Women in Computer Science Research

What is the Bibliography Data Telling Us?

By Swati Agarwal, Nitish Mittal, Rohan Katyal, Ashish Sureka, and Denzil Correa

The low participation by women authors in research is an important equity issue in Computer Science Research (CSR). There are various parameters and methodologies that can be used to measure the gender imbalance. In this paper, we present a study on gender gap, imbalance and women participation in CSR. We conduct our experiments on DBLP bibliographical database and analyze several years of publication dataset across various domains of CSR. We perform Exploratory Data Analysis on the bibliographical dataset and study the trend of gender imbalance over several years. We propose eight research questions across various facets and our results shows a significant gender imbalance in different sub-fields within CSR and low rate of women participation across various regions of world.

Keywords: Women in Computing, Gender Inequality in Science and Technology, Gender and Society, DBLP Computer Science Bibliography Database, Gender Imbalance, Scientific Publication Mining, Scientometric and Bibliometrics
Categories: Social and Professional Topics – Gender; Human-Centered Computing-Visualization
Corresponding Author: Swati Agarwal
Email: swatia@iiitd.ac.in

Research Motivation and Aim

Studies on gender imbalance and women participation as authors in research papers and scientific publications are conducted to understand gender inequality, reasons for gender differences and impact of policies and practices[1, 2, 3]. Scientometric and Bibliometrics consisting of statistical analysis of scholarly and scientific publications is a widely used methodology for extracting measurable data on academic literature. It has also been used to measure gender imbalance in science. DBLP[4] is a well-known Computer Science bibliography database indexing more than four million publications (as of November 2015) from Journals, Books, Thesis, Conferences, Workshops and Informal Publications. DBLP has been widely used in several bibliometrics based studies as a general dataset for

[1] The evolution of female authorship in computing research. Cavero, José María, Belén Vela, Paloma Cáceres, Carlos Cuesta, and Almudena Sierra-Alonso. Scientometrics 103, no. 1, 85-100 (2015)
[2] Gender and computing conference papers. Cohoon, J. McGrath, Sergey Nigai, and Joseph Jofish Kaye. Communications of the ACM 54, no. 8, 72-80 (2011).
[3] Participation of women in software engineering publications. Vela, Belén, Paloma Cáceres, and José María Cavero. Scientometrics 93, no. 3: 661-679 (2012).
[4] http://DBLP.uni-trier.de/

Computer Science publications[5]. The study presented in this paper is motivated by the need to measure scientific output and contributions of women, gender gap and imbalances in the field of Computer Science by conducting a bibliometric study on DBLP bibliographic database. We define eight research questions across various facets such as the number of articles published by all women authors, contributions of women authors in leadership positions (for example, editors in CSR conferences), percentage of women authors making among top k most prolific authors in CSR and the rate of research productivity of women authors across various sub-fields within CSR and across various regions of world. Our research motivation and aim is to provide answers to those research questions by analyzing several years of scientific publication data in the field of Computer Science. The objective of this study is to measure the number and trend of gender imbalance in CSR and not to investigate the bias or reasons behind this imbalance.

Related Work and Research Contributions

We conduct a literature survey in the area of studying gender imbalance and women participation in various research domains. Cavero et al. analyze the evolution of women in computing research by analyzing Computer Science publications from year 1936 to 2010. They use DBLP dataset for their study and present insights on various aspects of the evolution of female authorship in computing research[1]. Cohoon et al. conduct a study on measuring trends and influences on women authorship of computing conference papers. They analyze data from over 3000 ACM-affiliated conferences, workshops, symposia and forums held between 1966 and 2009[2]. Vela et al. conduct an empirical study for female participation in 12 leading Software Engineering (SE) journals. They conduct an analysis of the gender of the authors, editorial board members, associate editors and editors-in-chief 12 journals over a two-year period (2007 and 2008)[3].

Novel Research Contributions: In context to existing work, the study presented in this paper is the first in-depth and focused work on mining 4.7 million records of DBLP database spanning 16 years on the topic of women contribution and gender imbalance in Computer Science scholarly publications. We conduct a series of experiments to answer several research questions and present empirical results and insights.

Experimental Setup

We conduct experiments on an open source bibliography database. We download September 17, 2015 snapshot of DBLP dataset[6] consists of 4.7 million records approximately including 15.9K books, 1.7 million conference proceeding articles and 1.34 million journal articles and other publications records. We implement an XML parser to convert DBLP database into MySQL containing carefully designed database for faster querying. We split our dataset into four parts named as three different domains of

[5] The DBLP computer science bibliography: Evolution, research issues, perspectives. Ley, Michael. In String Processing and Information Retrieval, pp. 1-10. Springer Berlin Heidelberg, 2002.
[6] http://dblp.uni-trier.de/xml/

Computer Science Research (CSR) and one additional Computer Science & Engineering (CSE) itself as a whole. We choose Software Engineering (SE), Data Engineering (DE) and Theory (TH) as three distinct sub-fields of CSR and select 81 unique conferences (CSE: 29, DE: 20, SE: 11, TH: 30) based on the parameters used in previous literature[7, 8, 9, 10]. We select top conferences in these domains (for example, VLDB, CIKM, KDD from Data Engineering, ICSE, ASE from Software Engineering and STOC, FOCS, CRYPTO from Theory domains) and extract their bibliography data starting from 2000 to 2015. To understand the trend of gender gap and imbalance in CSR, we restrict our analysis to 16 years of data. If two distinct authors have same name in DBLP database, the second name has four padding bits added as a suffix, for example, "Wang Yi" and "Wang Yi 0001". We collect the name of all authors and editors in a list and apply Genderize API[11] (open source) to determine the gender of authors from their first name. As per February 2016 statistics, Genderize API uses a database containing 2,16,286 distinct names across 79 countries and 89 languages. The API classifies names into three classes i.e. Male, Female and NA (if the gender cannot be determined). It returns a confidence score *s* referred as a certainty factor or probability of gender that varies from 0 to 1. If the confidence score for a name is below 60% we classify it as NA. Further, if an author has only initials in name, we classify the gender as NA. For example, "X. J. Chen". For a data sample of 81 conferences, we were able to determine the gender of 3,57,436 authors (2,87,936 as Male and 69,500 as Female) while 59,009 authors' gender cannot be determined using the API. As discussed in Section 0, we analyze bibliography data for studying the rate of women research participation across various regions of the world. Therefore, our next step involves the extraction of affiliation information for all the authors. DBLP database contains "ee" field that redirects to the corresponding publisher site under which the paper is published. DBLP database contains the articles published via 1,182 distinct publishers. We split our database and select only five high rank publishers (Springer, IEEE, IEEE Computer Society, ACM, AAAI) to obtain the affiliation information as each publisher has a different page format and requires an individual parser to crawl the data. We scrape the publisher page of article and map the authors in database with corresponding affiliation. In DBLP database, Springer, IEEE, IEEE CS, ACM and AAAI have 27,362, 14,316, 21,363, 50,193 and 4,855 unique publications respectively. The next step involves obtaining country and coordinates of the affiliation of authors so obtained. If the country name is not available on publisher page, we extract the geocodes of affiliation by using an open source API by Data Science Toolkit[12]. We further reverse geocode these affiliations to obtain the country name. For all the publishers, we reverse geocode 6551 entries and obtain a total of 2,14,865 unique affiliations. We create a complete database of gender and affiliations of authors in DBLP dataset and provide it publicly: Dataset Link.

[7] http://projects.csail.mit.edu/dnd/ranking/
[8] A citation-based system to assist prize awarding. Sidiropoulos, Antonis, and Yannis Manolopoulos. ACM SIGMOD Record 34, no. 4 (2005): 54-60.
[9] How healthy are software engineering conferences?. Vasilescu, Bogdan, Alexander Serebrenik, Tom Mens, Mark GJ van den Brand, and Ekaterina Pek. Science of Computer Programming 89 (2014): 251-272.
[10] Peer-Selected "Best Papers"—Are They Really That "Good"?. Wainer, Jacques, Michael Eckmann, and Anderson Rocha. PloS one 10, no. 3 (2015): e0118446.
[11] https://genderize.io
[12] http://www.datasciencetoolkit.org/

Experimental Results

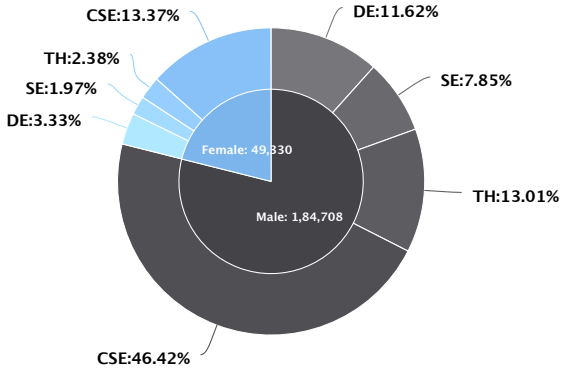

Figure 1 Variation in Number of Male and Female Researchers in Three Different Domains of CSR

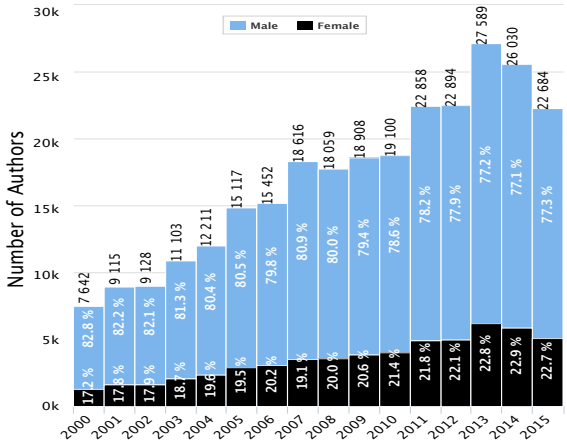

Figure 2 Trend of Gender Imbalance in Computer Science Research over a Period of Time from 2000 to 2015

In this Section, we analyze the DBLP bibliography dataset and study the gender imbalance in CSR by defining various facets. We propose 8 research questions (RQs) and perform a statistical analysis on the dataset from different perspectives. In following subsections, we define these RQs and present our experimental results.

RQ1: Is Computer Science Research a male dominated field?

Figure 1 shows the distribution of male and female authors in CSE, DE, TH and SE domains. Figure 1 reveals that in Computer Science Research, only 21% of the researchers (actively publishing in CSR conferences) are female authors and rest 79% are male authors. Software Engineering has the least number of authors publishing papers in in SE conferences (7.85% of male authors and 1.97% of female authors). Figure 1 also reveals that both female and male authors are publishing actively in both Data Engineering and Theory (difference of approximate 1%). However, women participation is relatively lower than men authors. For example- Data Engineering has a participation gap of 8.29% and similarly, in TH domain women participation is 10.63% lower than men participation.

RQ2: Is there an upward or downward trend on gender imbalance? Figure 2 shows a stacked column chart showing the relationship of number of female and male authors to the whole (total number of authors). The stacked column chart in Figure 2 consists of two categories: number of female authors and number of male authors. The contribution of each of the category to a total across categories is shown for a period of 16 years. For example, for the year 2010, the total number of authors is 19100 out of which 21% are female and 79% are male. Figure 2 reveals an increase in the percentage of female authors from 17% in 2000 to 23% in 2013. The percentage of female authors from 2000 to 2013 is not monotonically increasing (as there is a slight decrease from 2004 to 2005 as well as 2006 to 2007) but in general has an upward trend. Experimental results show that there is an increase in the percentage of female authors over a period of 13 years but

the increase is marginal. Results for year 2013 shows a significant gender gap in authorship as the women representation is still less than 25%. It is interesting to note that for the last 3 years (2013, 2014 and 2015), the percentage of female authors is 23%. The rate of women as a percentage of the rate for men for past 3 years is 23/77=29.87%. A value of 100 percent indicates perfect gender balance and a value less than 100 indicates that women are underrepresented. The further away the value is from 100%, the more imbalances.

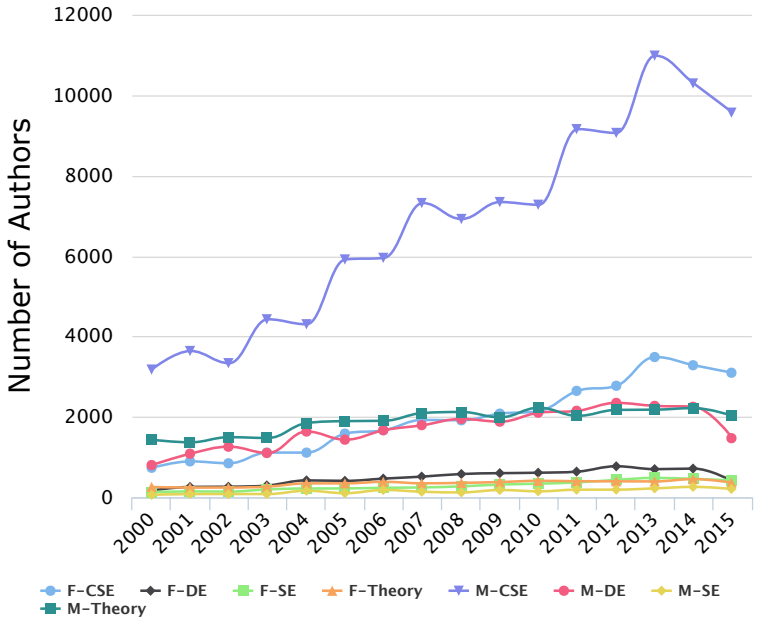

Figure 3 Domain Specific Trend of Gender Imbalance Over a Period of 16 Years (2000 to 2015)

RQ3: Is the gender imbalance problem unique to some fields or is it a common concern? We perform a trend analysis on gender imbalance and as discussed in Section **3**, we analyze the data extracted for a period of 16 years. Figure 3 shows a timeline-based graph showing the trend of number of male and female authors in different fields with-in Computer Science Research. Figure 3 reveals that there is an upward trend for both male and female authors in all four domains. However, initially for first 10 years the rate of the increment in female authors is relatively low in comparison to male authors. For example, in 2000, number of male authors is increasing by 2.83% while female authors are increasing only by 0.86%. Similarly, from 2005 to 2007 male authors are increasing by 2.36, 1.25, 1.54 percentage while increment in female authors is always less than 1%. It is interesting to see that the pattern of increment rate is same for female researchers in all three domains (DE, SE, and TH) and the values vary with a difference between 0.5 and 0.2 percentage. Figure 3 reveals that in 2015, there is a large decrement in number of authors across all sub-fields of CSR. We assume that it might be possible since we extract September 2015 snapshot of DBLP database and proceedings of some of conferences happened in 2015 year are missing or not yet available in database.

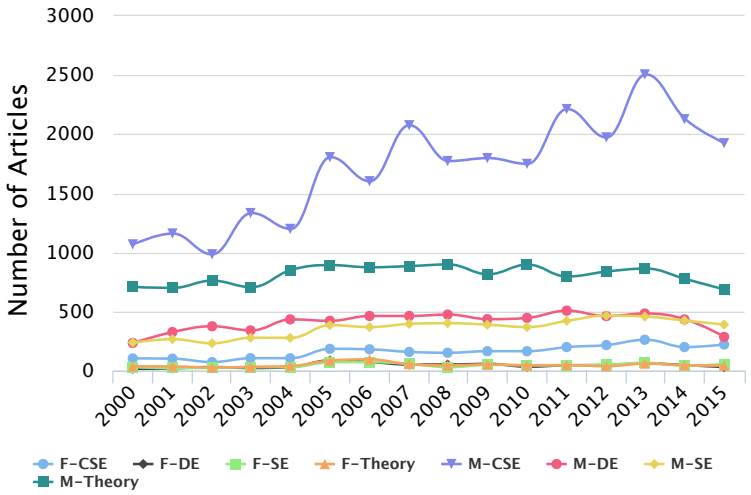

Figure 4 Trend of Domain Specific Articles Co-Authored by only Men and only Women Researchers

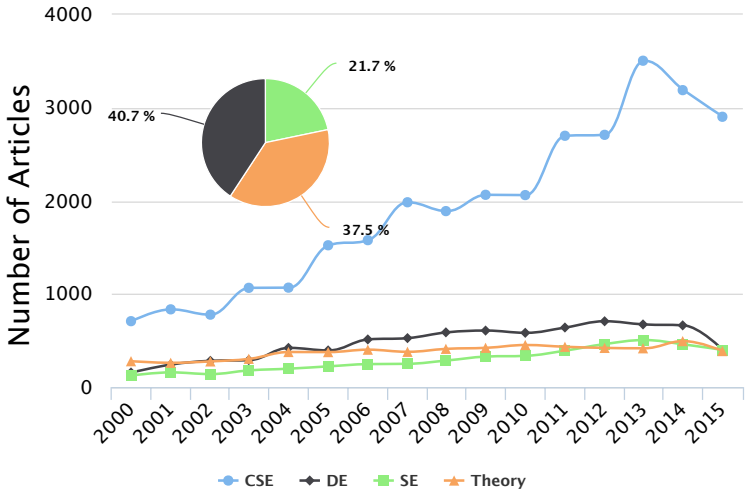

Figure 5 A Timeline Based Review of Number of Articles Co-authored by At-least One Female Researcher

RQ4: What are the different trends of women authored papers in CSR? Figure 4 shows a line chart showing the number of articles co-authored by only men and only female authors. The graph in Figure 4 shows a timeline-based review of number of articles varying over a period of 16 years across various research domains within Computer Science. Figure 4 reveals that given the ratio of male and female authors in CSE, the number of articles co-authored by only men authors is also relatively higher (27,291) in comparison to articles co-authored by only female authors (2583). As illustrated in Figure 2 and Figure 3, number of female authors publishing in CSE conference is relatively lesser than male authors. However, in Figure 4, we see that articles published by only female authors are even lesser. Lines for articles published by only female authors are overlapping for DE, SE and TH domains showing that they all have same patterns fluctuating only with a difference of 0.05%. In absolute Figures, we found that for all three domains, number of articles co-authored by only female authors

are mostly below 50; for example, from 2000 to 2004 and from 2007 to 2015. We see a rapid increment in number of articles in 2005 and 2006 (0.56, 0.46, 0.42 for DE, TH and SE respectively). Figure 4 reveals that the trend of number of all men authored articles published in Theory domain conferences are higher in comparison to all other domains in over the years. For example, the maximum number of publications by female authors in CSE (261), DE (86), TH (95) and SE (72) are at least 62% lower than the minimum number of publications by men authors in Theory conferences (685 in 2015 that further doesn't include all the proceedings of year).

In Figure 4, we observer that only-women authored articles are lesser in comparison to only-men authored articles. However, the number of articles co-authored by both men and women researchers are relatively higher. Figure 5 shows the relationship and trend of number of articles published over past 16 years that are co-authored by at least one female researcher. The graph in Figure 5 shows the combination of pie chart and timeline graph showing a trend of papers published in all selected domains of CSR. As shown in the Figure 4, in year 2000, number of articles published by only-women authors is below 200 while the number of articles also co-authored by other male authors is 700 approx. (refer to Figure 5). In Figure 5, we see an upward trend in all domains with relatively high margin (varying between 0.2% and 0.5%) unlike Figure 4. Initially, the line graph of articles published in Theory and DE domains are overlapping (difference <0.50%) while after 2005, this gap increases with a difference between 1.50 and 1.80 percentage. Further, later in 2012, it reaches at 2.90%. As we compare Figure 3, Figure 4 and Figure 5, we notice that the number of authors publishing in Software Engineering conferences is relatively less in comparison to Data Engineering and Theory. The trend of number of articles co-authored by only-men and only-women researchers are always below 500. We find a similar trend in number of articles that are co-authored by both male and female researchers (also evident in the pie chart- SE domain shares only 21.7% of articles). Figure 5 shows that line for Software Engineering has the minimum publications over past 16 years except in 2012 (0.41% higher than TH) and 2013 (0.88% higher than TH). Further, we add one more dimension in our analysis i.e. "primary author" and review the

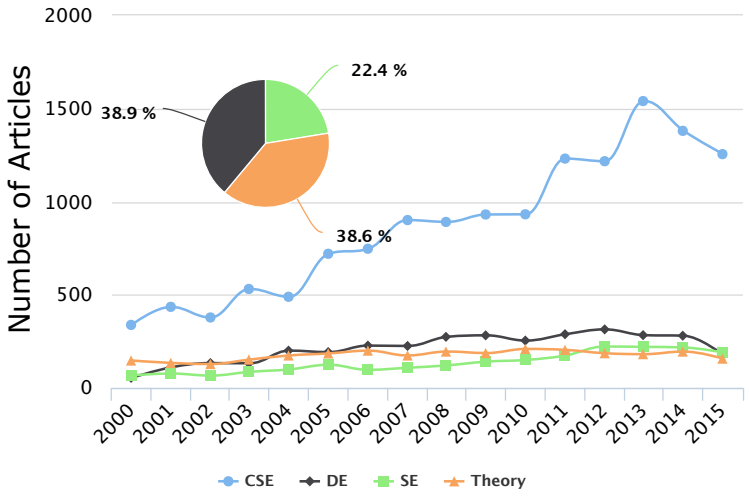

Figure 6 Domain Specific Trend Analysis of Articles Primarily Authored by a Female Researcher

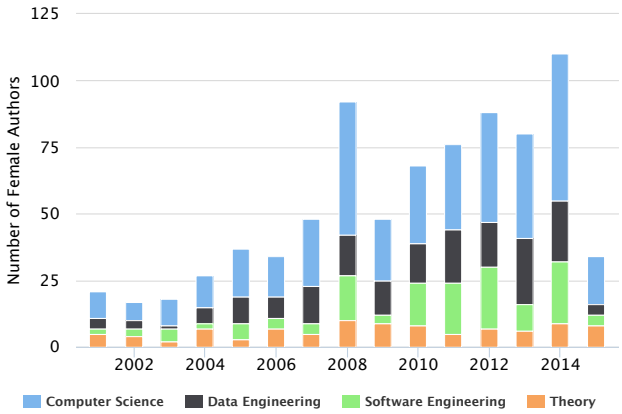

Figure 7 A Timeline Based Review of Female Authors Holding a Leadership Position in CSR Conferences

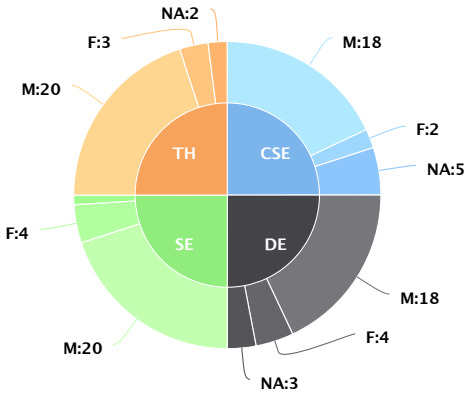

Figure 8 A Donut Presentation of Domain Specific Distribution of Top 25 Authors Publishing in CSR Conferences

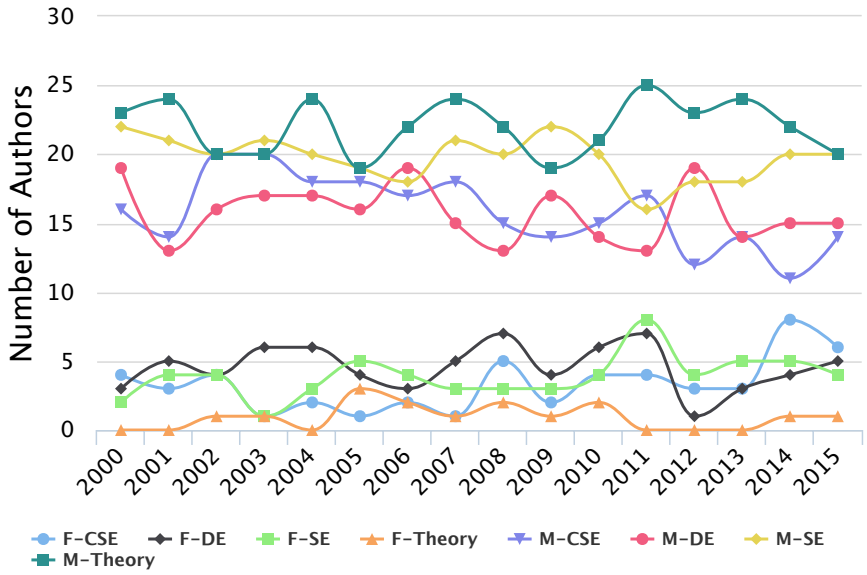

Figure 9 A Timeline Based and Domain Specific Review of Top 25 Authors Publishing in CSR Conferences

number of articles published over past 16 years that are primarily authored by female researcher. Figure 6 shows the relationship of number of articles published in CSR domain conferences and their trend from 2000 to 2015. This analysis is an extension of Figure 5; therefore, despite the variation in absolute Figures (huge decrement by 54.5%) we see a similar pattern of trend. In Figure 6, we observe a decrement of 1.8% in total number of articles in Data Engineering domain while there is an increment of 0.7% and 1.1% in Software Engineering and Theory domains respectively.

RQ5: What percentage of women researchers is in conference leadership positions? We analyze the DBLP bibliography dataset from another perspective and check the number of female researchers that hold a leadership position in various conferences of CSR. Figure 7 A Timeline Based Review of Female Authors Holding a

Leadership Position in CSR Conferences shows the trend of such numbers over past 16 years in four different field of Computer Science. DBLP snapshot provides the information of researchers/authors who hold editor position in various conferences of CSR. The numbers reveal that maximum ratio of these editors are male researchers. For example, in Data Engineering conferences, numbers of male editors are 71.1% more than female editors and similarly in Software Engineering and Theory domains, the difference is 59% and 86.5% respectively. These numbers are a strong evident of gender imbalance on leadership positions in CSR. In Figure 7, CSE has the maximum number of female authors and is an outlier. Number of female authors in CSE has an upward and smooth trend except a sudden peak in 2008 (from 4 to 17) and immediate fall in 2009 (from 17 to 3). However, the numbers increase gradually after 2009 (between 15 and 17%). As we compare the number of female editors in various domains, Figure 7 reveals that despite the variation in numbers, DE and TH domains have same patterns till 2010 except two major gaps in 2005 (70%) and 2007 (64.3%). After 2010, both domains follow an invert pattern and values are reciprocal to each other. For example, after 2010, number of female editors in DE increases by a percentage of 25 while in TH the number falls down by 37.5%. Similarly, in 2012 where the number increases by 28.5% in TH domain, it decreases in DE by 15%. Further, despite the minimum publications by female authors in Software Engineering domain, we observe a relatively large percentage of female editors in SE in comparison to Theory domain. However, these numbers have been fluctuating in every year including sudden peaks and fall. For example, after 2007, the number of female editors increases by 76.5% and immediately falls by 82.35% in 2009.

RQ6: How many women are among Top 25 most prolific authors for a particular field? Previous RQs reveal that there is a gender imbalance in CSR in terms of number of publications by female authors and the leadership positions. In RQ6, we analyze the women participation among most prolific authors publishing in CSR conferences. We define most prolific profile based on the number of publications by an author. Figure 8 shows a donut presentation of number of Male, Female and NA authors who are among top 25 authors in CSR. We extract these numbers for all selected conferences of CSR (81). Figure 8 reveals the numbers of females (the ones who are determined) in top 25 authors are always below 5, which is a strong evidence of gender imbalance. While comparing with other domains, despite the maximum numbers of publications and editor positions, the numbers of most prolific females are minimum in CSE conferences. We perform a domain specific trend analysis of top 25 profiles publishing papers in CSR conferences every year and observe a significant variation in numbers. Figure 9 shows a timeline-based review of number of top 25 authors (both male and female) in CSE, DE, SE, and TH domains. Figure 9 shows that the numbers of top female researchers are always lower than number of male researchers. Following a different pattern from previous graphs, Figure 9 reveals that despite several sudden decrements over 16 years, Theory domain has the maximum number of male authors making among top 25 researchers of year while, numbers of female researchers are always below 5 including six zeros. For example, in 2000 number of male, female and NA authors are 23, 0 and 2 respectively. Similarly, in 2011, there are only male authors in top 25 prolific researchers in Theory domain. However, unlike TH, we see a substantial increment in number of female researchers among top 25 authors publishing in DE conferences (from 10 to 20%).

We observe that the number of male and female authors follow a pattern in DE domain for a while. For example, from 2005 to 2013 the lines are inversely related to each other. However, as we compute the median of these numbers, we find that for both male and female authors in DE, only 31.25% of the numbers are above median value while 50% of the numbers are below the median edge. Similarly, in Software Engineering domain, 37.5% (occurs 6 times) of the times number of female authors are below the median value while only 18.5% (occurs 3 times) times the male authors are below their median value. Figure 9 reveals that even after taking all the numbers collectively, maximum number of female authors (8) never made it up to the minimum value of male authors (11), which is below 50%.

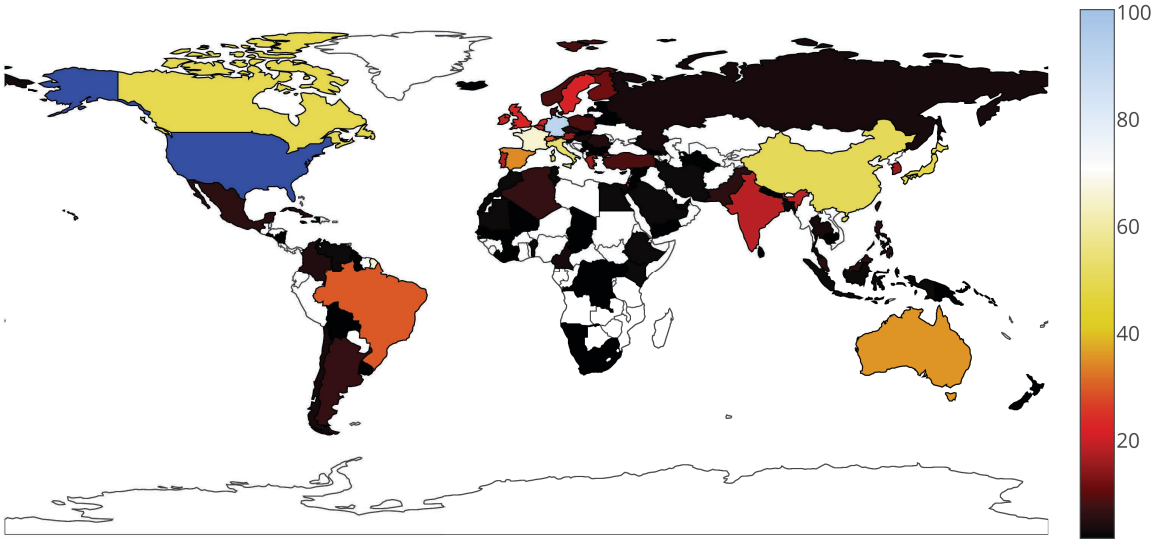

Figure 10 Unique Locations of All Female Authors Publishing Papers in Top Five Publishers of CSR

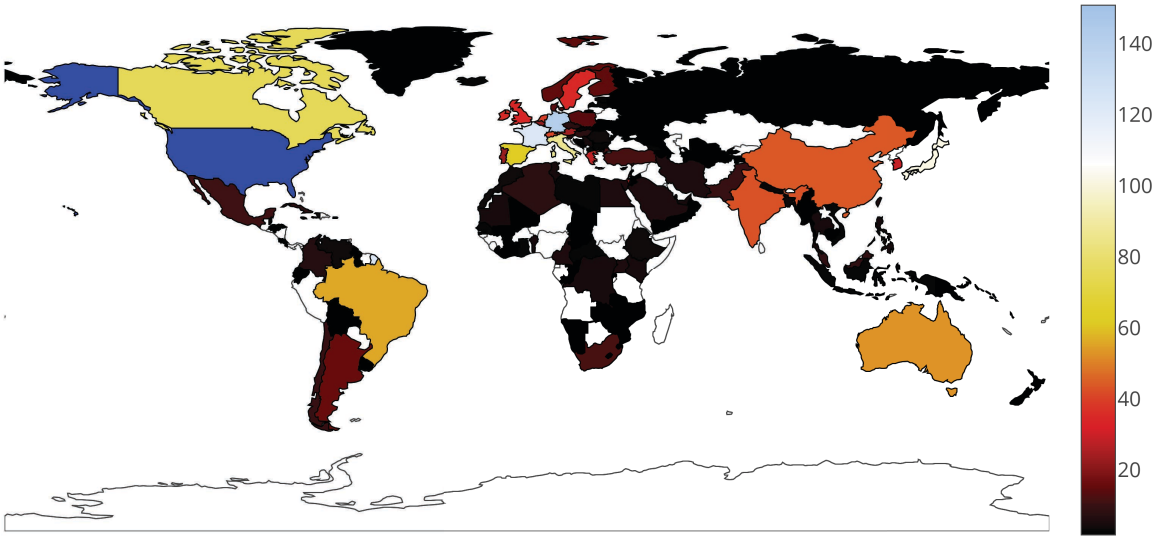

Figure 11 Unique Locations of All Men Authors Publishing Papers in Top Five Publishers of CSR

RQ7: Is the gender imbalance specific to only developing countries? In this research question, we study the collective participation of women researchers in Computer Science Research and analyze their rate of research productivity in various developing and developed regions across the world. For conducting this experiment, we take a sample of DBLP dataset from 2010 to 2015 and extract the affiliations' locations of men and women authors publishing papers in top rank conferences of Springer, IEEE and IEEE Computer Society, ACM and AAAI. We analyze the rate of women participation in all 81 conferences that fall under the publishers mentioned above. Figure 10 shows the choropleth map showing 2,131 women authors publishing from 109 countries. Similarly, Figure 11 shows distribution of 3,494 men authors publishing from 137 countries across the world. As shown in the Figures, USA has the maximum number of male and female authors (1920 and 1162 respectively), which is approximately 92% higher than the country from top 2. Figure 10 and Figure 11 reveal that France and Germany have the second highest number of male and female authors. While Canada, Australia and Spain also have a large number of authors publishing in top rank CSR publishers. However, despite less number of authors in comparison to other advanced developed countries, China has maximum rate of women participation in top rank CSR publishers i.e. 53% while USA has only 37% women participation. Similarly, Japan, France and South Korea have women research participation below 35%. Figure 10 reveals that despite being a less developed country, Brazil has higher number of female authors (29) having articles in top rank publishers than South Korea (15) and various other advanced developed countries of Europe continent. Similarly, South Korea (an advanced developed country) has lesser number of male and female authors than India (17% and 33% respectively). From Figure 10 and Figure 11, we also find that other developed countries like Japan (32%), France (34%), Canada (39%) and South Korea (34%) have approximately similar rate of women participation in top rank CSR publishers as less developed countries like India (33%).

RQ8: Is the gender imbalance in top tier CSR conferences specific to developing countries? Figures 12 and 13 show the world thematic maps showing the relationship between authors publishing papers in high rank CSR conferences (WWW, SIGIR, VLDB, SIGCOMM, CHI, and ICSE) and their affiliation from various regions of world. Figures 12 and 13 demonstrate a choropleth map for the number of unique women and men authors respectively. DBLP dataset reveals that there are 1828 women and 3253 men authors publishing articles in top CSR conferences from a total of 81 and 93 countries respectively. Similar to Figure 10, USA has the maximum number of male and female authors, which is an outlier (728 and 1,328 respectively). Therefore, we define second highest value (Germany- 219 and 112 respectively) as maximum range in graduate colors (shown as legend in heatmaps). The bibliography data reveals that despite being an advanced developed country there is only 35% (728 out of 2056 authors) women research participation in USA. While despite being less advanced country, China has 48.4% (61 out of 126 authors) women participation, which is nearly half of the researchers' population publishing in top rank conferences. Similarly, Canada and South Korea have 42.3% women participation in Top 6 CSR conferences, which is reasonably higher than USA (a gap of 7%).

Figures 12 and 13 also reveal that among all advanced developed countries, Japan, Germany and USA have minimum rate of women

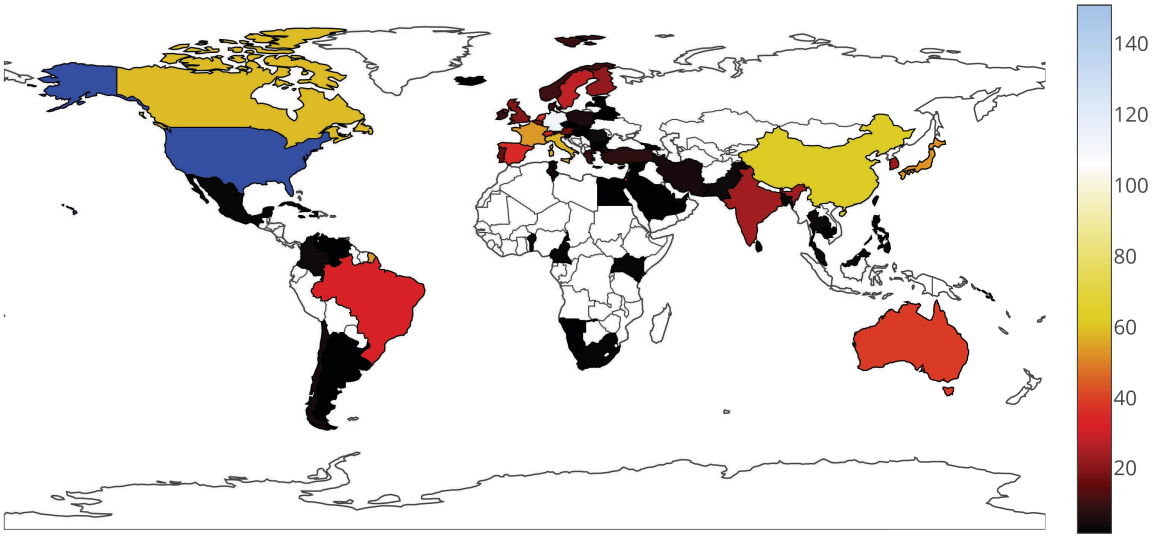

Figure 12 Unique Locations of All Women Authors Publishing Papers in Top 6 Conferences of Computer Science Research (WWW, CHI, ICSE, SIGIR, SIGCOMM, VLDB) in a Span of Six Years (2010-2015

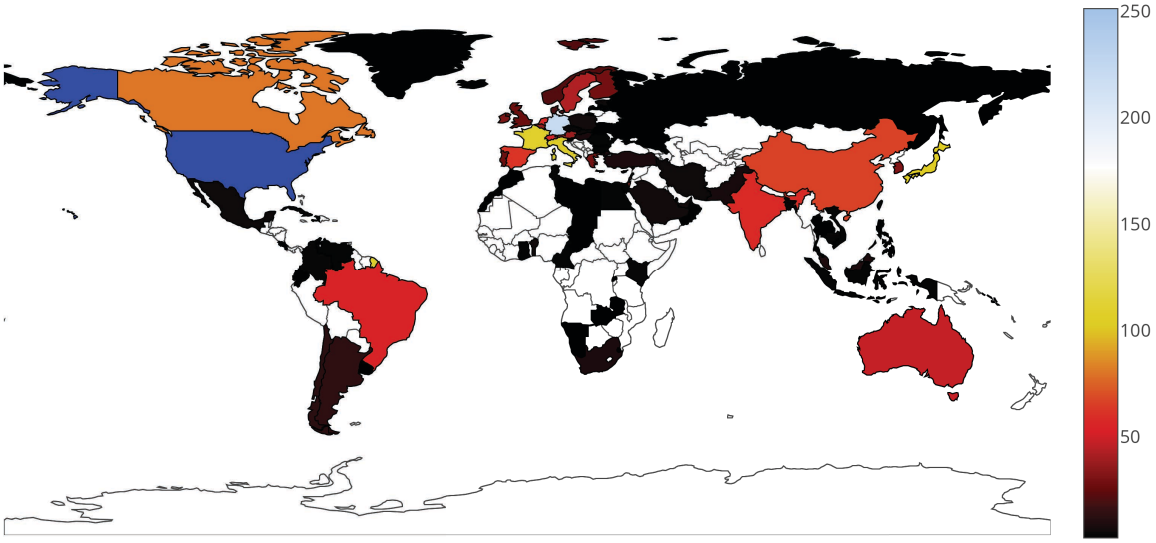

Figure 13 Unique Locations of All Men Authors Publishing Papers in Top 6 Conferences of Computer Science Research (WWW, CHI, ICSE, SIGIR, SIGCOMM, VLDB) in a Span of Six Years (2010-2015)

participation (between 33 and 35%). Further, despite having only 4% of total number of articles published by authors from USA, Australia has 45.2% of women authors making among top tier conferences of CSR. In comparison to other developing countries, some of the highly advanced developed regions have a significantly low rate of research productivity from women authors. For example, Brazil has higher number of unique women contributing in top rank CSR conferences in comparison to Japan and Germany (37%- 31 out of 84 authors). Similarly, China has 61 unique women authors publishing papers in top rank CSR conferences while only 22 women authors from South Korea have papers in top CSR conferences in last 6 years. India being a less developed country,

it has 29% women research participation while France- an advanced developed country has 0% women authors publishing in prestigious conferences of CSR. **Error! Reference source not found.** and **Error! Reference source not found.** also reveal that there are various countries of South Africa continent and other regions across the world where women participation in highly reputed conferences of CSR is either zero or insignificant. For example, Sudan, Libya, United Arab Emirates, Russia, and Indonesia.

Conclusions

In this paper, we perform an exploratory data analysis on DBLP database for studying gender gap and imbalance in various sub-fields of CSR (CSE, DE, SE and TH). We define eight research questions across various parameters and our experimental results reveal that CSR is a male-dominated field including all selected sub-domains. Based on DBLP statistics, we conclude that the research participation and collaboration of women authors is reasonably lower than the male authors. The trend of this participation is increasing over the past decade. However, the rate of increment is below 0.1%. Further, in terms of making substantial contributions by holding a leadership position (editor) in CSR conferences, women participation is significantly low including least participation in Theory domain. We also conclude that the gender imbalance is not specific to only developing countries. Our analysis reveals that despite having the maximum number of publications, USA has low women participation than China and Brazil that are less developed countries. DBLP publication dataset demonstrates that Computer Science Research and various other domains of CSR are male dominated areas and there is a significant variation in women participation across various regions of the world.