, variety and velocity of scholarly documents generated satisfies the big data definition, so that we can now talk of *big scholarly data* [KLSA17].

Sensemaking in this huge reservoir of data calls for platforms adding an element of automation to standard procedures – such as literature search, expert finding, or collaborators discovery – to reduce the time and effort spent by scholars and researchers. In particular, there has been an increase in the number of visual approaches supporting the analysis of scholarly data. Visualization techniques were proposed to help stakeholders to get a general understanding of sets of documents, to navigate them, and to find patterns in publications and citations. Federico et al. [FHKM17] survey about 109 visual approaches for analysing scientific literature and patents published in-between 1991 and 2016. Most of the works focused on the visualization of document collections and citation networks. A more ambitious goal for visualization platforms would be to enable users get enough understanding to make decisions.

In this paper, we focus on the problem of reviewer finding by journal editors or International Program Committee (IPC) members, who are required to search for reviewers who know well a subject, yet are not conflicted with the authors of the paper under scrutiny. Finding good candidate reviewers requires to analyse topic coverage (possibly over time), stage of career, and past and ongoing collaborations. Every member of the community has its own approach to reviewer finding, which usually involves bibliographic research, and frequent visits to public repositories like DBLP [Ley02] and researchers' home pages. In any case, one has to confront possibly large collections of data to make decisions, and a user may easily get lost after following a few links.

We propose ReviewerNet, a visualization platform which facilitates the selection of reviewers. The intuition behind ReviewerNet is that the authors of relevant papers are good candidate reviewers. ReviewerNet offers an interactive visualization of multiple, coordinated views about papers and researchers that help assessing the expertise and conflict of interest of candidate reviewers (Figure 1). One of the main advantages of ReviewerNet is that it only relies on citations, to analyse the literature, and on co-authorship relations, to analyse conflicts. Citations are an essential part of research: they represent a credible source of information about topic similarity and intellectual influence. Moreover, since citations have author-chosen reliability, they are a very robust cue to relatedness. Similar reasonings hold for co-authorship relations. Therefore, an important contribution is the demonstration that a well-combined visualization based on citation and co-authorship relations only can support the reviewer search process, without the need for more complicated content analysis techniques.

ReviewerNet builds on a reference database including papers, authors and citations from selected sources (journal articles and conference papers) taken from the Semantic Scholar Research Corpus [AGB*18]. We demonstrate the platform usage in the field of Computer Graphics, with a reference dataset containing 17.754 papers, 108.155 citations, 23386 authors. We show how ReviewerNet can be used to search for reviewers who are expert on a certain topic, are at a certain career stage, who have a certain track of publishing records, who are not conflicting with neither the submitters nor other reviewers, and who are well-distributed in the scientific community. The tool is free to use and open source; the source code is available at https://github.com/cnr-isti-vclab/ReviewerNet, while the demonstration platform is available at https://reviewernet.org/.

2. Related work

Concerning the reviewer selection process, the literature mostly focused on the automatic reviewer *assignment* task, which is a different problem than ours. Indeed, the reviewer assignment problem requires finding the best assignment between a finite set of reviewers (e.g., the members of the Programme Committee of a conference) and a finite set of papers (the papers submitted to the conference). This is usually done using bi-partite graph matching and taking into account pertinence of the reviewers with the papers and fair distribution of loads; [WSC10] provides an overview of this problem. In what follows, we briefly review the state-of-the art about the search, analysis and recommendation services offered by scholarly data platforms, and the visualization of bibliometric networks.

Scholarly data platforms Many applications have been developed on top of the big scholarly data platforms to search for authors, documents, venues, and analyse statistics about for example distribution per research area, citations, and other bibliometric indices. Most academic search engines also provide research paper recommendations according to one's research interests. Microsoft Academic provides a semantic search engine that employs natural language processing and semantic inference to retrieve the documents of interest [SSS*15]. It also provides related information about the most relevant authors, institutions, and research areas. Scopus enables one to search for authors or documents, track citations over time for authors or documents, view statistics about an author's publishing output, and compare journals according to different bibliometric indices [sco18]. These and similar applications offer basic functionalities and static visualizations which researchers do use while looking for reviewers. Though, none of them offers an integrated service to support the higher level tasks of fine-tuned reviewer selection, where both expertise and conflicts of interest have to be taken into account.

Visualization of bibliometric networks The visualization of bibliometric networks is an active area of research [Che13, FHKM17]. Bibliometric networks include citation, co-citation, co-authorship, bibliographic coupling and keyword co-occurrence networks [IIS*17, HSZ13]. Concerning visualization of citations, most part of the literature focused on co-citation and bibliographic coupling networks, rather than on direct citations. One of the first visualization of citation networks is Garfield's historiography [GPI03], a node-link diagram where citation links are directed backwards in time. Garfield and colleagues underline how citation networks enable one to analyse the history and development of research fields. CiteNetExplorer [vEW14] is a software tool to visualize citation networks which builds on Garfield and colleagues' work: it improves the graph layout optimization to handle a larger number of papers, and offers network drill-down and expansion functionalities. PaperVis [CY11] is an exploration tool for literature review, which adopts modified Radial Space Filling and Bullseye View techniques to arrange papers as a node-link graph while saving the screen space, and categorizes papers into semantically meaningful hierarchies. [GLK*13] describes a visual analytics system for exploring and understanding document collections, based on computational text analysis; it supports document summarization, similarity, clustering and sentiment analysis, and offers recommendations on related entities for further examination. Rexplore [OMM13] is a web-based system for search and faceted browsing of publications. Rexplore also includes a graph connecting similar authors, where similarity depends on research topics as extracted from document text. At any rate, using keywords as proxies for research topics can be noisy. Therefore, in ReviewerNet we only rely on co-authorship relations. Finally, we observe that many of the approaches for bibliographic network visualization make limited use of user interaction, and often use a loose coupling of views [FHKM17]. With ReviewerNet, we propose an integrated environment which facilitates a high-level task (reviewer discovery and selection) by means of coordinated, interactive views.

3. ReviewerNet description

ReviewerNet supports the various actions that journal editors and IPC members perform while choosing reviewers, namely, searching the literature about the submission topic, looking for active experts in the field, and checking their conflicts of interest. ReviewerNet does so by integrating an overview visualization of the literature with a visualization of the career of potential reviewers, their conflicts of interests, and their nets of collaborators. This combined visualization helps to make sense of scholarly data, and rapidly get enough understanding to make a sensible decision. In detail, ReviewerNet integrates the visualization of three main classes of data in a single window (Figure 1):

Paper Network (PN): a chronologically ordered graph visualization of the literature related with the submission topic. The nodes represent papers, while arcs represent in- and out-citation relations between papers. The horizontal dimension represents time. By means of interactive graph expansion functionalities, the PN supports the rapid exploration of key papers in the literature with respect to the topic of the submitted paper. The authors of the key papers identified will define the set of the candidate reviewers. The PN is built by the users, starting from a small number of seed papers of their choice;

Researcher Timeline (RT): a time-based visualization of the academic career of researchers, through horizontal lines (spanning their overall career) and bars (highlighting the years in which they focused on the submission topic). The RT helps assessing the suitability of potential reviewers, showing thier topic coverage, productivity over years, and stage of career. Also, visual cues help the user to tell apart candidate reviewers from conflicting researchers. The RT is built automatically by ReviewerNet while the user builds the PN;

Researcher Network (RN): a graph visualization of co-authorship relations: the nodes represent the authors in the PN and their collaborators in the dataset; the arcs connect authors who have publications in common. The aim of the RN is to visualize the research communities: indeed, the identification of network of collaborators helps looking for sets of independent, non-conflicting reviewers. As with the RT, the RN is built online by ReviewerNet.

The basic pipeline for finding reviewers with ReviewerNet involves building the Paper Network starting from a small set of seed documents; evaluating possible choices of reviewers, by navigating the Reviewer Timeline and the Reviewer Network; and finally obtaining a justified list of chosen reviewers, along with possible substitutes suggested by ReviewerNet in case of decline. The user can navigate the different views and interact with the system through simple actions, to drive his/her investigation. Each view in ReviewerNet is linked to the other views, so that so that any action in a view is reflected in the others. Visual cues are used to improve the comprehension during interactive sessions: the colour, size, boundary, and style of visual elements visually represent important characteristics of the entities they stand for. To better explain how ReviewerNet works, the next Section presents an example user scenario. In addition, since a static description may not adequately convey the dynamic nature of the investigation with ReviewerNet, we refer the reader to the accompanying video at `https://www.youtube.com/watch?v=JnomPO8QI28`, which illustrates the scenario described below. Finally, the technical details about the system are reported in Section 4.

3.1. Usage scenario

We introduce Robert, a fictitious academic researcher. Robert is in the IPC of a conference in the field of Computer Graphics; he is the primary reviewer for a paper, and he is in charge of finding three additional reviewers, plus alternative reviewers in case of decline.

Data collection To construct the reference dataset for this scenario, we collected papers, authors and citations from eight selected sources in the field of Computer Graphics, taken from the Semantic Scholar Research Corpus [AGB*18]. The dataset includes data from the journals and conference proceedings listed in Table 1, spanning the years in-between 1995 and 2018. After an automatic cleaning steps to remove non-papers (such as acknowledgments to reviewers, prefaces, etc.), the final reference dataset contains 17.754 papers, 108.155 citations, and 23386 authors.

Starting ReviewerNet Robert is in charge of finding reviewers for a paper about polycube maps, authored by Marco Tarini and Daniele Panozzo. In the Control Panel area, he inputs their names in the *Submitting Authors* field, also with the help of a drop-down menu. The authors are now shown in the Researcher Timeline and the Reviewer Network, marked as purple, and the rest of the interface becomes active.

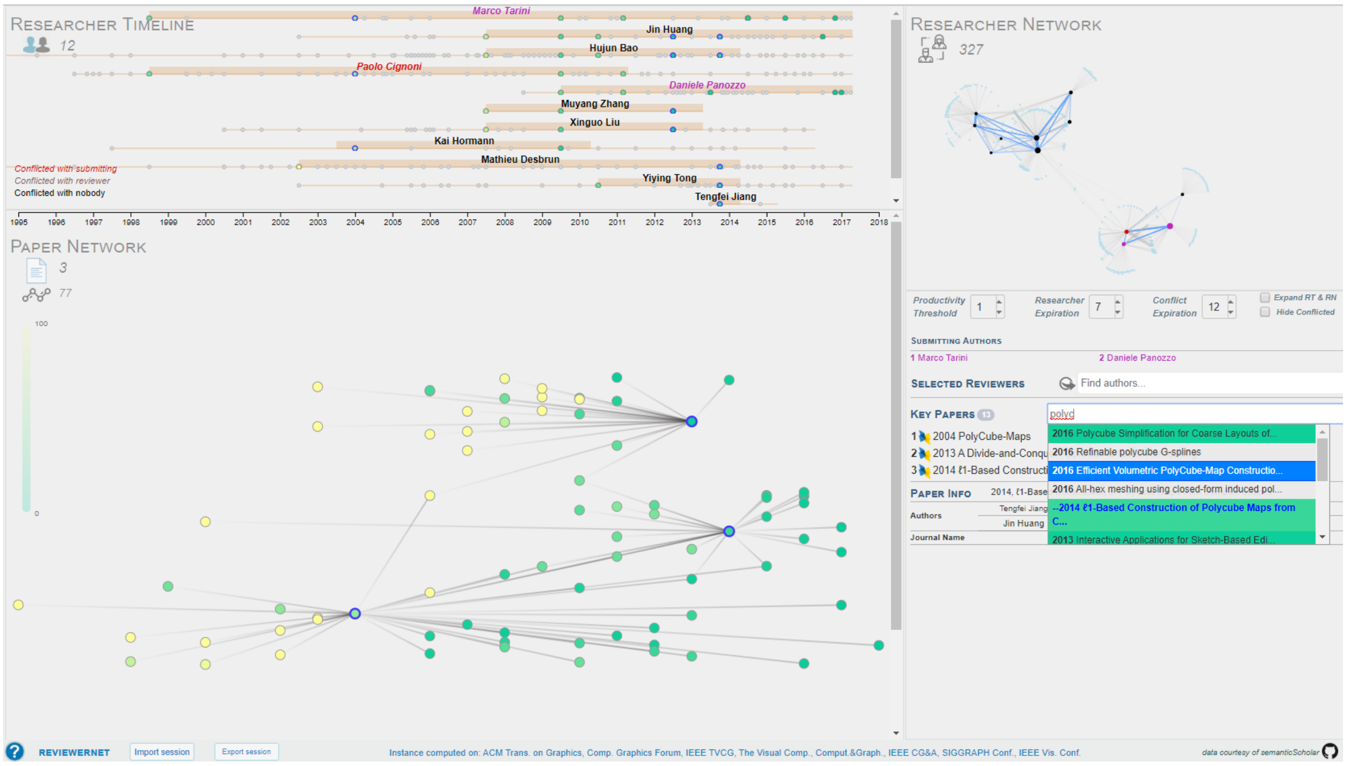


Figure 2: *When the user inputs the seed papers (bottom right), ReviewerNet starts building the Paper Network (bottom left), the Researcher Timeline (top left), and the Researcher Network (top right). The dots representing papers in the Paper Network and the Researcher Timeline are colored according to their citation count, from green (few citations) to yellow (many citations). Grey dots in the Researcher Timeline are papers in the reference database, but not included in the current Paper Network. Selected papers are circled in blue.*

Table 1: *The selected sources from the Semantic Scholar Research Corpus used in our demonstration scenario, and the number of papers for each source. The final reference dataset contains 17.754 papers, 108.155 citations, and 23386 authors.*

ACM Trans. on Graphics	2833
Computer Graphics and Applications	1983
Computer Graphics Forum	3238
Computers & Graphics	2155
IEEE Trans. on Visualization and Computer Graphics	3236
Visual Computer	2107
Proc. IEEE Conference Visualization (pre 2006)	501
Proc. ACM SIGGRAPH (pre 2003)	1701

Building the Paper Network The first step is to build the Paper Network (Figure 2), that is, a set of key papers which are relevant to the submission topic. Later on, Robert will chose his reviewers among the authors of those key papers. Robert thinks of a first set of three documents about polycube maps, which serve as seeds for building the network (*PolyCube-Maps*, 2004; *A divide-and-conquer approach for automatic polycube maps construction*, 2009; *L_1-based construction of polycube maps from complex shapes*, 2014). He inputs their titles in the *Key papers* field. His knowledge of the domain helps him in this initial step, though he can also take advantage of title-based suggestions, which are shown in a drop-down menu, listed by publication year. The three papers are now included in the Paper Network, along with their in- and out-citations. While Robert builds his Paper Network, ReviewerNet automatically adds the authors of selected papers in the Researcher Timeline and the Researcher Network, as candidate reviewers.

Robert can now expand the Paper Network, to discover additional documents and therefore additional candidate reviewers. With a double click, he selects interesting nodes, i.e., papers he deems relevant to polycube maps. The Paper Network then updates with the in- and out-citations of the selected papers, so that Robert can further explore the literature. Robert navigates the network, and decides to reduce its size by deselecting a paper he realizes he is no longer interested in, because its citations suggest it addresses a different topic than the submission. Robert continues until he feels the selected papers and their citations offer a good coverage of the literature about the topic at hand. Robert checks the paper details, including the link to the respective DBLP page, shown in the bottom right corner of the interface. A quick keyword search with *polycube maps* in the *Key papers* field let him notice that there is an important paper he was missing (*Efficient volumetric poly-cube*

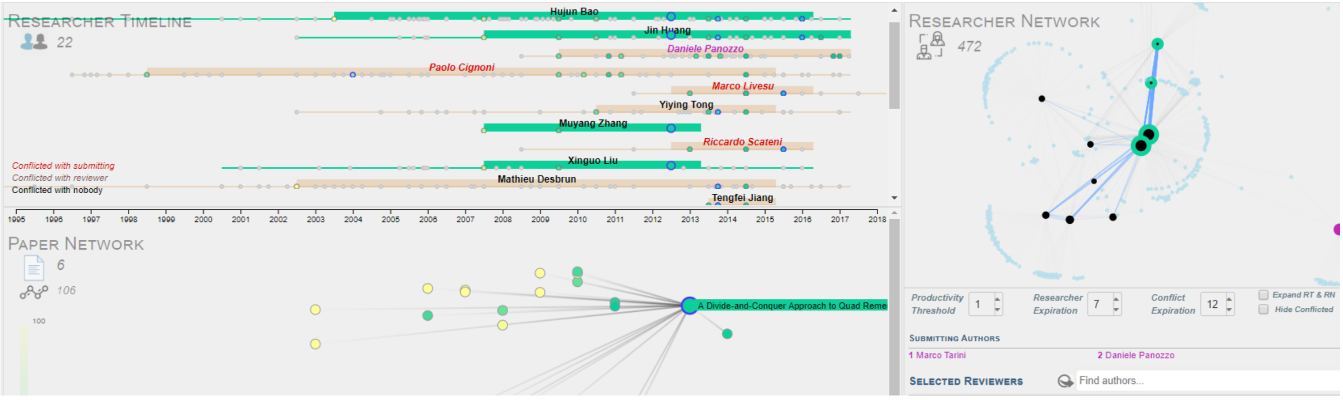


Figure 3: *In the Researcher Timeline, researchers are represented as horizontal lines, spanning their academic career; the bars over the lines indicate the years in which the authors published about the submission topic, namely, the years for which they have papers in the Paper Network. When hovering over an entity representing a paper, the authors of that paper are highlighted as green in the other views.*

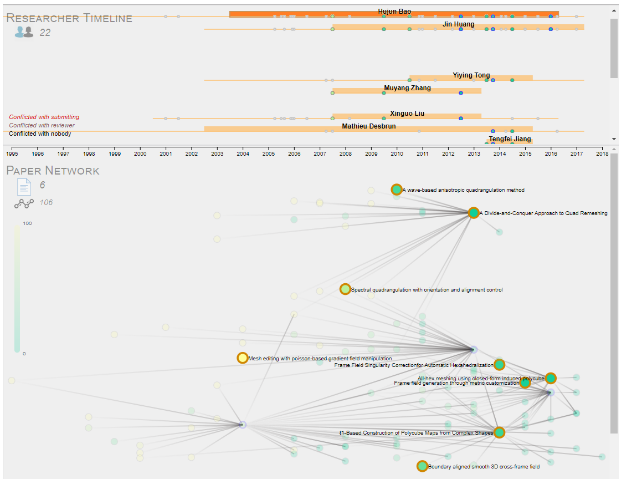


Figure 4: *Focusing on a researcher by clicking on her/his name in the Researcher Timeline allows one to highlight her/his co-authors and production in the Paper Network.*

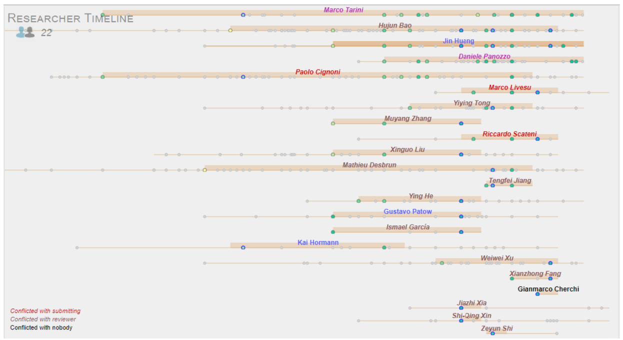


Figure 5: *The Researcher Timeline lists the names of potential reviewers. The name coloring emphasizes the distinction among roles: submitting authors (marked as purple), their co-authors (red), selected reviewers (blue), their co-authors (brown), and non-conflicting, candidate reviewers (black). The font style of names further helps to tell apart conflicting researchers (italic) from non-conflicting candidate reviewers (normal). The researchers are ordered vertically according to their relevance (cf. Section 4 for details).*

map construction, 2016); the paper can be easily told apart from papers already in the network, thanks to visual cues in the drop-down menu. Finally, the selection of 6 papers produces a list of 22 candidate reviewers.

Exploring the Researcher Timeline and the Researcher Network Robert now explores the Researcher Timeline to assess the suitability of candidate reviewers. In the Researcher Timeline, researchers are represented as horizontal lines, spanning their academic career. Robert checks the expertise of candidate reviewers by looking at their stage of career, and production over years. Since each view is linked to the other views, Robert checks topic coverage by looking at who published what, by hovering the mouse over papers to highlight their authors in all the views (Figure 3). With a mouse click on a researcher, ReviewerNet highlights both his/her co-authors and papers (Figure 4). While looking for candidate reviewers, Robert can always checks conflicts of interests, thanks to colours and font style (Figure 5).

The visualization also helps Robert analysing the network of collaborators of candidate reviewers. This is fundamental to find sets of independent, well distributed reviewers. Robert can navigate the Researcher Network, a graph visualization of co-authorship relations among the candidate reviewers and their collaborators in the dataset. He pans and zooms and uses the different handlers available to discover the communities of collaborators. He founds that there are three distinct groups of collaborators dealing with the topic at hand (Figure 6).

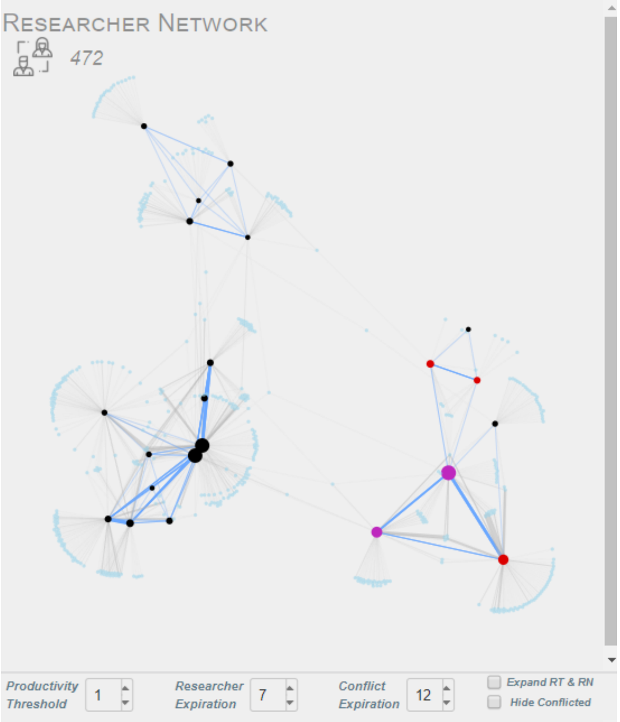
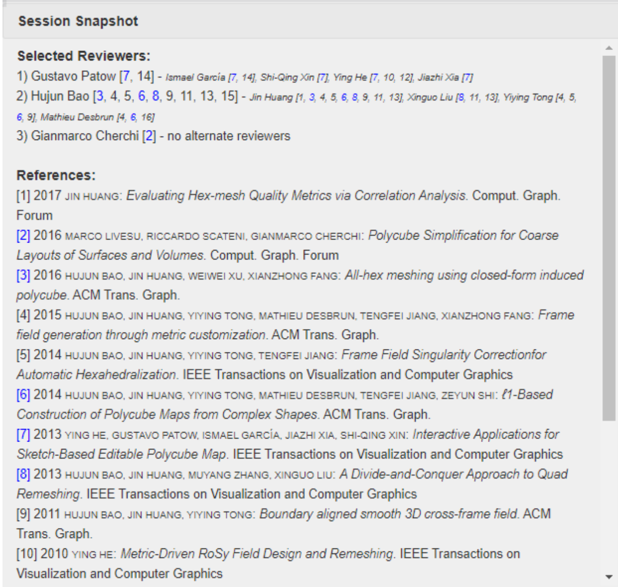


Figure 6: *In the Researcher Network, arcs connect authors who have publications in common. Arcs are blue when the co-authored papers include a selected paper. The coloring of nodes emphasizes roles as in the Researcher Timeline, while the relevance of authors is rendered through the dimension of nodes. The tickness of arcs renders the number of co-authored papers. The visualization can be fine-tuned by adjusting a set of parameters defining the criteria on productivity to be included in the visualization, or the criteria that define conflicts (cf. Section 4 for their definition).*


Session Snapshot

Selected Reviewers:
1) Gustavo Patow [7, 14] - *Ismael García [7, 14], Shi-Qing Xin [7], Ying He [7, 10, 12], Jiazhi Xia [7]*
2) Hujun Bao [3, 4, 5, 6, 8, 9, 11, 13, 15] - *Jin Huang [1, 3, 4, 5, 6, 8, 9, 11, 13], Xinguo Liu [8, 11, 13], Yiying Tong [4, 5, 6, 9], Mathieu Desbrun [4, 6, 16]*
3) Gianmarco Cherchi [2] - no alternate reviewers

References:
[1] 2017 JIN HUANG: *Evaluating Hex-mesh Quality Metrics via Correlation Analysis.* Comput. Graph. Forum
[2] 2016 MARCO LIVESU, RICCARDO SCATENI, GIANMARCO CHERCHI: *Polycube Simplification for Coarse Layouts of Surfaces and Volumes.* Comput. Graph. Forum
[3] 2016 HUJUN BAO, JIN HUANG, WEIWEI XU, XIANZHONG FANG: *All-hex meshing using closed-form induced polycube.* ACM Trans. Graph.
[4] 2015 HUJUN BAO, JIN HUANG, YIYING TONG, MATHIEU DESBRUN, TENGFEI JIANG, XIANZHONG FANG: *Frame field generation through metric customization.* ACM Trans. Graph.
[5] 2014 HUJUN BAO, JIN HUANG, YIYING TONG, TENGFEI JIANG: *Frame Field Singularity Correctionfor Automatic Hexahedralization.* IEEE Transactions on Visualization and Computer Graphics
[6] 2014 HUJUN BAO, JIN HUANG, YIYING TONG, MATHIEU DESBRUN, TENGFEI JIANG, ZEYUN SHI: *ℓ1-Based Construction of Polycube Maps from Complex Shapes.* ACM Trans. Graph.
[7] 2013 YING HE, GUSTAVO PATOW, ISMAEL GARCÍA, JIAZHI XIA, SHI-QING XIN: *Interactive Applications for Sketch-Based Editable Polycube Map.* IEEE Transactions on Visualization and Computer Graphics
[8] 2013 HUJUN BAO, JIN HUANG, MUYANG ZHANG, XINGUO LIU: *A Divide-and-Conquer Approach to Quad Remeshing.* IEEE Transactions on Visualization and Computer Graphics
[9] 2011 HUJUN BAO, JIN HUANG, YIYING TONG: *Boundary aligned smooth 3D cross-frame field.* ACM Trans. Graph.
[10] 2010 YING HE: *Metric-Driven RoSy Field Design and Remeshing.* IEEE Transactions on Visualization and Computer Graphics

Figure 7: *The list of selected reviewers, together with substitutes in case of decline, and a bibliography. A substitute reviewer is a researcher who has authored a similar set of publications and has the same conflicts as the original reviewer. The bibliography motivates the reviewer selection, since it lists, for each selected reviewer, the papers he/she has authored.*

Selecting reviewers Once Robert identifies one or more candidate reviewers, he inputs their names in the *Selected Reviewers* field (also with the help of the drop-down menu). He decides to chose Gustavo Patrow, whose expertise fits with his requirements. The colouring of the selected reviewer switches to blue both in the Researcher Timeline and the Researcher Network, and the colouring of his co-authors switches to grey, to identify them as conflicting potential reviewers, and tell them apart from the remaining available candidates. Then, Robert selects Hujun Bao, a senior researcher, and Gianmarco Cherchi, a younger researcher who belongs to a different community than the previous two, and has been working very recently on the subject at hand. The icon beside the reviewers name links to their respective DBLP pages, so that Robert can further check about conflicts of interest possibly deriving from the co-authorship of papers published on venues not included in the dataset.

Robert downloads his list of three reviewers with a click on the download button. The list reports reviewers' names and bibliographic references to their papers (Figure 7). After contacting the reviewers, Robert finds that one of them declines his invitation. Fortunately, for each reviewer selected by Robert, ReviewerNet has automatically added a list of potential alternative reviewers, in case of a negative answer by the original reviewer. Alternative reviewers are chosen from the candidate ones, so that they only conflict with the declining reviewer. Robert evaluates possible substitutes, again taking advantage of ReviewerNet functionalities, and finds his best replacement.

Discussion This abreviated scenario shows how ReviewerNet can support investigating the literature, learning who are the experts in a field, and exploring relationships among them. The description above necessarily simplified a typical intercation process: Robert could of course switch back and forth between different tasks; as the coherence of visual cues across different views enforces their meaningfulness, it is easy for him to switch between different views without losing focus. Robert could have also refined the Paper Network after having examined the list of candidate reviewers. He could have adjusted the size of the list by fine tuning the optional parameters. The process is iterative in nature, and the desiderata may evolve as the search proceeds. Thanks to the user-friendly interface which leaves the user control over the process, ReviewerNet enables the user to narrow down as well as widen the scope of analysis. In turn, the combined visualization of different aspects of the problem at hand well supports the decision making process.

4. Technical details

In ReviewerNet, the visualized data pertain to three types of entities: *papers*, *researchers*, and *citations*. The data attributes are both quantitative and qualitative, and the time dimension is central.

Concerning papers, let $\mathcal{P}$ denote the set of papers in a reference dataset, and let $\mathcal{P}_V \subseteq \mathcal{P}$ be the set of papers relevant to a submission. $\mathcal{P}_V$ is built by the users starting from a small number of seed papers of their choice. A paper $p \in \mathcal{P}_V$ is marked as *selected*, if it is considered as a key paper by the user; we denote by $\mathcal{P}_S$ the set of selected papers, with $\mathcal{P}_S \subseteq \mathcal{P}_V \subseteq \mathcal{P}$. The attributes of a paper which are visualized are standard *bibliographic attributes* (title, authors, publication year, venue) and its *citation count*: if $\mathcal{C}(p)$ is the set of papers citing p, the *citation count* $c(p)$ is its cardinality: $c(p) = |\mathcal{C}(p)|$. Papers are related through *direct citations*.

Concerning researchers, let $\mathcal{A}(p)$ be the set of authors of a given paper p, and $\mathcal{R}$ the set of authors of papers in $\mathcal{P}$. Then, the set $\mathcal{R}_C \subseteq \mathcal{R}$ of *candidate reviewers* is given by the set of researchers who authored a selected paper:

$$\mathcal{R}_C = \{r \in \mathcal{R} \; s.t. \; \exists \; p \in \mathcal{P}_S : r \in \mathcal{A}(p)\}$$

Researchers have two attributes in ReviewerNet: relevance and conflict of interest. We define a researcher's *relevance* as a reviewer according to the authorship of relevant papers. The concept of relevance can be tuned according to the user needs (e.g., looking for highly-specialized reviewers, as opposed to generalists). For a candidate reviewer r, let $\mathcal{P}_S|_r$ be the set of papers in $\mathcal{P}_S$ authored by r. Then, the *relevance score* $s(r)$ of the candidate reviewer r is defined as a weighted sum of the number of selected and non-selected papers in $\mathcal{P}_V$ authored by r:

$$s(r) = \alpha|\mathcal{P}_S|_r| + \beta|\{\mathcal{P}_V - \mathcal{P}_S\}|_r|$$

with α and β real-valued coefficients summing up to one. We set $\alpha = 0.7$ and $\beta = 0.3$ as default parameters. The set of candidate reviewers will be visualized in the Researcher Timeline in order of their relevance; relevance will also define the dimension of nodes in the Researcher Network.

The second attribute of researchers is their *conflict of interest*, with either the submitting authors or other reviewers. We model the conflict of interest after *co-authorship* relations: two researchers have a conflict of interest if they have papers in common. In what follows $\mathcal{CA}(r)$ denotes the set of co-authors of a researcher r, or, in other words, the set of researchers who have a conflict with him/her. We let the degree of conflict, and hence the availability as a reviewer, be modulated according to the number of papers in common, and the years passed since the last co-authored paper, again according to the user intent or the specific policies of conferences and journals.

4.1. User interface

The visual composition of the four regions in the interface (Figure 1) helps the user to gain different perspectives on the problem at hand, within a single visualization. Each region is resizable in height.

The nodes in the Paper Network (PN), at the bottom-left hand side of the screen, represent papers in $\mathcal{P}_V$, while the arcs represent in- and out-citation relations between them. Papers are ordered horizontally according to their publication year, while a force-directed graph drawing algorithm determines the layout in the vertical direction [BOH11].

Each line in the Researcher Timeline (RT), at the upper-left side of the screen, represents a candidate reviewer r in $\mathcal{R}_C$, that is, the author of a selected paper in $\mathcal{P}_S$. The dots over the line represent the set $\mathcal{P}|_r$ of papers authored by r in the reference database $\mathcal{P}$.

The nodes in the Researcher Network (RN), at the upper-right hand side of the screen, are the researchers in $\mathcal{R}_V$ along with their collaborators in $\mathcal{R}$. The arcs connect authors who have publications in common: for each node representing a researcher r, the node degree is the cardinality $|\mathcal{CA}(r)|$. A force-directed graph drawing algorithm determines the graph layout. Both the Researcher Timeline and the Researcher Network are built automatically by ReviewerNet while the user builds the Paper Network.

The Control Panel (CP), at the bottom-right hand side of the screen, allows the user to input and manage the names of submitting authors, the names of selected reviewers, and the titles of key papers. The CP area also displays information about papers, upon request. The DBLP icon beside reviewers' names and paper titles links to their respective DBLP page. Moreover, the CP includes parameters boxes and checkboxes to fine-tune the visualization:

Size of data visualized: To limit the number of candidate reviewers visualized in the RT and the RN, the user can set two thresholds a researcher has to meet to be considered as a candidate reviewer:

- **Productivity threshold:** the minimum number of authored selected papers in $\mathcal{P}_S$ (i.e., $|\mathcal{P}_S|_r|$ has to be greater than the threshold, for a researcher r to be included in the set $\mathcal{R}_C$ of candidate reviewers);
- **Researcher expiration:** the maximum number of years since the last authored paper in the reference dataset $\mathcal{P}$ (i.e., the number of years has to be lower than the threshold for a researcher to be considered active and included in $\mathcal{R}_C$).

The user can also remove conflicting authors and their co-authors from the visualization, by ticking the *Hide Conflicted* checkbox. To augment instead the number of potential reviewers visualized, the user can tick the *Expand RT & RN* checkbox: the visualization will include all the researchers in $\mathcal{R}_V$ (all the authors of relevant papers) instead of the researchers in $\mathcal{R}_C$ only (the authors of selected papers only). Note that visualizing a large number of researchers can slow down the interface.

Conflict-of-interest: Finally, to modulate the conflict of interest, the user can set a threshold for two researchers to be considered as co-authors, namely:

- **Conflict expiration:** the maximum number of years since the last co-authored paper in $\mathcal{P}$.

A larger threshold will increase the number of candidates marked as conflicted. Conversely, a smaller threshold will increase the number of available reviewers.

Finally, the CP enables the user to download the list of selected reviewers, along with references to the reviewers' publications in the dataset $\mathcal{P}$. The list also includes substitute reviewers suggested by ReviewerNet, in case of negative answers from the selected one. For a selected reviewer r, the alternative reviewers are chosen in the set $\mathcal{R}_C$ of candidate reviewers, so that they only conflict with r, and with no other selected reviewer. The list of substitutes is ordered according to the number of common papers between the reviewer and his/her substitute.

4.2. Visual consistency

Visual cues include the colour, size, boundary, and style of visual elements representing papers, researchers and their relations across the different views.

Visual cues for papers For a paper $p \in \mathcal{P}_V$, the color corresponds to the citation count $c(p)$, from yellow (few citations) to green (many citations). This colormap applies to both nodes in the PN and dots in the RT. Dots corresponding to papers in $\mathcal{P} - \mathcal{P}_V$ (papers in the reference database, but not included the PN) are marked as grey. Selected papers in $\mathcal{P}_S$ are circled in blue, both in the PN and the RT. Arcs are blue in the RN when the co-authored papers include a selected paper.

Visual cues for researchers For researchers in the RT, the name coloring emphasizes the distinction between roles: submitting authors (marked as purple), their co-authors (red), selected reviewers (blue), their co-authors (brown), and non-conflicting, candidate reviewers (black). The nodes in the RN corresponding to researchers in the RT follow the same rule, whereas nodes representing their co-authors in $\mathcal{R}$ are light blue. For researchers in the RT, the font style of names further helps to tell apart conflicting researchers (italic) from non-conflicting candidate reviewers (normal). The same colour/font rules apply to the names suggested in the selected reviewers' drop-down menu in the CP. The researchers in the RT are ordered vertically according to their relevance score $r(s)$. The same score is rendered in the RN through the dimension of nodes.

4.3. Actions

Each view (PN, RT, RN, CP) is linked to the other views, so that any action in a view is reflected in the others.

Actions on Papers The user inizializes the Paper Network by inputing the titles of a small set of seed papers in the *Key papers* field, with the help of title-based suggestions. The seed papers are visualized in the PN, along with their in- and out-citations. The user can now expand the network, to discover additional documents. With a double click, he selects interesting nodes, i.e., papers he/she deems relevant to the submission topic. The PN then updates with the in- and out-citations of the selected papers. Papers can be deselected with a double click.

When the users focuses on a paper in one of the views by mouse hovering, the same paper is highlighted in the other views. For example, when hovering the mouse over a node in the PN, the corresponding dot in the Researcher Timeline is highlighted, and viceversa. Also, the paper details (title, publication year, venue) are shown in the Control Panel on a mouse click. Likewise, by hovering over or clicking on the title in the CP, the corresponding node and dot are highlighted in the PN and the RT. When hovering the mouse over an entity representing a paper (a node in the PN, a dot in the RT bars, the title in the CP), the paper authors are highlighted in the RT and RN, if present. A mouse click on the focused paper lets the user navigate the visualization with highlighted items. A single click restores the previous visualization. The icon beside paper titles in the Control Panel links to DBLP pages.

Actions on Researchers In a similar fashion to papers, when the user focuses on a researcher in one of the views by mouse hovering, the same researcher is highlighted in the other views. When hovering the mouse over a node in the Researcher Network, the name of the corresponding researcher appears on the upper-right corner. When hovering the mouse over an entity representing a researcher (a bar in the RT, a dot in the RN, the name in the CP), the papers authored by the researcher are highlighted in the Paper Network view. A mouse click on a researcher puts the focus on him/her, his/her production and his/her personal net of collaborators. The user can navigate a visualization with selected items and additional functionalities. Only the set of co-authors is visualized in the Researcher Timeline and the Researcher Network. While hovering on one of the co-authors, the common publications are shown in the PN, and the arc representing the co-authorship relation is visualized in the RN. Another mouse click will get the user back the previous visualization. When hovering the mouse over an arc in the RT, a pop-up on the upper-right corner shows the pair of co-authors names, the number of common papers in the dataset $\mathcal{P}$, and the number of common relevant papers in $\mathcal{P}_V$. In turn, for blue arcs, the common papers are highlighted in the PN.

The icon beside the researcher name in any of the fields in the Control Panel links to the DBLP page of that researcher. A researcher can be removed from the list of selected reviewers with a double click. The user can exchange a reviewer with one of his/her substitutes by clicking on the name of the substitute. The export button enables the user to download the list of reviewers and their potential substitutes. Work sessions can be saved for later re-use and re-assessment.

5. Conclusions

We have presented ReviewerNet, a novel system for choosing reviewers by visually exploring scholarly data. ReviewerNet enables scientific journal editors and members of IPCs to search the literature about the topic of a submitted paper, to identify experts in the field and evaluate their stage of career, and to check possible connections with the submitting authors and among the reviewers themselves. This helps to avoid conflicts and to build a fairly distributed pool of reviewers. To do so, ReviewerNet features a combined visualization of the literature, the career of potential reviewers, their conflict of interests, and their nets of collaborators. Interestingly enough, the system is able to help the process even without exploiting any content-based analysis of the papers.

A fisrt evaluation of the demonstration platform with both in-house testers and members from the Computer Graphics community confirmed that the users were able to get acquainted with the system even with a very limited training, and appreciated the different functionalities of ReviewerNet and its capability of improving the reviewer search process. Some of the users also highlighted the potential of ReviewerNet as a tool to support bibliographic research, besides the reviewer selection process.

The evaluation also highlighted that there is room for improving the system. In particular, we are working at the automation of the choice of the key papers, whose manual insertion was signalled as a weakness by some users. Our strategy will be to import the bibliography of papers from pdf files. Also, we are working at a user-friendly procedure to generate instances of the platform with customizable data coverage. Indeed, some of the testers commented on the choice of venues on the demonstration platform. Therefore, in the future ReviewerNet will include a functionality to automatically build the dataset, by selecting a list of venues of interest. Once the platform development will be complete, we will carry out a formal user study with end users.

References

[AGB*18] Ammar W., Groeneveld D., Bhagavatula C., Beltagy I., Crawford M., Downey D., Dunkelberger J., Elgohary A., Feldman S., Ha V., Kinney R., Kohlmeier S., Lo K., Murray T., Ooi H.-H., Peters M., Power J., Skjonsberg S., Wang L. L., Wilhelm C., Yuan Z., van Zuylen M., Etzioni O.: Construction of the literature graph in semantic scholar. In *NAACL* (2018). 2, 3

[BM15] Bornmann L., Mutz R.: Growth rates of modern science: A bibliometric analysis based on the number of publications and cited references. *Journal of the Association for Information Science and Technology 66*, 11 (2015), 2215–2222. 2

[BOH11] Bostock M., Ogievetsky V., Heer J.: D3 data-driven documents. *IEEE Transactions on Visualization and Computer Graphics 17*, 12 (Dec. 2011), 2301–2309. 7

[Che13] Chen C.: *Mapping Scientific Frontiers - The Quest for Knowledge Visualization*, revised 2nd edition ed. Elementary Differential Geometry Series. Springer, 2013. 2

[CY11] Chou J.-K., Yang C.-K.: Papervis: Literature review made easy. *Computer Graphics Forum 30* (2011), 721–730. 3

[DBL] Dblp statistics - new records per year. Accessed on February 11th, 2019. URL: https://dblp.uni-trier.de/statistics/newrecordsperyear.html. 2

[FHKM17] Federico P., Heimerl F., Koch S., Miksch S.: A survey on visual approaches for analyzing scientific literature and patents. *IEEE Transactions on Visualization and Computer Graphics 23*, 9 (2017), 2179–2198. 2, 3

[GLK*13] Görg C., Liu Z., Kihm J., Choo J., Park H., Stasko J.: Combining computational analyses and interactive visualization for document exploration and sensemaking in jigsaw. *IEEE transactions on Visualization and Computer Graphics 19*, 10 (2013), 1646–1663. 3

[GPI03] Garfield E., Pudovkin A. I., Istomin V. S.: Why do we need algorithmic historiography? *Journal of the American Society for Information Science and Technology 54* (2003), 400–412. 2

[HSZ13] Hoonlor A., Szymanski B. K., Zaki M. J.: Trends in computer science research. *Communications of the ACM 56*, 10 (2013), 74–83. 2

[IIS*17] Isenberg P., Isenberg T., Sedlmair M., Chen J., Möller T.: Visualization as seen through its research paper keywords. *IEEE Transactions on Visualization and Computer Graphics 23*, 1 (2017), 771–780. 2

[KLSA17] Khan S., Liu X., Shakil K. A., Alam M.: A survey on scholarly data: From big data perspective. *Information Processing and Management 53* (2017), 923–944. 2

[Ley02] Ley M.: The dblp computer science bibliography: Evolution, research issues, perspectives. In *International symposium on string processing and information retrieval* (2002), Springer, pp. 1–10. 2

[MA] Microsoft academic. Accessed on February 11th, 2019. URL: https://academic.microsoft.com/. 2

[OMM13] Osborne F., Motta E., Mulholland P.: Exploring scholarly data with rexplore. In *The Semantic Web – ISWC 2013* (Berlin, Heidelberg, 2013), Alani H., Kagal L., Fokoue A., Groth P., Biemann C., Parreira J. X., Aroyo L., Noy N., Welty C., Janowicz K., (Eds.), Springer Berlin Heidelberg, pp. 460–477. 3

[sco18] Scopus fact sheet, 2018. URL: https://www.elsevier.com/__data/assets/pdf_file/0008/208772/ACAD_R_SC_FS.pdf. 2

[SSS*15] Sinha A., Shen Z., Song Y., Ma H., Eide D., Hsu B.-J. P., Wang K.: An overview of microsoft academic service (ma) and applications. In *Proceedings of the 24th International Conference on World Wide Web (WWW '15 Companion)* (2015), ACM New York, USA. 2

[vEW14] van Eck N. J., Waltman L.: Citenetexplorer: A new software tool for analyzing and visualizing citation networks. *Journal of Informetrics 8* (2014), 802–823. 3

[WoS] Web of science platform: Summary of coverage. Accessed on February 11th, 2019. URL: https://clarivate.libguides.com/webofscienceplatform/coverage. 2

[WSC10] Wang F., Shi N., Chen B.: A comprehensive survey of the reviewer assignment problem. *International Journal of Information Technology & Decision Making 9*, 4 (2010), 645–668. 2