EasyChair Preprint

№ 4111

A Bibliometric Overview of the International Symposium on Symbolic and Numeric Algorithms for Scientific Computing Between 2005 and 2018

Teodor-Florin Fortis and Alexandra Emilia Fortis

August 30, 2020

A bibliometric overview of the International *Symposium on Symbolic and Numeric Algorithms for Scientific Computing* between 2005 and 2018

Teodor-Florin Fortiș
Department of Computer Science
West University of Timișoara,
Faculty of Mathematics and Informatics,
Timișoara, Romania
https://orcid.org/0000-0002-3143-8908

Alexandra-Emilia Fortiș
Department of Computer Science
West University of Timișoara,
Faculty of Mathematics and Informatics,
Timișoara, Romania
https://orcid.org/0000-0002-7899-3415

***Abstract*—Current research offers a bibliometric overview of the International *Symposium on Symbolic and Numeric Algorithms for Scientific Computing*, from 2005 to 2018, from different perspectives, in order to highlight the generated impact, the dimensions and strength of international collaborations, as well as a statistical study of conference papers, typical structure of collaboration groups, evolution of research trends, and others. Associated findings are presented either as raw data, or processed via VOSViewer.**

***Index Terms*—bibliometrics, conference analysis, VOSViewer, Web of Science, Scopus**

I. Introduction

The International *Symposium on Symbolic and Numeric Algorithms for Scientific Computing* (SYNASC) started in 1999 as "a workshop that aims to be an interaction forum between the two communities of symbolic and numeric computing", under the organization of the Faculty of Mathematics, West University of Timișoara, Romania and the Research Institute for Symbolic Computation RISC, from Johannes Kepler University, Linz, Austria. Currently at its 22^{nd} edition, SYNASC is an international event with a more diverse list of research topics, held each year in September, Timișoara, Romania.

With a relatively constant number of contributions, the number of participating researchers and countries has grown since 1999, along with the diversification of the research topics covered. Currently, the tracks of the conference and its satellite events cover topics from symbolic computation, numerical computing, logic and programming, distributed computing, advances in the theory of computing, artificial intelligence, among others.

Initially published as special issues of the "*Annals of the University of Timișoara*", Mathematics and Informatics series, starting with its 7th edition the conference proceedings are edited and indexed by the IEEE Computer Society Press, later by Conference Publishing Service (CPS), with full coverage and indexing in Web of Science and Scopus since 2007.

TABLE I
THE SYNASC 2005-2018 INFORMATION, AS EXTRACTED FROM SCIMAGOJR

Year	SJR	DOC	CIT	CITD	INT
2008	0.138	91	28 (0.308)	23	6.45%
2009	0.164	60	27 (0.45)	16	16.13%
2012	0.276	71	74 (1.042)	32	13.7%
2014	0.183	75	46 (0.613)	24	16.88%
2015	0.158	62	36 (0.581)	24	12.5%
2016	0.147	62	59 (0.952)	28	18.46%

A. Motivation

The *International Symposium on Symbolic and Numeric Algorithms for Scientific Computing* (SYNASC) was included in various scientific repositories, including IEEE Xplore, Web of Science and Scopus, since 2005, with a constant coverage since then. Based on existing Scopus information, Scimago [1] was able to extract conference data and compute their journal ranking indicator (Scimago Journal Ranking – SJR) since 2008. The information included in ScimagoJR is covering targeted indicators and offer partial coverage for the different events associated with SYNASC.

The computed values of these indicators are included in Table I, thus offering a broad image of the different conference editions. Even if there is no relation between the different conference events included in the ScimagoJR analysis, and for some editions there is no available/computed information, we can notice that there is a relatively constant number of documents (DOC), there is a growing tendency for the 3 years citation rate (CIT), even if the number of citable documents (CITD) is usually less than 30% of total documents (DOC). Also, there is an increasing rate of articles coming from international collaborations.

On the other hand, we may notice that the most recent SJR value that was computed for a SYNASC event was 0.195[1],

[1]SYNASC 2016, https://www.scimagojr.com/journalsearch.php?q=21100797845&tip=sid&clean=0

thus ranking in top 900 of 2366 entries. Worth mentioning that the *2015 IEEE/ACM 19th International Symposium on Distributed Simulation and Real Time Applications*[2] is situated immediately after the SYNASC 2016 event in ScimagoJR, while the *IEEE 14th International Symposium on Network Computing and Applications*[3] have a SJR score of 0.199, suggesting that there is a good quality of the generated impact of the conference papers.

We may notice that SYNASC, which is a C ranked conference in CORE2018 and a 'National' classified outlet in CORE2020, is comparable in terms of recent SJR or Google Scholar h5-index computed values with higher ranked events, like DS-RT (a B ranked CORE2020 conferences, with a h5-index value of 10). However, CORE conference rankings are based on "a mix of indicators, including citation rates, paper submission and acceptance rates, and the visibility and research track record of the key people hosting the conference and managing its technical program."[4] It is out of the scope of this paper to evaluate the set of indicators used for various conference rankings.

B. Bibliometric investigations

Bibliometric studies have been carried out to provide different perspectives related with a specific scientific outlet: journal or conference. For the case of journals, these investigations are usually oriented towards the identification of the leading trends for a specific period of time (like in [2]–[5]), analysis of the impact of a scientific outlet (e.g., an analysis for co-authorships and citations in [2], a citation and publication landscape in [4]), or as a means of comparison between similar scientific outlets or analysis of a scientific domain (like [3] with a comprehensive bibliometric analysis on computer networking research, or [6] with an analysis of fuzzy techniques in Big Data).

For carrying out such investigations, miscellaneous tools for bibliometric analysis, statistical approach and data visualization have to be used on the large amount of metadata associated to scientific outputs, such those provided in [3], [4], [7], [8]. Results arising from these investigations can support various management strategies, to improve visibility and accessibility of research data, to establish and consolidate research groups and to better share experiences and expertise on specific domains.

The study proposed in our paper offers insights on the growth and evolution of SYNASC 2005–2018 from different perspectives. An initial set of findings offers an overview of the conference impact and coverage, based on existing WoS and Scopus data. A second perspective is based on the collaboration measures, emphasizing the structure of the typical conference paper, and its evolution during the 14 editions under investigation, that could be used by both organizers and contributors. These findings are complemented by some semantic investigations, in order to highlight the international dimensions of the conference and the strength of the established international collaboration.

TABLE II
NUMBER OF RETRIEVED DOCUMENTS PER YEAR

Year	IEEE Xplore	WoS	Scopus
2005	79	67	67
2006	80	68	68
2007	89	78	82
2008	101	93	91
2009	72	62	60
2010	94	84	83
2011	67	58	57
2012	80	71	71
2013	85	75	73
2014	89	79	79
2015	79	68	67
2016	78	63	63
2017	80	70	70
2018	79	N/A	63

II. DATA COLLECTION

We describe now the data sources that were considered for our analysis, and data collection methodology. Most of the SYNASC entries are classified as 'conference papers', while there are several entries classified as 'conference review' or 'editorial', entries which were excluded from the various data sources.

A. Data sources

To perform our analysis, we investigated the available options for three different data sources, in order to discover and obtain the most relevant bibliographic data and associated metadata. These data sources include IEEE Xplore[5], Scopus[6] and Clarivate Web of Science[7].

IEEE Xplore: The first choice was for IEEE XPlore, as this database offers the best available coverage for SYNASC 2005–2018 papers. In the case of IEEE XPlore, a query for 'Conferences' with the query string 'SYNASC' was performed, retrieving the entire set of available entries. The query returned 1152 entries for the reference time interval of 2005-2018, with options to export the full set of information as CSV (comma separated values), or a limited set of entries in RIS (Research Information Systems) format. While the resulted download offers a limited amount of metadata, we are using this information only as a reference.

Web of Science: In the case of Clarivate Analytics database, Web of Science (WoS), the query that was used is `CF=(Symposium on Symbolic and Numeric Algorithms for Scientific Computing) AND PY=(2004-2018)`. By using this query, we were able to fetch a set of 936 entries ('proceedings paper'). The export option is available for 'Other file formats' with record content defined as 'Full record and Cited References', but this is

[2]DS-RT 2015, ranked B CORE2020
[3]NCA 2015, ranked A CORE2020
[4]https://www.core.edu.au/conference-portal
[5]https://ieeexplore.ieee.org/search/searchresult.jsp?queryText=synasc
[6]https://www.scopus.com/
[7]http://www.webofknowledge.com/

limited to 500 entries at once. Consequently, in the case of Web of Science, we defined two distinct queries in order to retrieve the full set of available records.

Scopus: The query that was used for retrieving data from Scopus was `SRCTITLE(SYNASC) AND PUBYEAR > 2004 AND PUBYEAR < 2019 AND ( LIMIT-TO ( DOCTYPE,"cp" ) )` , which returned a set of 994 documents classified as 'Conference Paper', with an option to export the full set of available data, including 'Citation information', 'Bibliographic information', 'Abstract and keywords', and 'References'.

B. Data preparation

Further processing of the exported data sets was required for both Web of Science and Scopus, as the two databases link entries with publishing year, which was not necessarily the year when the conference took place. We may notice that the number of retrieved documents, as specified in Table II is consistent for the different indexing databases, with slightly higher values in the case of IEEE XPlore.

Information was retrieved as Tab-delimited values from Web of Science and comma separated values (CSV) from Scopus. Corrections were made for the reported conference year/volume. Additionally, the full author list was used in order to extract the full list of author data for each document in the collection.

III. Results

In this section we present some basic findings regarding the SYNASC editions, included in the analysis period of 2005–2018. First, we have a set of initial findings that are based on raw data, as extracted from Web of Science and Scopus, offering a statistical overview of the conference for the entire interval.

The next set of results are built on top of Scopus data, as we were able to collect more fields from this repository, and cover additional results related with general information about the conference: the distribution of papers, distribution of the number of papers by authors, or various geographical distributions. For this set of results, together with the overall analysis we are including results based on three different intervals of SYNASC events (namely, 2005–2010, 2011–2014, and 2014–2018), in order to capture the tendencies in the evolution of the conference.

A final set of results include additional findings that can be extracted from the Scopus dataset, such as co-authorship, co-occurrence of keywords, co-citation links, and others. These findings will allow us to identify collaboration patterns in the context of SYNASC events.

A. Overview of results

An initial set of results can be extracted directly from the raw data obtained from the various sources. These initial findings can offer an overview of the various dimensions of the conference: coverage of authors, institutions and countries, as well as an initial estimation of SYNASC 2005–2018 impact.

TABLE III
H-INDEX VALUES AND CITATION RATE FOR SYNASC 2005–2018

Interval	h-index and citation rate		h5-index
	WoS	Scopus	GS
2005–2010	10 (*1.36*)	15 (*3.29*)	–
2011–2014	10 (*1.86*)	15 (*3.31*)	–
2015–2018	6 (*0.76*)	8 (0.81)	–
2005–2018	12 (*1.38*)	19 (*2.64*)	11

1) Conference Impact: We are going to have an initial evaluation of conference impact based on the estimated h-index, based on Web of Science and Scopus data, and presented in Table III. Together with these results we list the top cited documents, which are included in the composition of the h-index for the year intervals under analysis, in Table IV. For our analysis we limited citations to 2018[8]. We may notice that, as there is a larger coverage for Scopus, the total number of citations and the computed values for h-index are higher than in the case of Web of Science. Additionally, there is a higher citation rate for documents included in SYNASC 2011–2014, compared with the other two time intervals (2005–2010 and 2015–2018), as it is shown in Table V and Table VI.

TABLE IV
TOP 5 CITED DOCUMENTS

Document title	Citations (rank)	
	WoS	SCO
Proposal of Business Process and Rules Modeling with the XTT Method (2007)	22 (1)	28
Optimization of resource scheduling in cloud computing (2010)	7	53 (1)
Automatic State-Based Test Generation Using Genetic Algorithms (2007)	19 (2)	39 (3)
Synthesis from examples: Interaction models and algorithms (2012)	10	44 (2)
Optimized zero false positives perceptron training for malware detection (2012)	17 (3)	25
Archeology of code duplication: Recovering duplication chains from small duplication fragments (2005)	17 (4)	30 (4)
Efficient Computation of the Isotropy Group of a Finite Graph: a Combinatorial Approach (2013)	16 (5)	18
Proposal of business process and rules modeling with the XTT method (2007)	22	28 (5)

2) Geographic coverage: The entire set of SYNASC 2005–2018 documents counts authors from around 48 different countries or regions, of which for 22 countries there are at least 8 submitted documents. An overview of top contributing countries is described in Fig. 1. As Romania is by far the largest contributor for this type of investigation, counting authors or co-authors for around 577 of 994 documents (based on Scopus data), the country was not considered for this analysis.

When we limit our investigation for geographic coverage to the set of documents receiving at least 10 citations – set

[8]Currently, due to some subscription disruptions, we cannot access WoS data after December 31st, 2018

TABLE V
TOP CITED DOCUMENTS (WEB OF SCIENCE, PER YEAR INTERVAL)

Document title	Citations
2005–2010	
Proposal of Business Process and Rules Modeling with the XTT Method (2007)	22
Automatic State-Based Test Generation Using Genetic Algorithms (2007)	19
Archeology of code duplication: Recovering duplication chains from small duplication fragments (2005)	17
Enhanced Rule-based Phonetic Transcription for the Romanian Language (2009)	13
2011–2014	
Optimized zero false positives perceptron training for malware detection (2012)	17
Efficient Computation of the Isotropy Group of a Finite Graph: a Combinatorial Approach (2013)	16
Cloud4SOA: Multi-Cloud Application Management Across PaaS Offerings (2012)	16
On computing mesh root systems and the isotropy group for simply-laced Dynkin diagrams (2012)	15
2015–2018	
Gesture Recognition on Kinect Time Series Data Using Dynamic Time Warping and Hidden Markov Models (2016)	13
Using Machine Learning to Decide When to Precondition Cylindrical Algebraic Decomposition With Groebner Bases (2016)	8
A CUDA Implementation of the Standard Particle Swarm Optimization (2016)	8
Malware Classification Based on Dynamic Behavior (2016)	7

TABLE VI
TOP CITED DOCUMENTS (SCOPUS, PER YEAR INTERVAL)

Document title	Citations
2005–2010	
Optimization of resource scheduling in cloud computing (2010)	53
Automatic state-based test generation using genetic algorithms (2007)	39
Archeology of code duplication: Recovering duplication chains from small duplication fragments (2005)	30
Proposal of business process and rules modeling with the XTT method (2007)	28
2011–2014	
Synthesis from examples: Interaction models and algorithms (2012)	44
Supporting the development and operation of multi-cloud applications: The MODAClouds approach (2013)	26
Assessing SLA compliance from Palladio component models (2013)	25
Optimized zero false positives perceptron training for malware detection (2012)	25
2015–2018	
Evaluating weighted round robin load balancing for cloud web services (2015)	24
A streamlined difference ring theory: Indefinite nested sums, the alternating sign, and the parameterized telescoping problem (2015)	15
Gesture recognition on kinect time series data using dynamic time warping and hidden markov models (2017)	14
Business Reviews Classification Using Sentiment Analysis (2016)	13

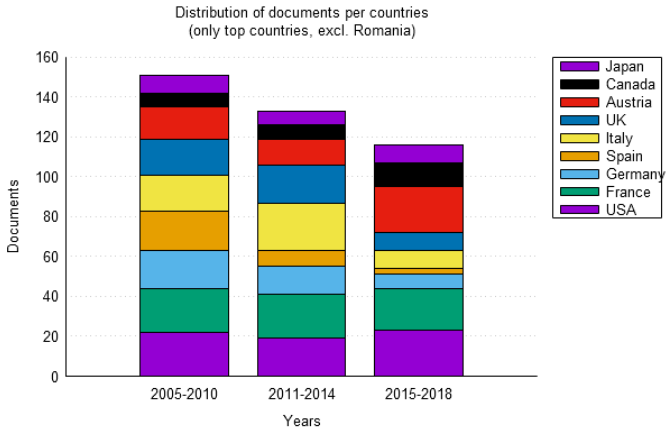


Fig. 1. Country distribution for SYNASC 2005–2018 documents

which include a total of 80 documents, based on Scopus data, – we may notice that these documents register authors from 23 different countries or regions, and for 9 of these countries there are at least four highly cited documents. The leading country is, again, Romania, with authorship or co-authorship for 33 of the 80 documents. An overview of this additional investigation is figured in Fig. 2. Top authors in this set include *Ardagna, D.* (4 documents, Italy), *Di Nitto, E.* (3, Italy), *Davenport, J.H.* (3, UK), *Casale, G.* (3, UK), *England M.* (3, UK), *Simson, D.* (3, Poland), *Paprzycki, M.* (3, Poland), and *Bădică, C.* (3, Romania).

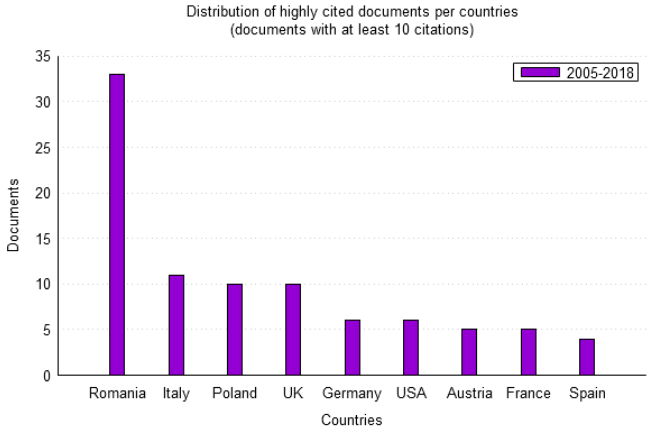


Fig. 2. Country distribution for SYNASC 2005–2018 (top cited papers).

3) Distribution of papers and collaboration measures: An interesting analysis that can be performed links authorship with the number of papers. For SYNASC 2005–2018, a total of 2511 authors published 995 papers (according to Scopus). Only 22% (219 of 995) of the total number of papers have one author, around 71.55% (712 of 995) have two to four authors, while for about 6.43% (64 of 995) there are at least five authors recorded. Interesting to notice that two papers are registered with twelve authors. The mean number of authors per paper is 2.524. The distribution of papers based on the

TABLE VII
DOCUMENT STATISTICS AND COLLABORATION METRICS FOR SYNASC 2005–2018

Year	Papers	Number of authors per paper					Total authors	Measures		
		1	2	3	4	>4		CI	DC	CC
2005	67	16.42%	44.78%	25.37%	4.48%	8.96%	167	2.493	0.836	0.500
2006	68	22.06%	35.29%	25.00%	13.24%	4.41%	168	2.471	0.779	0.479
2007	82	21.95%	36.59%	20.73%	17.07%	3.66%	203	2.476	0.780	0.479
2008	91	21.98%	37.36%	28.57%	8.79%	3.30%	213	2.341	0.780	0.470
2009	60	31.67%	28.33%	16.67%	16.67%	6.67%	144	2.400	0.683	0.432
2010	84	13.10%	30.95%	22.62%	23.81%	9.52%	252	3.000	0.869	0.563
2011	57	21.05%	38.60%	29.82%	5.26%	5.26%	137	2.404	0.789	0.475
2012	71	18.31%	35.21%	29.58%	8.45%	8.45%	186	2.620	0.817	0.507
2013	73	17.81%	36.99%	28.77%	9.59%	6.85%	193	2.644	0.822	0.506
2014	79	29.11%	18.99%	32.91%	11.39%	7.59%	200	2.532	0.709	0.462
2015	67	22.39%	29.85%	28.36%	10.45%	8.96%	176	2.627	0.776	0.491
2016	63	26.98%	28.57%	23.81%	15.87%	4.76%	155	2.460	0.730	0.460
2017	70	28.57%	30.00%	25.71%	10.00%	5.71%	166	2.371	0.714	0.443
2018	63	19.05%	42.86%	23.81%	7.94%	6.35%	151	2.397	0.810	0.483

number of authors is synthetized in Table VII and illustrated in Fig. 3a.

In order to have a better understanding of some collaboration patterns, several metrics were computed, as depicted in Table VII, and their graphical evolution is exposed in Fig. 3b. The metrics used for our investigation include the degree of collaboration (DC, [9]), the collaborative index (CI, [10]), and the collaboration coefficient (CC, [11]). These metrics are computed as follows:

$$\begin{aligned} CI &= \frac{\sum_{j=1}^{A} j f_j}{N}, \\ DC &= 1 - \frac{f_1}{N}. \\ CC &= \frac{\sum_{j=1}^{A} 1/j f_j}{N} \end{aligned}$$

where f_i represents the number of papers authored by i authors, N is the total number of research papers and A is the largest number of authors per paper. One may notice that the value of CI is, in fact, the mean number of authors per paper.

In order to understand the information behind the three measures, we have the following observations:

- when there is no collaboration, we have $DC = 0.00$, $CC = 0.00$ and $CI = 1.00$;
- when most of the entries are with two authors, the values are close to $DC = 1.00$, $CC = 0.50$, and $CI = 2.00$;
- for three authors, the threshold is $DC = 1.00$, $CC = 0.667$, and $CI = 3.00$;
- while for four authors the target values are $DC = 1.00$, $CC = 0.75$, $CI = 4.00$.

For the case of SYNASC 2005–2018, the computed values are $CI = 2.524$, $DC = 0.780$ and $CC = 0.484$, we can conclude that even if the mean number of authors is 2.524 (an indication that most of the papers have one to four authors), the collaboration coefficient indicates (CC) suggests that most of the registered papers are with two authors, while the high value of the degree of collaboration (DC) indicates that the majority of documents in a specific year are a combination of one-two or two-three authored documents. Also, as it is indicated in Fig. 3b, this pattern applies for all of the SYNASC editions in the analyzed interval.

B. Conference trends

Next, we are going to offer investigations over several semantic indicators, and exploit VOSViewer's[9] capabilities [7], [8]: co-authorship, bibliographic coupling and co-occurrences of keywords. As our study was primarily oriented towards the international dimensions of SYNASC 2005–2018 events, the co-citation analysis and bibliographic coupling are directed towards the discovery of SYNASC collaborations, while the investigation based on the co-occurrence of keywords can be used for finding the major research trends and their evolution.

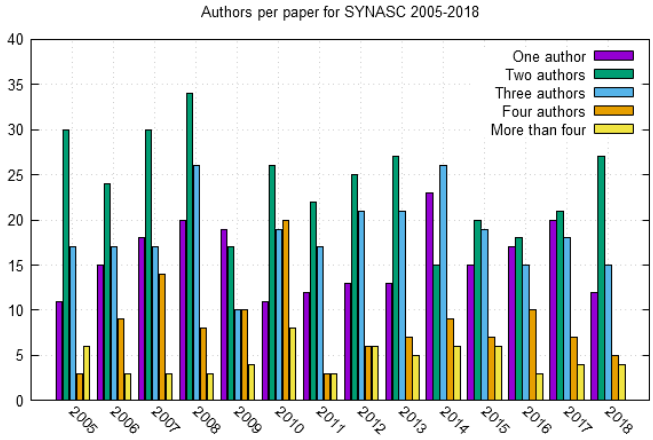
1) Co-authorship (countries): In the case of co-authorship analysis, one can link researchers, institutions, countries or other entities based on the number of publications they have together, as an indicator of existing international scientific collaboration [12]. In order to discover the strength of international collaborations, we used this type of analysis at countries level.

Two different perspective were used to capture these international collaborations:

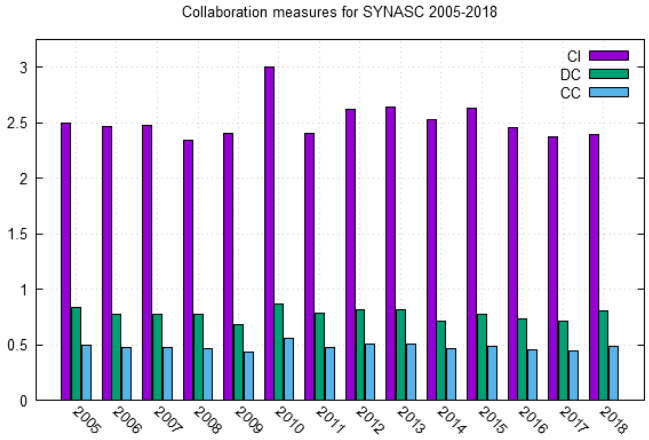
- co-authorship (countries), based on the set of SYNASC 2005–2018 documents. For this perspective we selected countries with at least 20 entries, 11 of 63 countries met this criterion and were included in Fig. 4a.
- co-authorship (countries), based on the set of documents citing SYNASC 2005–2018 papers. For this perspective we selected countries with at least 40 entries, 16 of 94 countries met this criterion and were included in Fig. 4b.

Even if the discovered cluster composition is slightly different, one can notice that there is an important overlap for the identified list of countries, with similar collaboration patterns, and these findings also aligned with the results previously collected in Fig. 1. Based on the identified clusters from the

[9]https://www.vosviewer.com/

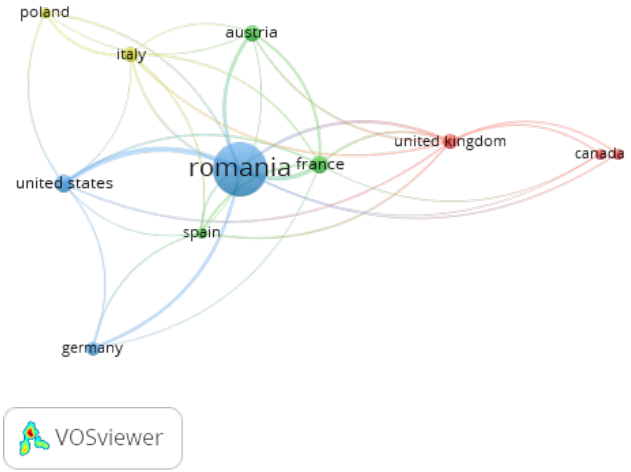


(a) Evolution of the number of authors per paper.

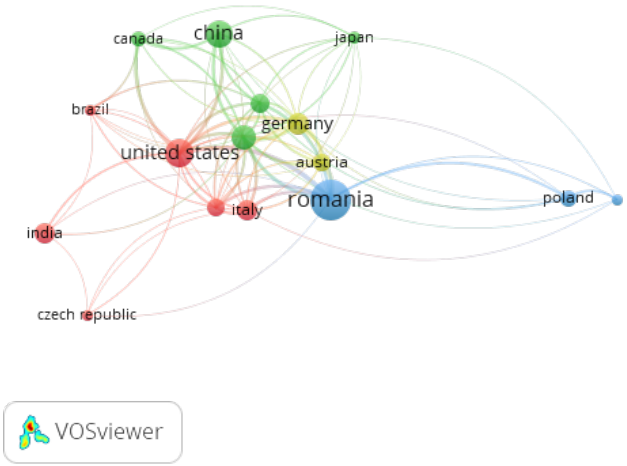


(b) Evolution of the collaborative measures.

Fig. 3. Authorship measures for SYNASC 2005–2018.



(a) Co-authorship (countries) based on SYNASC 2005–2018 documents.



(b) Co-authorship (countries) based on documents citing SYNASC 2005–2018 papers

Fig. 4. Comparative study for co-authorship (countries).

two analyzed perspectives are, we can emphasize the core set of countries with a constant interest for SYNASC: Austria, Canada, France, Germany, Italy, Japan, Poland, Romania, Spain, UK, and USA.

2) Bibliographic coupling: Like co-citations, which was described as "*a new form of document coupling [...] defined as the frequency with which two documents are cited together*" [13], bibliographic coupling can be used establish the relatedness of two documents [14]. However, in order to capture the international dimensions of the conference, we used bibliographic coupling to measure the relatedness of participating countries, offering an indication of the common research interests, when the country bibliographic coupling strength is higher.

Given the similarity results obtained for the co-authorship (countries) analysis, for the case of bibliographic coupling we consider a hybrid approach, which is built over the union between the set of documents for SYNASC 2005–2018 and the set of papers citing SYNASC documents. For this representation we considered only countries with at least 30 documents and at least 30 citations, resulting in an extended list of 33 countries. The results of this analysis and the corresponding clusters are presented in Fig. 5, with a clear view of the previously identified core set of countries.

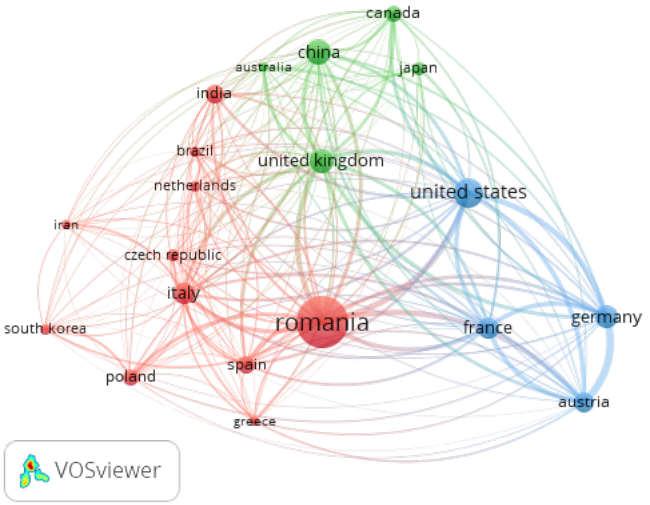


Fig. 5. The hybrid approach for bibliographic coupling.

3) Co-occurrences keywords: The co-occurrences analysis on the set of keywords extracted from the SYNASC 2005–2018 documents can be used to discover the major research areas and research trends, and to estimate the impact that was generated by the research interests and collaborations on conference publications. By using such an analysis, one can follow the evolution of the research interest in conference publications.

Initially, the analysis on the co-occurrence of keywords was carried out on the main set of conference data, limited to keywords with at least 20 occurrences (from the set of 7466 identified keywords). This initial investigation was figured in Fig. 6a. Additionally, we repeated this investigation for each of the three time periods, with lower filtering values: keywords with 10 occurrences for 2005–2010 documents (from a total of 3648 keywords); 7 occurrences for 2011–2014 (2645 keywords) and 2015–2018 (2544 keywords).

Following these investigations, we can identify several clusters, which follow the development of the different conference events, with 'computation theory' as one of the major keywords, followed by keywords from distributed computing and applications (such as 'grid computing' before 2010, 'cloud computing' after 2010), theoretical computer science and mathematics (such as 'graph theory', 'set theory' before 2010, 'semantics', 'model checking' after 2010), and a growing interest for artificial intelligence especially after 2010.

An additional investigation which was based on co-occurrences of keywords was carried out on the full set of documents citing SYNASC 2005–2018 papers. Even if the filtered set of keywords is larger (the threshold was set on 45 occurrences to select the most frequently used keywords from the set of 15111), we can notice from Fig. 6 that there is a nice overlap between these findings and the previously identified ones, showing that the generated impact by SYNASC 2005–2018 is uniformly distributed among the various conference research topics.

IV. Conclusions

In this paper we developed a bibliometric overview of of the International *Symposium on Symbolic and Numeric Algorithms for Scientific Computing* (SYNASC), from 2005 to 2018. While in the CORE rankings portal, the rank of SYNASC was C, for the analyzed year interval[10], Scimago Journal Ranking data suggests that the impact of the conference for specific years is similar with that of higher ranked conferences (DS-RT or NCA, according to Scimago web-site). This performance was possible due to the uniformly distributed impact that was generated by conference papers, both from the point of view of year coverage and spread across research domains. However, the CORE ranking exercise

Hence, we first investigate the typical structure of SYNASC papers and the composition of the set of papers which offer the necessary citations for the h-index value of the conference. The geographical distribution of conference papers, as well as of the set of h-index papers are investigated, offering an indication that there is a constant interest for conference outputs. Next, an analysis of several collaboration measures reveals the average number of authors per conference paper (measured by the collaborative index – CI), and that most of the papers records one to four authors, with minor variations in recent years. However, there is no impact of these variations on the degree of collaboration (DC) or the collaboration coefficient (CC).

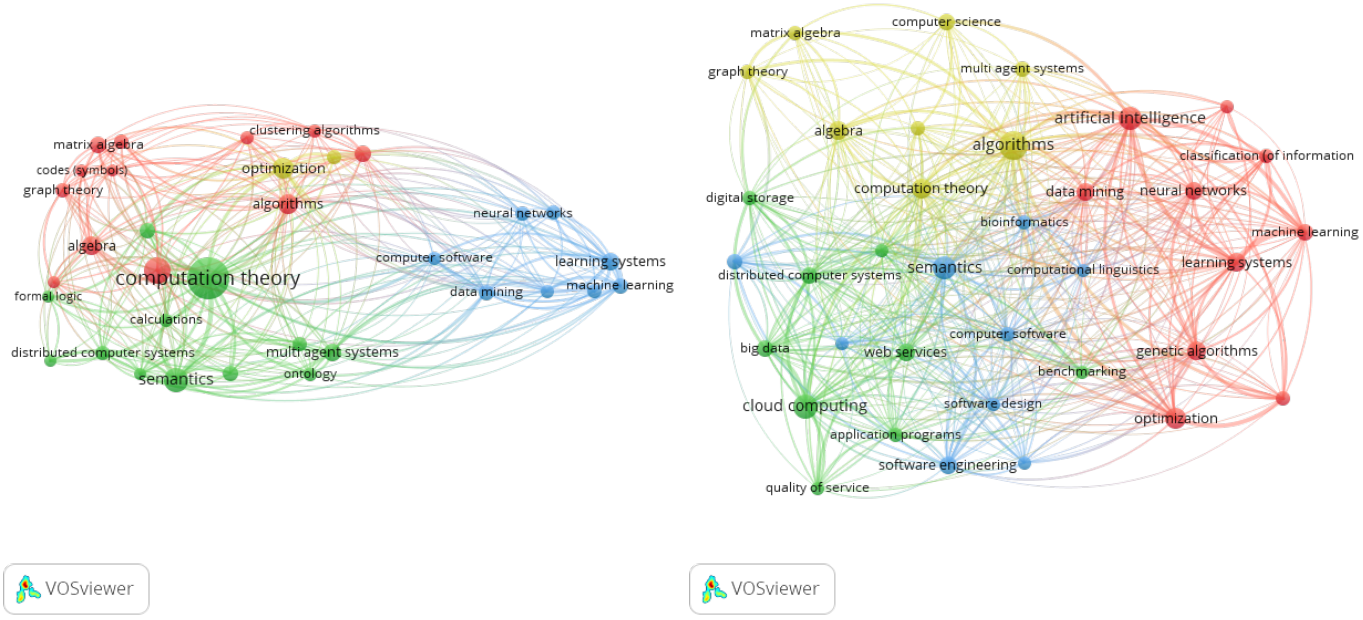
With an increased number of papers authored by at least two researchers, we focus next on the dimensions of international collaborations (country level). The investigations we used are first directed towards the discovery of co-authorship patterns, with a confirmation that previous country collaborations can be discovered in more recent SYNASC events. Next, an investigation based on bibliographic coupling reveals that there are strong links between the previously discovered clusters of country, a confirmation of common research interests between participating institutions. Finally, the co-occurrence analysis was carried out to reveal those common research interests and the major research trends that can be linked with the various conference events.

More detailed research based on conference data is considered for future extensions of this paper, including a more detailed analysis for co-authorship and bibliographic coupling, in-depth citation and co-citation investigations, with details at author, institution or institution level. In order to support these activities it will be required to improve the quality of data collected from both Web of Science and Scopus.

Acknowledgment

The research leading to these results was partially supported by the BID project (Big Data Science), PNIII-28-PFE.

[10]http://portal.core.edu.au/conf-ranks/1421/, ranked C until CORE2018, currently ranked as a 'National' outlet



(a) Co-occurrence of keywords in SYNASC 2005–2018 documents (b) Co-occurrence of keywords in documents citing SYNASC 2005–2018 papers

Fig. 6. Comparative study for co-occurrence of keywords.

REFERENCES

[1] SCImago, "SJR — SCImago Journal & Country Rank.." Portal, July 2020.

[2] J. M. Merigó, A. Mas-Tur, N. Roig-Tierno, and D. Ribeiro-Soriano, "A bibliometric overview of the journal of business research between 1973 and 2014," *Journal of Business Research*, vol. 68, pp. 2645–2653, dec 2015.

[3] W. Iqbal, J. Qadir, G. Tyson, A. N. Mian, S. ul Hassan, and J. Crowcroft, "A bibliometric analysis of publications in computer networking research," *Scientometrics*, vol. 119, pp. 1121–1155, mar 2019.

[4] G. Zurita, A. K. Shukla, J. A. Pino, J. M. Merigó, V. Lobos-Ossandón, and P. K. Muhuri, "A bibliometric overview of the journal of network and computer applications between 1997 and 2019," *Journal of Network and Computer Applications*, vol. 165, p. 102695, sep 2020.

[5] H. Clausen and I. Wormell, "A bibliometric analysis of IOLIM conferences 1977-1999," *Journal of Information Science*, vol. 27, pp. 157–169, jun 2001.

[6] A. K. Shukla, P. K. Muhuri, and A. Abraham, "A bibliometric analysis

[12] T. Luukkonen, R. J. W. Tijssen, O. Persson, and G. Sivertsen, "The measurement of international scientific collaboration," *Scientometrics*, vol. 28, pp. 15–36, sep 1993.

and cutting-edge overview on fuzzy techniques in big data," *Engineering Applications of Artificial Intelligence*, vol. 92, p. 103625, jun 2020.

[7] N. J. van Eck and L. Waltman, "Visualizing bibliometric networks," in *Measuring Scholarly Impact*, pp. 285–320, Springer International Publishing, 2014.

[8] N. J. van Eck and L. Waltman, "Software survey: VOSviewer, a computer program for bibliometric mapping," *Scientometrics*, vol. 84, pp. 523–538, dec 2009.

[9] K. Subramanyam, "Bibliometric studies of research collaboration: A review," *Journal of Information Science*, vol. 6, pp. 33–38, jan 1983.

[10] S. M. Lawani, *Quality, Collaboration and citations in cancer Research: a bibliometric Study.* PhD thesis, 1981.

[11] I. Ajiferuke, Q. Burell, and J. Tague, "Collaborative coefficient: A single measure of the degree of collaboration in research," *Scientometrics*, vol. 14, no. 5, pp. 421–433, 1988.

[13] H. Small, "Co-citation in the scientific literature: A new measure of the relationship between two documents," *Journal of the American Society for Information Science*, vol. 24, pp. 265–269, jul 1973.

[14] J. MARTYN, "BIBLIOGRAPHIC COUPLING," *Journal of Documentation*, vol. 20, pp. 236–236, apr 1964.