

A Review of Ten Years of the Symposium on Search-Based Software Engineering

Thelma Elita Colanzi¹, Wesley Klewerton Guez Assunção²,
Paulo Roberto Farah^{3,4}, Silvia Regina Vergilio⁴, and Giovani Guizzo⁵

¹ DIN - State University of Maringa, Maringa, Brazil thelma@din.uem.br

² Federal University of Technology - Paraná, Toledo, Brazil wesleyk@utfpr.edu.br

³ Santa Catarina State University, Ibirama, Brazil paulo.farah@udesc.br

⁴ DInf - Federal University of Parana, Curitiba, Brazil silvia@inf.ufpr.br

⁵ CREST centre - University College London, UK giovaniguizzo@gmail.com

Abstract. The year 2018 marked the tenth anniversary of the Symposium on Search Based Software Engineering (SSBSE). In order to better understand the characteristics and evolution of papers published in SSBSE, this work reports results from a mapping study targeting the ten proceedings of SSBSE. Our goal is to identify and to analyze authorship collaborations, the impact and relevance of SSBSE in terms of citations, the software engineering areas commonly studied as well as the new problems recently solved, the computational intelligence techniques preferred by authors and the rigour of experiments conducted in the papers. Besides this analysis, we list some recommendations to new authors who envisage to publish their work in SSBSE. Despite of existing mapping studies on SBSE, our contribution in this work is to provide information to researchers and practitioners willing to enter the SBSE field, being a source of information to strengthen the symposium, guide new studies, and motivate new collaboration among research groups.

Keywords: Systematic mapping · SBSE · Bibliometric analysis.

1 Introduction ⁶

The year 2018 marked the tenth anniversary of the Symposium on Search Based Software Engineering (SSBSE), the premier event on Search Based Software Engineering (SBSE). SBSE is the research field that formulates Software Engineering (SE) problems as search problems. In this way, heuristic techniques are used to reach optimal solutions to efficiently solve a large variety of problems associated to different SE tasks. Over the past ten years, the symposium has drawn attention of researchers, academics and practitioners alike, contributing to strengthen the field and to integrate the SBSE community, gathering a large body of studies that serve as reference for researchers. To obtain and understand

⁶ This work was funded by CNPq (Grants 428994/2018-0 and 408356/2018-9) and by the ERC Advanced Grant 2016, ID 741278, Evolving Program Improvement PE6 London Collaborators (EPIC).

a big picture of SSBSE, we synthesized this ten-year history of research through a systematic mapping [9] conducted over all the SSBSE proceedings.

In the literature we can find surveys in the SBSE field [5–7] reporting applications of search-based algorithms on software bug fixing, project management, planning and cost estimation, software comprehension, refactoring, software slicing, service-oriented software engineering, compiler optimization, quality assessment, etc. Such surveys analyze the most used search-based algorithms and also point out research directions on SBSE. de Freitas et al. [3] present a bibliometric analysis of the SBSE field. Such works show a growing number of SBSE papers, and an increasing number of addressed SE activities.

Our work also analyzes the addressed SE tasks and used Computational Intelligence (CI) techniques, similarly to aforementioned SBSE studies. But, differently from related work, our focus is the SSBSE. In this way, we provide different analysis regarding the composition of steering and program committees, submission tracks, paper acceptance rate, and impact of the papers published. Such analysis allows a deeper view of SSBSE and contributes to comprehend how the symposium has been evolving along the years.

In our mapping⁷, we adopted the guidelines of Petersen et al. [9] and the following Research Questions (RQ):

RQ1: What are the basic SSBSE numbers? To answer this RQ we provide a quantitative analysis of the event: number of submitted and published papers along the years, acceptance rate, authors and committees characteristics, research groups and collaborations.

RQ2: What is the external impact of SSBSE? To answer this RQ we provide a citation analysis of the SSBSE papers, in order to evaluate the visibility and importance of publishing in the symposium.

RQ3: What are the most common addressed SE areas and CI techniques? To answer this question we provide a quantitative analysis of the addressed SE areas and number of papers in each, as well as the employed CI techniques. Besides, we analyze possible changes and trends over time.

RQ4: How have the SBSE approaches been evaluated? To answer this question we provide an analysis of the experimental evaluation carried out in the papers, identifying applied statistical tests and subjects. The main idea is to analyze experimental rigour employed in the studies published in the symposium and if such a rigour has changed over time.

In this way, the main contributions of this work are: i) to ascertain the impact and relevance of SSBSE, by reporting its main numbers and performing a citation analysis of the published works; ii) to devise a co-authorship network and depict the most prolific research groups and researchers, as well as the participation of the industry; iii) to point out the software engineering areas that have been most subjected to investigation as well as the ones that need more attention; iv) to identify the main CI techniques; and v) to analyze how SBSE approaches have been evaluated.

⁷ Raw data at <https://wesleyklewerton.github.io/SSBSE2019-DataCollection.ods>

Studies like ours are important to corroborate the importance of the symposium and if it has been following up the main changes pointed out by the existing surveys and mappings of the SBSE field, as well as to evaluate its representativeness. Besides, we list some recommendations to new authors who envisage to publish their work in SSBSE, providing information to researchers and practitioners willing to enter the SBSE field, being a source of information to strengthen the symposium, guide new studies, and motivate new collaboration among research groups. In Sections 2 to 5 we answered each posed RQ. Section 6 presents some recommendations to new authors. Section 7 concludes the paper.

2 RQ1 – SSBSE in numbers

The first edition of SSBSE occurred in 2009, in Windsor, United Kingdom (UK), and since then the symposium took place in six different countries in Europe, South America (2014), and North America (2016). Some editions were co-located with ESE/FSE, and with other events such as ICSME and ASE.

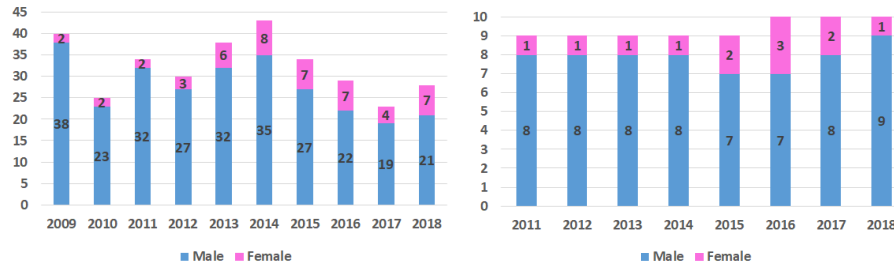
The symposium attracts researchers, students, lecturers and members from industry. Each SSBSE edition had the honor of having at least two keynotes, one from SE and other from the optimization field, in a total of 25 keynotes, as well as 11 tutorials and 4 panels.

Committee characteristics. Regarding the committee composition, the number of committee members varies from a minimum of 23 in 2017 to a maximum of 43 in 2014 (see Figure 1.(a)). Such members are from different countries, but we do not observe a great variation in the number of represented countries along the editions (minimum number of countries is 9, in 2009 and maximum number is 14, in 2014), average of 11.5.

A greater variation and significant gender imbalance are observed when we consider the percentage of women in the committee⁸. This percentage varies from a minimum of 5% (2009) to a maximum of 25% (2017). The gender imbalance has been decreasing in the last years. Considering the steering committee (Figure 1.(b)) such imbalance has also been decreasing. Such committee was composed for the first time in 2011. In the first four editions it had only 1 woman in a total of 9 members (percentage of 11%). The maximum percentage of women is 30% (2016). The number of countries represented in such committee has been kept almost constant (around 5, with a maximum number of 8 in 2012).

In spite of this gender imbalance, the percentage of women researchers in leadership positions in SSBSE is greater regarding other conferences and the Computer Science area [1]. We had a total of 43 chairs, 14 of which are women (33%). If we consider only the main track, this percentage is similar (35%), 7 women out of 20 chairs. This imbalance has been decreasing in the last five years. Considering the main track, we observe a perfect balance since 2014; a woman and a man have been chosen for chairs since then.

⁸ We manually checked the gender of committee members and authors by doing a web search in their profiles by using Google Scholar, Microsoft Academic, Research Gate, Linkedin, etc. We didn't find name and affiliation of two authors only; we used Genderize.io API and both were defined as females.



(a) Program Committee

(b) Steering Committee

Fig. 1: Committee Characteristics - Gender Imbalance

Table 1: SSBSE in Numbers. (COU: number of different countries submitting papers. TSUB: number of submissions including all tracks. SUB: number of submissions. ACC: number of accepted papers. Rate: percentage of acceptance. “-” means unknown or 0.)

Year	COU	TSUB	Full			Short/F.Abstract			Student			Challenge		
			SUB	ACC	Rate	SUB	ACC	Rate	SUB	ACC	Rate	SUB	ACC	Rate
2009	14	26	-	9	-	5	-	3	-	-	-	-	-	
2010	-	36	-	14	-	-	-	3	-	-	-	-	-	
2011	21	43	37	15	40.5	-	8	-	6	3	50	-	-	
2012	20	38	34	15	44.1	-	3	-	4	2	50	-	-	
2013	24	50	39	14	35.9	-	6	-	9	6	66.6	4	2	
2014	19	51	32	14	43.7	3	1	33.3	8	3	37.1	8	4	
2015	15	51	26	12	46.1	8	4	50	4	2	50	13	13	
2016	20	48	25	13	52	9	4	44.4	7	4	57.1	7	7	
2017	14	32	26	7	26.9	2	5	-	2	2	100	4	4	
2018	10	13	12	12	100	8*	6	75	-	-	-	1	1	

* with Hot of the Press Track

Number of submissions and acceptance rate. SSBSE has provided different tracks in its ten editions. Some statistics about such tracks are presented in Table 1. The main track of full research papers and the student track occurred in all editions with independent chairs. We can see that the total number of submitted papers considering all tracks is greater in the period of 2013-2016. A similar fact can be observed considering the number of submitted papers for the full research papers and the student track.

Regarding the main track of full research papers, the number of accepted papers varies from 7 (in 2016) to 15 (in 2011 and 2012). The acceptance rate of the main track falls in the range of 27% (2017) to 100% in (2018). These last two years are outliers. In 2018, a shepherding phase was added in the reviewing process, which may justify 100% of acceptance. In fact, we do not observed great variations in the acceptance rate before 2016, considering 2011-2016 the mean rate is 43%. After a period of growing and boom, we observed a decrease in the number of submitted papers, what might be justified by the recent inclusion of SBSE in the list of topics of several conferences.

The characteristics of the short papers track varied along the editions. In most editions, separated calls for short papers or fast abstracts were provided, with or without independent chairs. In some editions accepted short papers

Table 2: Most prolific authors

Name	Country	P	C
Andrea Arcuri	Norway, Luxembourg	8	250
Gordon Fraser	Germany, UK	7	208
Paolo Tonella	Italy, Switzerland	7	120
Shin Yoo	UK, Korea	6	176
Mark Harman	UK	6	157
Marouane Kessentini	USA	6	119
Jeffeson T. de Souza	Brazil	5	136
Giuliano Antoniol	Canada	5	67
Yann-Gaël Guéhéneuc	Canada	5	64
Enrique Alba	Spain	4	84
Ruilian Zhao	China	4	42
Silvia R. Vergilio	Brazil	4	34
Thelma E. Colanzi	Brazil	4	34
Betty H.C. Cheng	USA	4	30
Annibale Panichella	Netherlands, Luxembourg	4	14

Table 3: Author's churn

Year	New	Rep.	Left	Total	Churn
2009	24	0	0	24	0.0
2010	32	6	18	38	133.3
2011	34	8	30	42	89.4
2012	34	16	26	50	80.9
2013	34	8	42	42	68.0
2014	39	5	37	44	92.8
2015	35	4	40	39	79.5
2016	39	6	33	45	100.0
2017	13	3	42	16	28.9
2018	31	4	12	35	193.7

were originally submitted as full papers. The challenge track started only in 2013. Thus it is not possible to analyze the acceptance rate over the ten years of both tracks. The edition of 2017 included a journal-first papers track with 2 papers, and the last edition, in 2018, a Hot of the Press track that also included short/student papers with 6 papers. Including all tracks, we had a corpus of 220 papers, 125 associated to the full track, published by IEEE in the first two editions, and by Springer since the third one. Because of this variety in the tracks, the analysis conducted to answer our research questions includes only the 125 papers of the full research track collected from all the SSBSE proceedings.

Authorship. In 125 full papers, we found 271 distinct authors. Then we analyzed their affiliations, and identified the most prolific authors and collaborations. Table 2 presents a ranking of authors that have published at least four papers, ordered by number of papers and number of citations. The third and the fourth columns present the number of published papers and the total number of citations received of all published papers, respectively. Andrea Arcuri is the author with the greatest number of publications and citations. Another aspect to highlight is that greater productivity does not mean higher citations as some authors have fewer SSBSE papers and more citations (see also Section 3).

We can see that the great majority of authors (256 out 271 (94.5%)) published less than 4 papers. Table 3 quantifies unique authors who are new or returned to publish at SSBSE with at least one year without publication (column New), authors that published in the year before and maintained their position publishing at least one paper in the event (column Rep) and authors who did not publish anymore and left the event (Left). Additionally, we calculated the yearly churn rate, presented in the last column. The results indicate that few authors keep publishing along the years, 2012 was the edition with highest number of unique authors and 2017 was the year with the lowest. Churn rate is very high, the highest value was obtained for 2018 and the lowest for 2017.

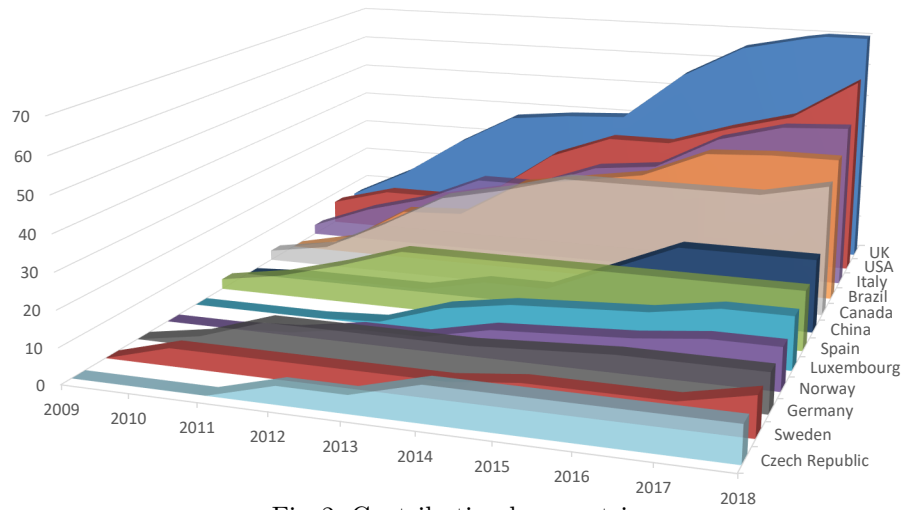


Fig. 2: Contribution by countries

We also investigated the number of countries represented by the authors. To this end we identified the country of all affiliations presented in the papers. Thus, if an author was affiliated to two countries, both were counted in our analysis. The analysis revealed that 24 different countries contributed to SSBSE. The top 3 countries (12.5%) had a contribution of 44.6% and the top 5 countries (20.8%) had a contribution of 64.7%. Figure 2 shows cumulative number of contributions of 12 country affiliations (50%). We can observe that authors from the UK contributed considerably more than other countries over the ten years. Next, there are four countries that have been disputing the second position in the period: USA, Italy, Brazil and Canada. The USA have been maintaining the second place since 2013, tied with other countries in some years. Following, there is a third block, composed by: China, Spain, Luxembourg, Norway, the Netherlands, Germany, Sweden and Czech Republic. There are some interesting aspects about these countries. First, the number of authors from China did a big jump in years 2015 and 2016, which made them lead the number of contributions of this third group. We can also highlight the fact that Spanish authors participated actively only in the first four years of the event and, since 2014, nobody from Spain has published any other paper. Authors from Germany presented a similar behaviour, 81.8% of the contributions were published in 2010 and 2011. Another important aspect of this group is that we observed an increase in the number of papers from Norway and Luxembourg, maybe due to collaborations with other countries. Norway collaborated with other countries in 41.6% of published papers and Luxembourg in 30%.

Another analysis shown in Figure 3 presents the percentage of women that published papers. It varies from 8.3% in 2009 to 29% in 2013. The results show a big gender imbalance that has not been decreasing along the years.

Some authors belong to more than one kind of institution. Figure 4 displays the percentage of authors from universities, research foundations and indus-

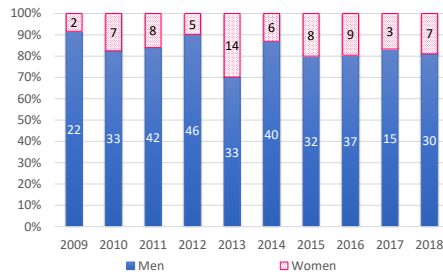


Fig. 3: Authors Gender Imbalance

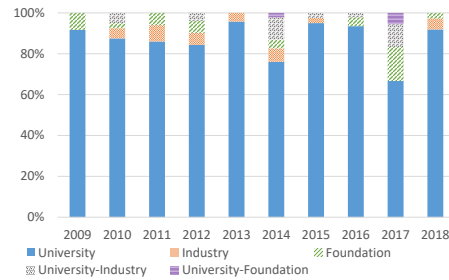


Fig. 4: Source of Contributions

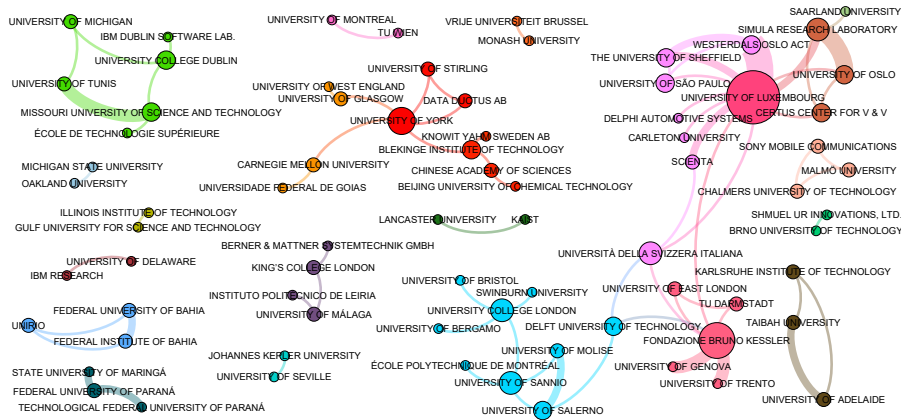


Fig. 5: Collaboration network.

try. We can observe clearly that the great majority are from universities. One interesting point is that the research foundation Fondazione Bruno Kessler contributed in almost all editions, except 2013 and 2015. The percentage of papers exclusively from universities is 87.7%, exclusively from industry and also exclusively from research foundations is 4.3%, from universities in collaboration with industry is 3.3% and from universities and research foundations is 0.5%. We noticed a modest participation of the industry.

Collaborations. We observed that 47.2% of papers have external collaboration, that is, were published by authors from different institutions, and in 28.8% the institutions are from different countries. To better identify the main SSBSE groups and collaborations we constructed a co-authorship network (Figure 5). We observed that the University of Luxembourg formed the main group, collaborating with 11 different institutions. Fondazione Bruno Kessler collaborated with 6, University of York with 5 and University College London, University of Sannio, Simula Research Laboratory and Università Della Svizzera Italiana with 4. Moreover, there are many other collaborations with fewer connections.

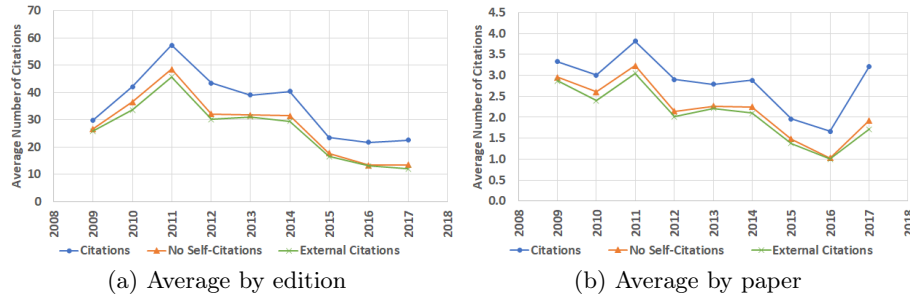


Fig. 6: Average number of citations per year.

3 RQ2 – Citations Analysis and External Impact

This section presents results regarding the total number of citations of SSBSE papers and citations by papers in order to evaluate the impact of the symposium.

The number of citations was collected from Google Scholar (GS) on 14th and 15th of March, 2019. All papers were individually evaluated, for which we collected their total number of citations (tagged as “Citations”), total number of citations excluding self-citations (tagged as “No Self-Citations”), and total number of citations excluding self-citations and citations by other SSBSE papers (tagged “External Citations”)⁹. Our citation analysis does not encompass the last edition of SSBSE, because by the time we collected this data, the citations of 2018 papers had not been computed by Google Scholar yet.

In the past 10 years, SSBSE papers have received a total of 2,080 citations, of which 1,692 (81.4%) account for no self-citations and 1,599 (76.9%) represent external citations. However, the difference between the number of no self-citations and external citations is only 93 (4.5%), i.e., there are only 93 citations of SSBSE papers by different SSBSE authors and the remaining citations are all from different authors in different venues.

Figure 6 depicts the average number of citations the papers received per year since they have been published. Figure 6.(a) shows the average citations by edition, whereas Figure 6.(b) shows the average by paper. Each SSBSE edition receives on average 35.5 citations per year (39 median), of which 27.9 are no self-citations (31.4 median), and 26.5 are external citations (29.4 median). Papers of SSBSE’11 are the most cited, considering both total number of citations and citations per year, with a total of 458 citations and 57.5 citations per year.

Next we present some statistics by paper. On average, each SSBSE paper has received 18.41 citations (11 median), of which 14.97 are no self-citations (8 median), and 14.15 are external citations (8 median). Moreover, each paper receives on average 2.83 citations per year of its publication (2.17 median), of

⁹ To ease this task, we used Publish or Perish (<https://harzing.com/resources/publish-or-perish>), a tool that helps researchers look up information about papers, conferences, journals and others researchers in several repositories, including GS.

Table 4: Ranking of the 10 most cited SSBSE papers. (C: citations, NS: no self-citations, E: external citations.)

Year	Title	Authors	C	NS	E
2011	On Parameter Tuning in Search Based Software Engineering	Arcuri and Fraser	135	121	112
2009	An Improved Meta-Heuristic Search for Constrained Interaction Testing	Garvin et al.	82	75	73
2012	Evolving Human Competitive Spectra-Based Fault Localisation Techniques	Yoo	69	56	52
2011	Highly Scalable Multi Objective Test Suite Minimisation Using Graphics Cards	Yoo et al.	68	59	56
2011	Ten Years of Search Based Software Engineering: A Bibliometric Analysis	de Freitas and Souza	59	58	55
2012	Putting the Developer in-the-Loop: An Interactive GA for Software Re-modularization	Bavota et al.	57	54	52
2009	A Study of the Multi-Objective Next Release Problem	Durillo et al.	56	51	47
2012	Reverse Engineering Feature Models with Evolutionary Algorithms: An Exploratory Study	Lopez-Herrejon et al.	52	35	33
2009	Search-Based Testing of AjaxWeb Applications	Marchetto and Tonella	51	47	46
2010	Genetic Programming for Effort Estimation: an Analysis of the Impact of Different Fitness Functions	Ferrucci et al.	49	39	37

which 2.22 are no self-citations (1.5 median) and 2.10 are external citations (1.44 median). As it happened to the cumulative number of citations of each SSBSE edition, SSBSE'11 has the most cited papers on average with 3.82 citations per year of publication, 3.23 of which are no self-citations and 3.05 are external citations. However, when we consider the median, SSBSE'11 is only the 8th in the rank with a median of 1.63, while SSBSE'17 gets the 1st position with a median of 3.00. This can be explained by the 135 citations the paper “On Parameter Tuning in Search Based Software Engineering” by Arcuri and Fraser [2] on SSBSE'11 has got. This paper single-handedly drags the average number of citations by paper per year from 2.88 to 3.82 considering all editions of SSBSE. Furthermore, the average citations per year of the 2011 edition would go down to 40.38 from 57.25 if we remove this paper from the average pool.

In fact, the paper authored by Arcuri and Fraser [2] is the most cited paper of the symposium. Table 4 shows the 10 most cited papers from all editions. It is worth mentioning that these 10 papers have 678 citations (32% of all SSBSE papers combined). Another interesting observation is that 5 of these papers are focused on Testing. This greater frequency of citations for testing papers can be explained by the greater number of testing papers in general (see Section 4).

Another interesting information is regarding SSBSE h-index and h5-index values [8]. The h-index counts the maximum number of h papers that have been cited at least h times. Similarly, the h5-index compute the h-index for the papers published in the last 5 complete years (we consider the 5 years between 2013 and 2017). The SSBSE h-index is 26 and the h5-index is 15. As a matter of comparison, according to GS, the h5-index of ACM/IEEE International Conference on Software Engineering (ICSE) is 74, IEEE Transactions on Software Engineering (TSE) is 56, IEEE Software is 37, IEEE/ACM International Conference on Automated Software Engineering (ASE) is 35 and ACM Transactions on Soft-

ware Engineering and Methodology (TOSEM) is 31. Considering only external citations, the SSBSE h-index and h5-index values are respectively 23 and 13.

This close gap between no self-citations and external citations (both count and h-index) may indicate that the SSBSE papers have got some substantial external visibility, as most of the citations are from different venues. Furthermore, this can also imply that such papers might have been used as source of inspiration for further research by the SE community.

We reported the number of citations as a measure of impact, however this might not be very accurate. As Ghezzi [4] stated in his keynote during the 31st edition of ICSE, the most cited papers will not always represent the most influential ones. Sometimes a paper is reported to be directly influenced by another paper, while having more citations. This can also be observed when comparing the rank of papers by citations count to the rank of most influential papers judged by the experts of the field. As shown by Ghezzi [4], the 8 most cited papers in their ICSE editions were elected as the most influential papers of that same year, but further down the rank, the most cited papers were not always selected as the most influential ones by experts.

This phenomenon actually happened for SSBSE. During the 10th edition of the symposium in 2018, the community was asked to vote on the most influential paper of the past 10 years. The award was given to “The Human Competitiveness of Search Based Software Engineering” by de Souza et al. [10]. However, the award-winning paper is only the 11th most cited paper (46 citations).

All in all, the number of citations seemed to be the best metric of impact in the context of our work. This metric can be of some value, as a greater number of citations can tell more than smaller numbers, even though only about the visibility of papers. The best approach to evaluate the external impact of SSBSE papers would be to actually check the experts opinion, however, that is not a trivial task. Indeed, this could be done in future work with a more carefully designed impact evaluation with experts of top-tier software engineering venues.

4 RQ3 – Software Engineering Areas and Tasks

SE Areas. To answer RQ3, the papers were grouped by SE areas¹⁰ as depicted in Figure 7. 54.4% of the papers are from Software Defect Analysis, which includes software testing and debugging, and 45.6% tackled some task related to software testing. Test data generation was addressed by papers in every SSBSE edition.

Defect prediction, test case evaluation and test management were tackled only in the first two editions. The last three editions contained papers on regression testing, stress testing, interaction testing and test suite minimization. Papers on tasks related to debugging addressed fault localization and program analysis over time and, more recently (2016 and 2017), program repair.

¹⁰ We used the four first levels of the 2012 ACM Computing Classification System (<https://www.acm.org/publications/class-2012>).

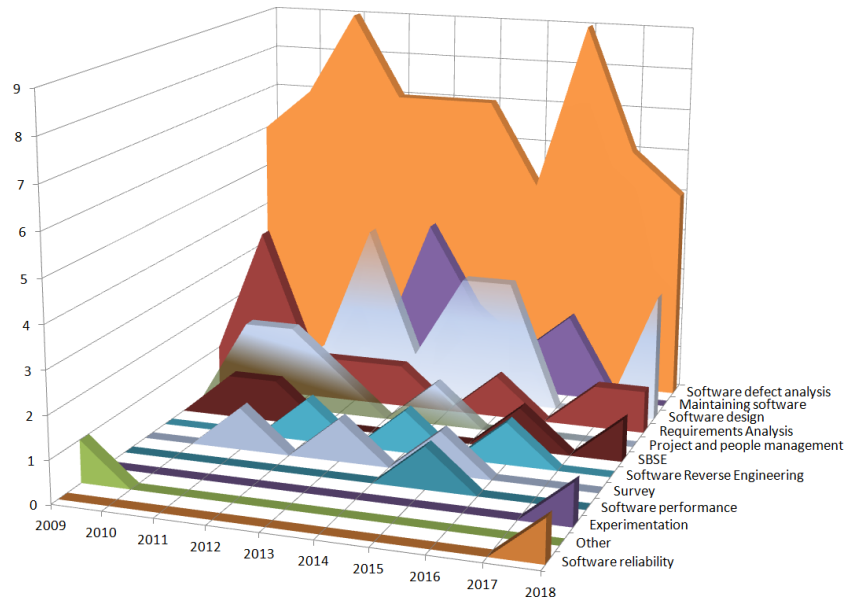


Fig. 7: Amount of Papers published by Software Engineering Area

Requirements Analysis, Software Design and Maintenance represent, respectively, 8.8%, 12% and 8% of the SSBSE papers. The Next Release Problem is the most addressed Requirements task, although the last publication about this task was in 2015. More recently, papers have focused on detection of incomplete requirements and non-functional requirements optimization. Regarding Software Design, most papers deal with architecture definition and model transformation (MDE), followed by automatic software configuration, architecture improvement and software modularization. After two years without publications in this area, three papers addressing MDE were published in 2018. Maintenance papers appeared between 2012-2016 and 55% of them addressed refactoring.

The other SE areas were focused in less than 5% of the papers. Most papers on Project and People Management deal with business process reduction and software project planning. Three surveys were published from 2011 to 2015. They addressed SBSE research analysis, metrics to search-based refactoring and software requirement selection and prioritization problems. Four papers treat SBSE over time, in the following order: SBSE evaluation, SBSE scalability, project decision making and online experimentation. Reverse engineering was applied to the software product line approach in 3 published papers. SE Areas such as Software Performance, Software Reliability, Experimentation and source-code authorship definition (identified as Other in Figure 7) had only one paper each.

Finally, tasks that have emerged in the last 4 years are the ones related to non-functional properties (software performance, software reliability, non-functional properties optimization and non-functional requirements optimization), as well as program repair, stress testing, MDE and experimentation.

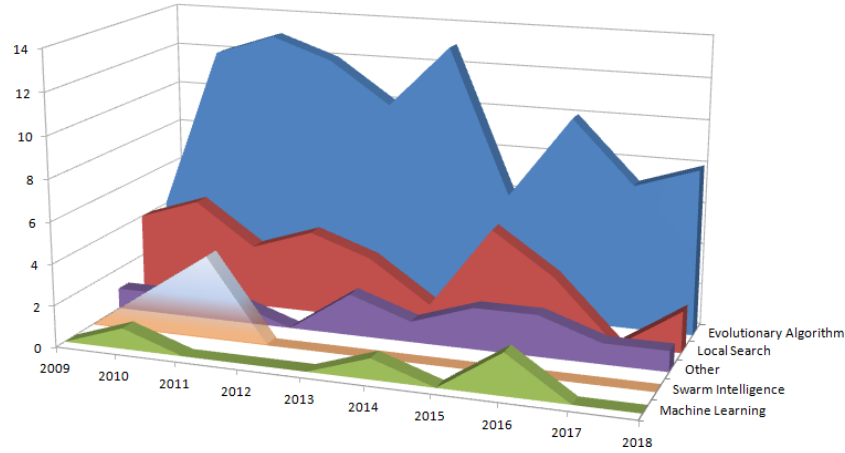


Fig. 8: Amount of Papers published by CI Technique

CI Techniques. Figure 8 shows the CI techniques used in the SSBSE papers over time. 76% of the papers applied (mono or multi-objective) evolutionary algorithms. 24% applied local search, such as Hill-Climbing, Greedy, Simulated Annealing and Tabu Search. 4.8% used swarm intelligence algorithms (ACO and PSO). The category named Other (almost 10% of the papers) includes algorithms such as Mathematical Optimization, Mixed Integer Linear programming, Error-Correcting Graph Matching algorithm, Constraint Programming, Artificial Immune Recognition Systems, Random Search, etc. 3.2% of papers have also applied Machine Learning algorithms (Artificial Neural Network, Greedy Agglomerative Clustering or Multiple Regression).

19 out 125 papers used more than one CI technique. In some cases, different algorithms were used to compare which one has the best performance to solve the addressed problem. In other cases, algorithms from different CI techniques were combined to better solve a problem, which happened with the 4 papers that combined evolutionary algorithms and machine learning. Each one addressed the following tasks: refactoring, test data generation, test management and automatic generation of maximally diversified versions.

As seen in Figure 8, since 2012 swarm intelligence has not been applied in SSBSE papers. The application of evolutionary algorithms have also decreased over time. On the other hand, other CI techniques and machine learning algorithms have been increasingly used.

5 RQ4 – Experimental Rigour

In this section, we discuss some aspects of the evaluations carried out in the SSBSE papers. We observed that 121 papers (96.6%) present evaluation results and among them 57.8% perform a statistical analysis. Almost all papers evaluate their proposed solution with a wide range of subjects and the attention on using statistical tests. Mainly in the last years, we can attest that the experimental

rigour has been taken into account by SSBSE authors. Further details about subjects and statistical tests, are presented next.

Subjects. A wide range of different subjects are used. Some of them are small computer programs typically used for educational purposes or proof of concept. Arcade Game Maker, Microwave Oven Software, and Service and Support System are examples of academic subjects. Other used subjects are real-world software, allowing an evaluation of how SBSE solutions work in practice. Such systems are in platforms like desktop (Microsoft Word and ReleasePlanner), Web (Tudu, Oryx and Softslate Commerce), mobile (Sony Mobile and Android programs), embedded software (Adaptive headlight control, door lock control and electric windows control modules), and MATLAB Simulink models. Open source projects are widely used in SSBSE papers. These projects are taken mainly from repositories such as SourceForge, GitHub, SPLOT, and Google Play. Examples of open source projects are Eclipse, Mozilla, Apache Commons project, Apache Ant, ArgoUML, Azureus, Xerces-J, JHotDraw, AJHSQLDB, Health Watcher, Toll System, JFreeChart, Rhino, and GanttProject.

Publicly available datasets and benchmarks also appeared in the evaluations. Example of a benchmark set is the one with faulty programs originally developed by Siemens: `print_tokens`, `replace`, `schedule`, `schedule2`, `tcas`, `tot_info`, and SF110 dataset. We also observed the use of synthetic data, non-real artifacts, sometimes randomly generated, used to represent difficult problems or large artifacts, allowing to expose the power of SBSE solutions.

Statistical tests and Effect size measures. Considering that SBSE approaches rely on CI techniques, which employ randomness in their search process, commonly proposed approaches are executed many times to identify a standard behaviour. The collected data results are evaluated with statistical tests to assess whether there are significant difference among results or not. Table 5 presents the Statistical tests and Effect size measures that were applied in at least 2 SSBSE papers. The last column of the table shows the number of papers using the tests. Among the 17 found tests/measures, Mann-Whitney-Wilcoxon U-test, Wilcoxon Ranked Sum Test, and Vargha-Delaney A12 effect size were by far the most commonly used. In Table 6 we can see the percentage of papers along the years which used those tests/measures. Tests and measures have been used since 2009, however, we can observe that after 2014 they have been used more frequently. In 2017 all papers used these tests/measures.

6 Recommendations to future SSBSE authors

During the screening of the 125 papers, we have realized that some pieces of information are not presented in several papers, what makes the SBSE approaches not completely clear to readers. Following, we present some recommendations to future SSBSE authors aiming at helping them to develop high quality studies, to improve text readability and to enable study replication.

Table 5: Statistical tests and Effect size measures used

Test/Measure	#Papers
Mann-Whitney-Wilcoxon U-test	26
Wilcoxon Ranked Sum Test	25
Vargha-Delaney A12 effect size	22
Student's T-test	5
Cliff's Delta	3
Kolmogorov-Smirnov test	3
Spearman's RC coefficient	3
Friedman test	2
Kruskal-Wallis	2
Two-Tailed Test	2

Table 6: Percentage of papers that apply statistical tests or Effect size per year

Year Percentage	
2009	11.1
2010	64.3
2011	46.7
2012	26.7
2013	28.6
2014	85.7
2015	83.3
2016	69.2
2017	100
2018	75

- Make it clear the ingredients of SBSE approaches that enable the application of CI techniques to solve the corresponding SE problem: problem representation and fitness function(s). Some (meta)heuristics also need operators to modify candidate solutions, which should also mentioned in the text;
- Illustrative examples so that readers can easily understand the problem and the proposed solution.
- Make it clear which are the CI techniques and algorithms used, not just mention the tool or framework name;
- For the evaluation, authors should prefer using real-word systems from different domains and sizes. This would make the findings more general;
- To avoid threats regarding randomness of CI techniques, run your approaches many times (at least 30 runs) and assess the results with statistical tests and effect size measures;
- Make the experimental package available providing, as much as possible, ways to other authors replicate your study and/or to ease comparison.

7 Conclusion

In this paper we presented an overview of the ten-year history of SSBSE as well as results from a systematic mapping involving all full papers of the ten proceedings. Our findings allow us to state that SSBSE papers have made some external impact on the SE research community. We found that most of the citations are from different venues and identified a close gap between no self-citations and external citations. This indicates SSBSE papers have got some substantial external visibility.

The women SSBSE leadership participation is rather good, but authorship is low. Gender bias is a major concern in software engineering discipline. Gender diversity is important because it can help sharing different skills, points of view and experiences, bringing and incorporating gender aspects of the customers and users in software engineering, expanding potential talents, among other aspects. Hence, increasing the participation of women in the symposium is of great value.

Regarding the area of SE problems solved with CI techniques, software testing is still the main addressed task, but other problems have emerged, mostly

related to non-functional requirements, program repair, MDE and experimentation. Evolutionary algorithms remained the most used CI technique.

We could observe along the ten years a wide range of subjects used by authors to evaluate their approaches. These subjects can also be used in new research. Besides, in recent years authors are paying more attention to the use of statistical tests to better evaluate their results. But it is important to increase industry participation and the creation of repositories containing benchmarks regarding the different SBSE sub-areas.

To call attention and guide new authors willing to publish their papers and to participate in SSBSE, we presented a set of recommendations to improve their publications on understandability, replicability, and experimentation soundness. However, the recommendations are limited to what we observed during the papers screening. Also, our findings are limited to SSBSE editions and they were not compared to other venues, which might be done in future studies.

SSBSE has been a representative venue to divulge studies and put together academics, researchers and practitioners to discuss SBSE. Currently, the SBSE field is explicitly listed as a topic of interest of important conferences and journals. Given what we reported and discussed in this paper, we can state that SSBSE has helped increase the popularity of SBSE in the SE research community and plays an important role to strengthen SBSE over the past ten years.

References

1. Agarwal, S., Mittal, N., Katyal, R., Sureka, A., Correa, D.: Women in computer science research: What is the bibliography data telling us? *SIGCAS Comput. Soc.* **46**(1), 7–19 (2016)
2. Arcuri, A., Fraser, G.: On parameter tuning in search based software engineering. In: SSBSE. pp. 33–47 (2011)
3. Freitas, F.G., Souza, J.T.: Ten years of search based software engineering: A bibliometric analysis. In: Cohen, M.B., Ó Cinnéide, M. (eds.) SSBSE. pp. 18–32 (2011)
4. Ghezzi, C.: Reflections on 40+ years of software engineering research and beyond an insider's view. In: <https://www.cs.uoregon.edu/events/icse2009/keynoteSpeakers/ICSEkeynote.pdf> (2009)
5. Harman, M., Jia, Y., Zhang, Y.: Achievements, open problems and challenges for search based software testing. In: International Conference on Software Testing, Verification and Validation (2015)
6. Harman, M., Mansouri, S.A., Zhang, Y.: Search Based Software Engineering: A Comprehensive Analysis and Review of Trends Techniques and Applications. Tech. rep., Department of Computer Science, King's College London (2009)
7. Harman, M., Mansouri, S.A., Zhang, Y.: Search-based Software Engineering: Trends, Techniques and Applications. *ACM Comput. Surv.* **45**(1), 1–61 (2012)
8. Hirsch, J.E.: An index to quantify an individual's scientific research output. *National Academy of Sciences* **102**(46), 16569–16572 (2005)
9. Petersen, K., Vakkalanka, S., Kuzniarz, L.: Guidelines for conducting systematic mapping studies in software engineering: An update. *Information and Software Technology* **64**, 1 – 18 (2015)
10. Souza, J.T., Maia, C.L., de Freitas, F.G., Coutinho, D.P.: The human competitiveness of search based software engineering. In: SSBSE. pp. 143–152 (2010)