A Retrospective Study on the 20 Years of the ADBIS Conference

Theodora Tsikrika[1] and Yannis Manolopoulos[2]

[1] Centre for Research and Technology Hellas, Thessaloniki, 57001 Greece
`theodora.tsikrika@iti.gr`
[2] Department of Informatics, Aristotle University, Thessaloniki, 54124 Greece
`manolopo@csd.auth.gr`

Abstract. The East-European Conference in Advances in Data Bases & Information Systems (ADBIS) spans 20 years of life. Here, by using simple statistic measures and bibliographic analysis, we illustrate basic characteristics and features of ADBIS, i.e. the venues, persons and countries involved. Also, its international character, its competitiveness and its status in the community are revealed. Finally, prolific authors and countries are extracted, topics of interest are deduced, collaboration issues and citing performance are illuminated.

1 Introduction

In September 1992, the Moscow ACM SIGMOD Chapter was chartered, chaired by Leonid Kalinichenko. Immediately after its formation, the Chapter founded the Annual Workshops "Advances in Data Bsases & Information Systems" (ADBIS). In May 1993, the ADBIS workshop was organised in Moscow under the subtitle "1st Joint Workshop of the Kiev and Moscow ACM SIGMOD Chapters". During 1994-1996, three ADBIS Workshops were organised in Moscow by the Moscow ACM SIGMOD Chapter under the subtitle "International Workshop". In 1996, after discussions with the ACM SIGMOD (mostly with its then Chair, Won Kim), it was decided to transform ADBIS into an East European forum for the exchange of scientific advancements in databases and information systems between the research communities in Eastern Europe and the rest of the world. Thus, for the next two years ADBIS was organised as an "East European Symposium", whereas since 1999 its status is as an "East European Conference".

This year the ADBIS conference becomes 20 years old since it was established as an East European event in 1997. In the meantime, the ADBIS conference has attracted the international interest of the research community and is being mentioned in several ranking lists and indexed in several digital libraries, such as DBLP[1], Google Scholar, Microsoft Academic Search and so on. In this paper, by using simple measures and bibliographic analysis, we illustrate basic characteristics and features of ADBIS, i.e. the venues, persons and countries involved. Also, its international character, competitiveness and status in the community

---

[1] Database Systems and Logic Programming, http://dblp.uni-trier.de/db/conf/adbis/

Table 1. Members of the ADBIS Steering Committee

Steering Committee 1997		
R. Bercaru (RO)	L. Kalinichenko (RU) *Chair*	B. Novikov (RU)
A. Čaplinskas (LT)	M. Kogalovsky (RU)	A. Stogny (UA)
J. Eiduks (LV)	T. Morzy (PL)	T. Weltzer (SI)
H.-M. Haav (EE)	P. Navrát (SK)	V. Wolfengagen (RU)
	J. Pokorný (CZ)	
Steering Committee 2016		
P. Atzeni (IT)	H. Jaakkola (FI)	M. Nikitchenko (UA)
A. Benczúr (HU)	L. Kalinichenko (RU) *Chair*	J. Pokorný (CZ)
A. Čaplinskas (LT)	M. Kogalovsky (RU)	B. Rachev (BG)
B. Catania (IT)	Y. Manolopoulos (GR) *Vice-chair*	B. Thalheim (DE)
J. Eder (AT)	R. Manthey (DE)	G. Vossen (DE)
T. Härder (DE)	M. Manukyan (AM)	T. Weltzer (SI)
M. Kirikova (LV)	J. Michaeli (IL)	V. Wolfengagen (RU)
H.-M. Haav (EE)	T. Morzy (PL)	R. Wrembel (PL)
M. M. Ivanović (RS)	P. Navrát (SK)	E. Zumpano (IT)
	B. Novikov (RU)	

are revealed. Finally, prolific authors and countries are extracted, topics of interest are deduced, collaboration issues and citing performance are illuminated.

2   Key-persons and Venues

The ADBIS Conference is managed by a Steering Committee, which initially (1997) was comprised by 13 persons from 11 countries and expanded until 2016 to 28 persons from 19 countries (see Table 1). In addition, Table 2 depicts for each year the venue, the main key organisers along with the invited persons and their countries. We notice that General Chairs and PC chairs are affiliated with 21 countries, whereas the invited speakers are affiliated with 21 countries as well. These figures indicate the sturdy international foundations of ADBIS. In the sequel, additional information will be provided to support this argument.

In the early years (until 2002), ADBIS events were organised in cooperation with ACM SIGMOD. ADBIS joined the ACM SIGMOD program "SIGMOD Digital Symposium Collection (DiSC)" aiming at collecting symposium anthologies on CDs. More information about this cooperation can be found in [2].

Since 1998 the ADBIS proceedings are published by Springer in the Lecture Notes in Computer Science (LNCS) series. In addition, a second volume with additional papers is published either by a local academic publisher, or by CEUR[3], or lastly in the AISC (Advances in Intelligent Systems & Computing) and CCIS (Communications in Computer & Information Science) series of Springer. Next, we focus on the main ADBIS proceedings and their bibliometric analysis.

---

[3] Sun SITE Central Europe, http://ceur-ws.org/

Table 2. Venue and key persons during the 20 years of ADBIS

Year-Venue	General Chairs	PC Chairs	Invited speakers
1997-St.Petersburg		R.Manthey (DE) W.Wolfengagen (RU)	H.Garcia-Molina (USA)
1998-Poznan	T.Morzy (PL)	W.Litwin (FR) G.Vossen (DE)	T.Imieliński (USA) C.Mohan (USA)
1999-Maribor	I.Rozman (SI)	J.Eder (AT) T.Welzer (SI)	S.Alagic (USA) E.Neuhold (DE) G.Pernul (DE)
2000-Prague	J.Pokorný (CZ)	Y. Masunaga (JP) J.Štuller (CZ) B. Thalheim (DE)	S.Nishio (JP) S.Spaccapietra (CH) H.Schweppe (DE)
2001-Vilnius	A.Zavadskas (LT)	J.Eder (AT) A.Čaplinskas (LT)	F.Garzotto (IT) L.Kalinichenko (RU) J.Schmidt (DE)
2002-Bratislava	L.Molnár (SK)	Y.Manolopoulos (GR) P.Navrát (SK)	P.Atzeni (IT) O.Guenther (DE) H.-J.Schek (CH)
2003-Dresden	B.Thalheim (DE) U.Wloka (DE)	L.Kalinichenko (RU) R.Manthey (DE)	F.Bry (DE) G.Bussler (DE)
2004-Budapest	A. Benczúr (HU)	J. Demetrovics (HU) G. Gottlob (AT)	J.Gray (USA) P.Revesz (USA)
2005-Talinn	J.Penjam (EE) A.Kajla (EE)	J.Eder (AT) H.-M. Haav (EE)	T.Härder (DE) N.Guarino (IT)
2006-Thessaloniki	Y.Manolopoulos (GR)	J.Pokorný (CZ) T.Sellis (GR)	S.Abiteboul (FR) Y.Ioannidis (GR) P.Zezula (CZ)
2007-Varna	B.Rachev (BG)	Y.Ioannidis (GR) B.Novikov (RU)	P.Atzeni (IT) T.Sellis (GR) G.Weikum (DE)
2008-Pori	H.Jaakkola (FI)	P.Atzeni (IT) A.Čaplinskas (LT)	H.Mannila (FI) Y.Mattias (IL) T.Öszu (USA)
2009-Riga	J.Grundspeņķis (LV)	T.Morzy (PL) G.Vossen (DE)	M.Brantner (DE)
2010-Novi Sad	M. Ivanović (RS)	B.Catania (IT) B.Thalheim (DE)	G.Antoniou (GR) W.Cellary (PL) S. Krčo (RS) S.Rizzi (IT)
2011-Vienna	A. Min Tjoa (AT)	J.Eder (AT) M.Bielikova (SK)	M.Dumas (EE) G.Gottlob (UK) M.Henzinger (AT)
2012-Poznan	T.Morzy (PL)	T.Härder (DE) R.Wrembel (PL)	Y.Ioannidis (GR) M.Middelfart (USA)
2013-Genoa	B.Catania (IT)	J. Pokorný (CZ) G. Guerrini (IT)	A.Ailamamki (CH) M.Theobald (DE) N.Stantic (AU)
2014-Ohrid	M. Kon-Popovska (FYROM)	Y. Manolopoulos (GR) G. Trajevski (USA)	J.Gamma (PT) M.Garofalakis (GR) M. de Rijke (NL)
2015-Poitiers	L.Bellatreche (FR)	P.Valduriez (FR) T.Morzy (PL)	S.Abiteboul (FR) J.Dittrich (DE)
2016-Prague	J. Pokorný (CZ)	M. Ivanović (RS) B. Thalheim (DE)	A.Gal (IL) E.Rahm (DE) P.Zezula (CZ)

3 Bibliometric Analysis Method

Bibliometric studies provide a quantitative and qualitative indication of the (scholarly) impact of research activities by analysing their associated publications and citations. In the field of Database & Information Systems, past investigations have analysed the publication behaviour [3] and citation frequencies [9] in major venues (journals and conferences) in the Database community and have reported on the impact of benchmarking activities in Information Retrieval, such as TRECVid[4] [10], CLEF[5] [11] and ImageCLEF[6] [12]. For ADBIS, a study with statistical findings about its publications and citations was conducted for the period 1994-2006 [7] to mark the first 10 years under its current status.

Bibliometric studies typically follow three steps: (i) publication data collection, (ii) citation data collection, and (iii) data analysis. Regarding the publication data, the complete lists of ADBIS proceedings can be obtained from bibliographic data sources, such as DBLP. To this end, the DBLP XML dump [6] created on May 6, 2016 was downloaded[7] and processed using XQuery queries to extract the DBLP records of the publications in the ADBIS proceedings. As author names may appear under several variations (e.g. Apostolos Papadopoulos vs. Apostolos N. Papadopoulos), the DBLP person records [6] were employed to automatically map all name variants to the "primary name" of each author, as this is identified by DBLP. Then, authors' names were manually examined to address all cases not currently identified by DBLP. In particular, authors with the same last name were examined to ensure that no duplicates existed due to variations in their first name (e.g. V. Kumar vs. Vijay Kumar).

Regarding the citation data, the most comprehensive citation data sources are: (i) Web of Science[8], (ii) Scopus[9], and (iii) Google Scholar[10]. Each follows a different data collection policy that affects both the publications covered and the number of citations found. Differences in their coverage may affect the assessment of scholarly impact metrics; the degree to which this happens varies among disciplines [1]. For computer science, where publications in peer-reviewed conference proceedings are highly valued and cited in their own right, the Web of Science greatly underestimates the number of citations found [1, 9], given that its coverage of conference proceedings is very partial. Scopus and Google Scholar offer broader coverage and were both employed in this work.

The citations were obtained as follows in a 24-hour period in May 2016. In Scopus, the query "CONFNAME (adbis) AND SRCTITLE (lecture notes in computer science)" was entered in the Advanced Search and the results were cross-checked against the DBLP publication lists. The publications that could not be retrieved in response to the aforementioned query were obtained by using

---

the query "SRCTITLE (lecture notes in computer science) AND VOLUME (*volume_number*)" again in the Advanced Search, where *volume_number* is the volume of the corresponding LNCS proceedings. The Google Scholar citation data collection was performed through *Publish or Perish (PoP)*[11]. In PoP, the "Advances in Databases and Information Systems" and "ADBIS" queries were both used in the Publication field, while the Year of publication was set between 1997 and 2016. In addition, the query "Current Issues in Databases and Information Systems" was also used in the Publication field for the particular Year 2000. The results were manually refined by removing duplicate entries and false positive matches and by merging entries deemed equivalent.

An initial examination revealed that Scopus does not cover the 1997 ADBIS proceedings, probably because these were not published by LNCS. Moreover, the number of citations varies greatly between Scopus and Google Scholar, with the latter finding around ten times more citations than Scopus (similarly to our earlier studies [11, 12]). As a result, this study employs Google Scholar (and in particular its PoP wrapper) as a citation data source. Scopus is employed in a different capacity, as a supplementary publication data source. More specifically, it is used for obtaining the affiliation data of the authors of the ADBIS publications since such affiliation data are very incomplete in DBLP (in our case they are available for around 20% of the ADBIS authors), whereas Scopus offers complete listings since it obtains them directly from the LNCS proceedings.

Finally, the analysis was performed using appropriate XQuery queries and R scripts along several axes, such as publications and citations, so as to identify trends over time and provide insights into the social network constructed by co-authorship relations and into the topics examined in ADBIS publications.

4   Results of the Bibliometric Analysis

4.1   ADBIS Publications

Table 3 focusses on the main proceedings of ADBIS 1997-2016 and provides several interesting measures. The number of submitted papers fluctuates over the years, ranging from 66 in 2008 (Pori) to a record 165 in 2010 (Novi Sad). An average of 30 papers per year has been accepted in the main ADBIS proceedings over these 20 years, with the main outliers being the 57 and 50 papers accepted in 1997 (St Petersburg) and 2010 (Novi Sad), respectively, while the absolute minimum of 21 accepted papers is observed both in the 2008 (Pori) and 2016 (Prague) ADBIS conferences.

The acceptance rate, i.e. the number of accepted papers over the number of submitted papers, ranges from 18.8% in 2005, (Tallinn) where the second highest number of submissions (144) was observed, to 33.7% in 2003 (Dresden). As the acceptance rate is considered to be correlated with the quality of the accepted material and thus it is often used as an indicator of the prestige of a conference, the 27.1% average acceptance rate indicates that ADBIS is competitive.

---

[11] http://www.harzing.com/pop.htm

Year-Venue	Submitted papers	Submitting countries	Accepted papers	Accepted countries	Paper % acceptance	Country % acceptance
1997-St Petersburg			57	25		
1998-Poznan	90		25	18	27.8	
1999-Maribor	94	33	25	17	26.6	51.5
2000-Prague	115	22	32	17	27.8	77.3
2001-Vilnius	82	30	25	20	30.5	66.7
2002-Bratislava	115	35	29	18	25.2	51.4
2003-Dresden	86		29	20	33.7	
2004-Budapest	130		27	18	20.8	
2005-Tallinn	144	40	27	17	18.8	42.5
2006-Thessaloniki	126	36	29	17	23.0	47.2
2007-Varna	77	29	23	14	29.9	48.3
2008-Pori	66	21	21	17	31.8	81.0
2009-Riga	93	28	25	15	26.9	53.6
2010-Novi Sad	165	38	50	26	30.3	68.4
2011-Vienna	105	31	30	16	28.6	51.6
2012-Poznan	122	31	32	20	26.2	64.5
2013-Genoa	92	43	26	12	28.3	27.9
2014-Ohrid	82	33	26	18	31.7	54.5
2015-Poitiers	135	39	31	23	23.0	59.0
2016-Prague	84	35	21	24	25.0	68.6
Average	105	33	30	18	27.1	57.1

**Table 3.** ADBIS main proceedings (1997-2016)

Again, we remark that there is steady international community paying attention to ADBIS; each year the submitted papers originate from 33 countries on average, whereas the accepted papers originate from 18 countries. In the remainder of this study, we focus our analysis on the 1997-2015 proceedings, since the 2016 proceedings were unavailable at the time of writing.

4.2 Authors in the ADBIS Proceedings

The 578[12] papers in the 1997-2015 ADBIS main proceedings were collaboratively produced by 1,065 unique authors. Figure 1 shows the cumulative number of authors, indicating a steady increase in the number of authors joining the ADBIS community each year. Figure 2 further provides the number of authors per year and their distribution across the two categories: the "newcomers" which appear for the first time in the ADBIS proceedings (upper dark part of the bar) and the "recurring" ones which have appeared before (lower light part of the bar).

On average, 53 new authors contribute to the accepted papers each year; the highest and lowest numbers of new authors publishing in the ADBIS proceedings

---

[12] In addition to the 569 papers listed in Table 3 for 1997-2015, we also consider 9 more papers listed in DBLP for 1998: 6 short papers and 3 in the industrial track.
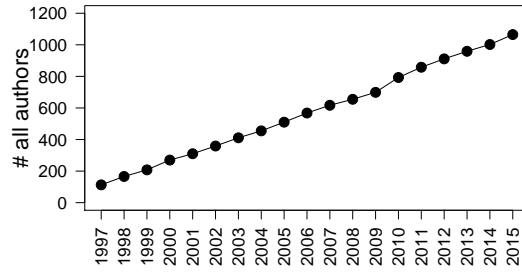
Fig. 1. Cumulative number of authors up to the indicated year
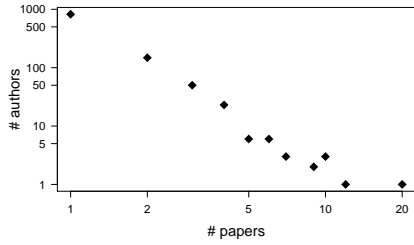

(a) number

(b) percentage
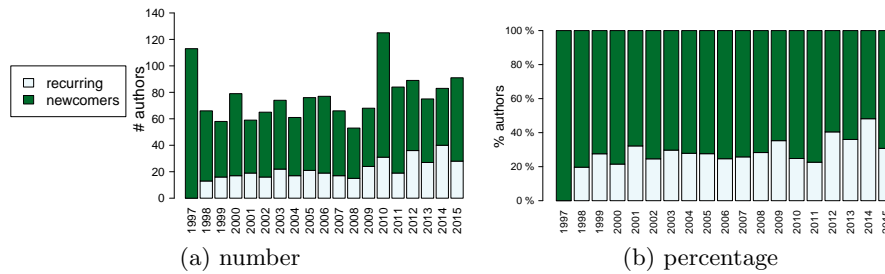
Fig. 2. Authors per year

(94 and 38, respectively) correspond to the conferences with the highest and lowest number of accepted papers, respectively, i.e. in 2010 (Novi Sad)[13] and 2008 (Pori). On average, 71% of authors are newcomers despite that this percentage has dropped lastly as the ADBIS community grows. However, in any case, at least half the authors have not previously published in ADBIS.

Overall, this indicates that ADBIS is an open community that is continuously being infused with new people and thus new ideas and perspectives. This is further corroborated by Figures 3 and 4 that depict the distributions on log-log scales of the number of papers per author and the number of years an author has published in ADBIS. Both follow a power law with the overwhelming majority of authors publishing only once in ADBIS and only about 20% of the authors publishing multiple times. The two distributions are actually quite similar, further indicating that most authors typically publish once within each year. As a matter of fact, the maximum number of papers published within a single year by one author is three (G. Saake in 1998 and Y. Manolopoulos in 2003); 61 authors have published twice in a year, while the majority publishes only once each year.

These results further indicate that ADBIS is an open community that constantly attracts newcomers, while it also maintains its appeal to several people that keep publishing in ADBIS over a number of years. Table 4 lists the top-10 authors in terms of the years they have published in ADBIS and of the number

---

[13] The highest number of accepted papers was actually observed in 1997. Since this was the first year of ADBIS, it is excluded, since all authors are considered as newcomers.

Fig. 3. Distribution of the number of papers per author
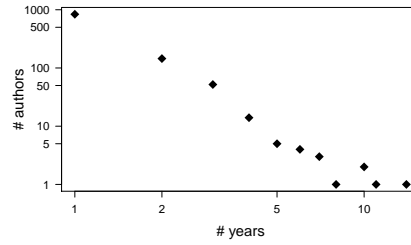
Fig. 4. Distribution of the number of years an author has published

**Table 4.** Most engaged and most prolific authors

# years		# papers		# papers (normalised)	
Y. Manolopoulos	14	Y. Manolopoulos	20	Y. Manolopoulos	6.73
T. Härder	11	T. Härder	12	T. Härder	5.17
L.A. Kalinichenko	10	L.A. Kalinichenko	10	L.A. Kalinichenko	4.92
M. Vassilakopoulos	10	M. Vassilakopoulos	10	K. Nørvåg	4.67
W. Lehner	8	W. Lehner	10	A.V. Zamulin	3.50
A. Corral	7	T. Morzy	9	T. Morzy	3.25
T. Morzy	7	G. Saake	9	M. Vassilakopoulos	3.25
K. Subieta	7	A. Behrend	7	M. Wojciechowski	3.17
A. Behrend, J. Pokorný	6	A. Corral	7	J. Pokorný	3.08
G. Saake, M. Zakrzewicz	6	K. Subieta	7	W. Lehner	3.07

of their papers. As expected the first two columns are highly correlated, with the top-5 authors being the same in both rankings. As publications typically involve multiple authors, and thus may require different amounts of effort, inversely proportional to the number of contributors, the number of papers attributed to each author is normalised by the number of authors in each paper to achieve a fair comparison among publications with varying numbers of authors. This results in a different ranking of the most prolific authors (3rd column) that has five authors in common with the ranking based on raw counts (2nd column). Interestingly, the top-3 authors are the same in both cases. This motivates us to further examine the relations among authors in the ADBIS proceedings.

4.3 Co-authors and their network

An indication that there is a variation in the number of co-authors over the years has already been provided in Figure 2a, where the highest number of authors in a year (125 in 2010 (Novi Sad)) does not correspond to the year with most accepted papers (57 in 1997 (St Petersburg)). Figure 5 shows the distribution of the number of authors per paper for each year, with the width of each "violin plot" reflecting the number of papers for a given number of authors. The line plot indicates the average number of authors which fluctuates between 2.1 and
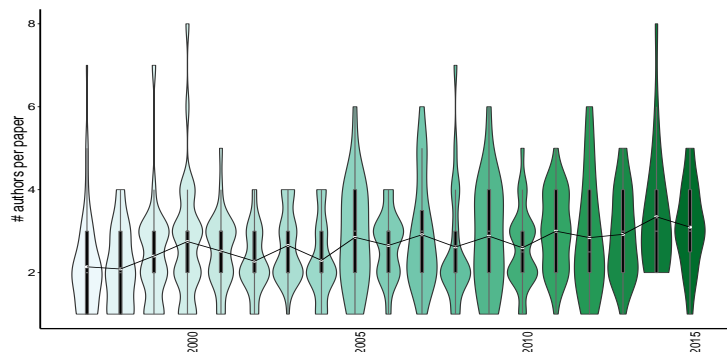
Fig. 5. Distribution of the number of authors per paper per year

3.3 with a slightly upward trend in recent years. This indicates a culture of collaboration, with only though a few papers (5.7%) having five authors or more.

Next, the co-authorship network is constructed to examine the ADBIS author community as a whole. It consists of the 1,065 authors as vertices, with 54 being isolates (i.e. authors who have not collaborated with anyone else) and 1,011 connected with 1,457 edges. The edge density is low (0.002) indicating a loosely connected community. This is further evidenced by the distribution of the size of the 226 components in the network. The largest connected component contains 78 vertices (7.7% of all vertices) and 158 edges, while only 16 components (7.1%) contain 10 vertices or more. Actually, more than half of the components (140 out of 226) contain only two or three vertices, likely to correspond to co-authors of papers who have not collaborated with other members of the ADBIS community.

Next, the key members of the ADBIS co-authorship network are identified based on the following centrality measures. Given an undirected network $G(N, L)$ with N nodes and L links, the *degree* of a node is the number of its neighbours, i.e. the number of co-authors in our case. The *betweenness centrality* [4] of node n_k is based on the number of paths $g_{ij}(n_k)$ from node n_i to node n_j that pass through node n_k, to the number of all paths g_{ij} from node n_i to node n_j, summed over all pairs of nodes and normalised by its maximum value $(N^2 - 3N + 2)/2$:

$$BC_k = \frac{2 \sum_{i<j}^{N} \frac{g_{ij}(n_k)}{g_{ij}}}{N^2 - 3N + 2} \tag{1}$$

Betweenness centrality is, in essence, based on a broker position connecting others. Finally, Google's PageRank (PR) [8], introduced to measure the importance of Web pages, is defined for node n_k as:

$$PR_k = \frac{1-d}{N} + d \sum_{n_i \in \mathcal{N}(n_k)} \frac{PR_i}{L(n_i)} \tag{2}$$

where d is the damping factor (set to 0.85), $L(n_i)$ is the number of links to node n_i and $N(n_k)$ is the set of nodes connected to node n_k, i.e. n_k's neighbourhood.

**Table 5.** Centrality measures for identifying key members of the co-authorship network

Degree		Betweeness		PageRank	
Y. Manolopoulos	26	Y. Manolopoulos	2223	Y. Manolopoulos	0.00553
W. Lehner	18	T. Morzy	1123	T. Härder	0.00427
G. Saake	15	T. Welzer	670	T. Morzy	0.00379
T. Morzy	14	A. Corral	497	W. Lehner	0.00378
C. Traina Jr.	13	Y. Theodoridis	496	G. Saake	0.00336
T. Härder	12	J. Eder	297	K. Subieta	0.00336
A. J. M. Traina	11	A. Vakali	296	P. Návrat	0.00281
A. Behrend, A. Corral,	10	D. Pfoser	292	E. Pitoura	0.00267
E. Pitoura, K. Subieta,	10	B. Brumen	258	J. Pokorný	0.00247
M. Vassilakopoulos, T. Welzer	10	W. Lehner	255	L.A. Kalinichenko	0.00246

Table 5 lists the key members of the co-authorship network based on the three centrality measures. Y. Manolopoulos has the largest number of co-authors (26) in the ADBIS community (W. Lehner follows with 18 co-authors). This number of co-authors resulted in forming a network that corresponds to the largest connected component in the ADBIS community. Also, he has the highest betweenness centrality and PageRank values indicating his mediator role. In addition, the number of co-authors appears to be correlated with the PageRank measure, since seven authors are present in both rankings. On the other hand, the correlation appears to be weaker with the betweenness measure, as only five authors appear in both rankings. Next, the (co-)authorship of the ADBIS papers is examined in terms of the countries where the authors are based.
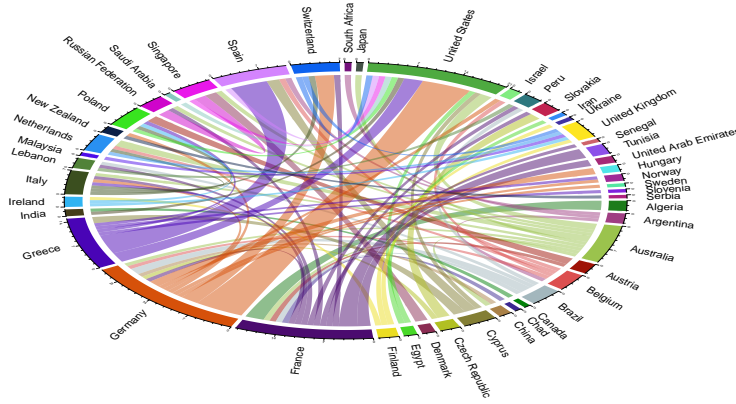
4.4 Countries

The 578 papers in the 1997-2015 ADBIS main proceedings were collaboratively produced by 1,065 unique authors affiliated with organisations in 60 countries. Figure 6 shows the cumulative number of countries, indicating an increase in the number of countries each year. As expected, this increase was more rapid in the early years and has slowed down in recent years; it has not reached a plateau yet, but it is very likely that it will increase only marginally in the future.
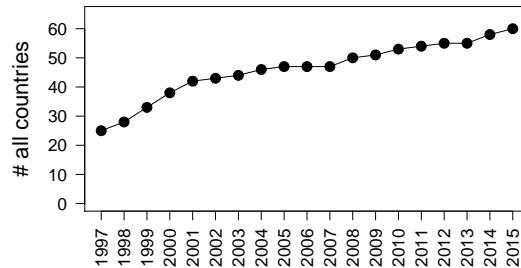
Fig. 6. Cumulative number of countries up to the indicated year

Table 6. Country statistics

# years		# papers		# papers (normalised)	
Germany	19	Germany	122	Germany	111.8
Greece	18	France	57	France	47.9
Poland	18	Greece	50	Greece	42.7
USA	18	Poland	44	Poland	40.8
France	16	USA	42	Russia	34.3
Australia	15	Russia	36	USA	30.9
Russia	15	Italy	28	Italy	25.0
Spain	14	Australia	24	Australia	18.6
Italy	13	Spain	22	Spain	17.7
United Kingdom	11	United Kingdom	15	South Korea	13.0

Fig. 7. Collaboration between authors in different countries in 2015

Table 6 lists the countries with the most solid presence over the years and also in terms of the number of papers with at least one author originating from them. Although ADBIS is supposed to be based in East Europe, it attracts researchers not only from neighbouring countries (i.e. in the rest of Europe), but also from other continents as indicated by the presence of USA and Australia in the top ranks. Similarly to before, the number of papers associated with a country is normalised by taking into account the number of authors originating from that country; e.g. a paper with three authors, two from Greece and one from Spain, counts as one both for Greece and Spain in the case of raw counts (2nd column), but counts as 0.67 and 0.33, respectively, in the normalised case (3rd column). Therefore, the largest the difference between the two columns for the same country, the largest the number of its international collaborations. Overall, the rankings are the same over the top countries, with the exception of South Korea appearing instead of the United Kingdom.

Figure 7 shows the collaborations between the different countries in the AD-BIS community up to 2015. It is encouraging to observe such a significant number of collaborating countries and collaborations. Moreover, the most prolific countries in terms of papers authored (i.e. Germany, France and Greece) are also (together with the USA) the most extrovert ones with the most collaborations.

4.5 Topics

Based on the titles of the 578 papers we have extracted tag clouds with uni-grams (see Figure 8a) and bi-grams (Figure 8b). It is interesting to notice that the keywords depicted span a large range of database topics. Further research is necessary to evaluate the temporal evolution in this respect.
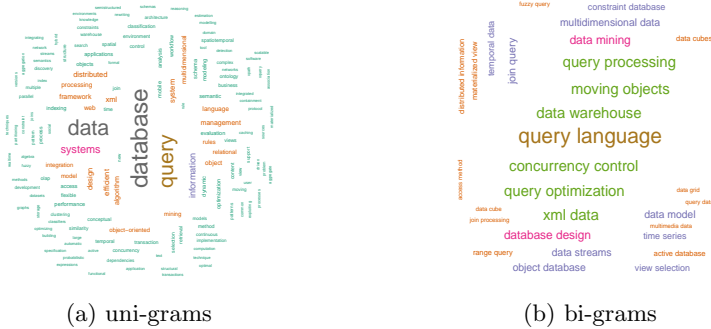

(a) uni-grams   (b) bi-grams

**Fig. 8.** Tag cloud extracted from the titles of the ABDIS papers

4.6 Citations

The 578 papers in the 1997-2015 ADBIS main proceedings have accumulated a total of 4,933 citations, resulting in 8.53 cites per paper on the average. As expected, the citations distribution (Figure 9) is exponential with a long tail, with about 12% of the papers attracting 20 citations or more and 32% of all papers attracting 80% of all citations. The skewness of the citation distribution is also observed by measuring its Gini coefficient, a measure of statistical dispersion that reflects the inequality among values of a frequency distribution. It corresponds to a nonnegative real number $\in [0, 1]$, with higher values indicating more diverse distributions. Its overall value of 0.63 in ADBIS indicates the high degree of variability in the citations of individual publications.

Figure 10 shows the citations distribution for the papers accepted in the main proceedings each year, with the years 2000-2004 attracting the most citations in total and also including some of the most-cited papers (see also Table 7). Among the most cited, there are also papers from 1997 and 2006; as expected,
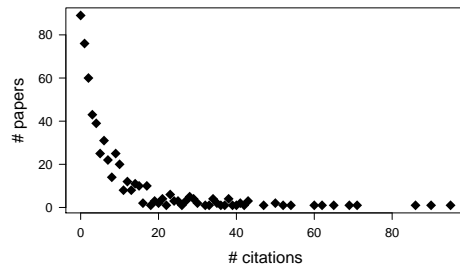
Fig. 9. Distribution of the number of citations per paper
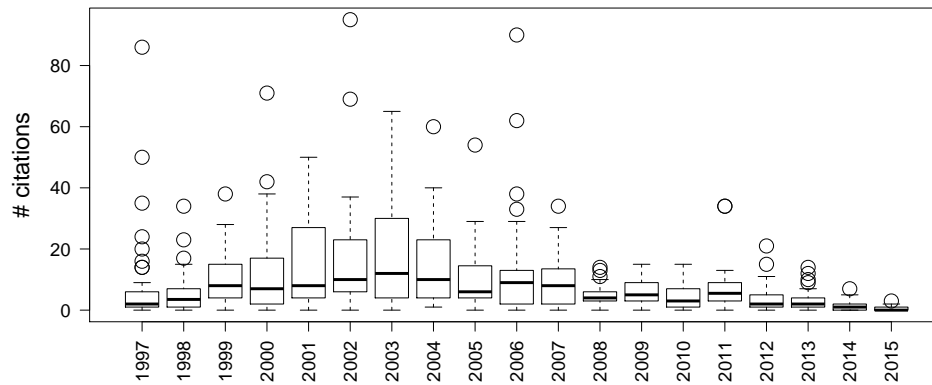
Fig. 10. Distribution of the number of citations over the years

more recent papers have not attracted yet significant numbers of citations. It is interesting to note that the ADBIS invited papers also have significant citations, with the 1997 Garcia-Molina et al. paper having the highest number of citations among all papers published in the ADBIS proceedings (accepted & invited).

5   Conclusions

This paper has narrated how the ADBIS conference was conceived and founded in the early 90s. ADBIS has matured under its present status ("East European") as it is now 20 years old. By using statistic measures and bibliographic analysis, we illustrate basic characteristics and features of ADBIS, i.e. venues, persons and countries involved. Also, its international character, competitiveness and status in the community are revealed. Finally, prolific authors and countries are extracted, topics of interest are deduced, collaboration issues and citations are illuminated. The ADBIS family should further work towards the next 20 years.

6   Acknowledgements

Thanks are due to L. Kalinichenko for information on the early years of ADBIS.

Table 7. Most cited papers in the ADBIS proceedings

Cites	Authors	Title	Year
papers accepted in the main proceedings			
95	Brakatsoulas et al.	Revisiting R-tree construction principles	2002
90	Aouiche et al.	Clustering-based materialized view selection in data warehouses	2006
86	Pozewaunig et al.	ePERT: Extending PERT for workflow management systems	1997
71	Nanopoulos & Manolopoulos	Finding generalized path patterns for web log data mining	2000
69	Akal et al.	OLAP query evaluation in a database cluster: a performance study on intra-query parallelism	2002
invited papers in the main proceedings			
162	Garcia-Molina et al.	Semistructured Data: The TSIMMIS Experience	1997
42	Schek et al.	Infrastructure for information spaces	2002
33	Boyens & Günther	Trust is not enough: Privacy and security in ASP and Web service environments	2002

References

1. J. Bar-Ilan. Which h-index? A comparison of WoS, Scopus and Google Scholar. *Scientometrics*, 74(2):257-271, 2008.
2. S. Ceri, L. Kalinichenko, M. Kitsuregawa, H. Lu, Z.M. Ozsoyoglu, R. Snodgras, and V. Vianu. SIGMOD sister societies. *ACM SIGMOD Record*, 29(1):4-15, 2000.
3. E. Elmacioglu and D. Lee. On six degrees of separation in DBLP-DB and more. *ACM SIGMOD Record*, 34(2):33-40, June 2005.
4. L. C. Freeman. Centrality in social networks conceptual clarification. *Social networks*, 1(3):215-239, 1978.
5. A.-W. Harzing. Citation analysis across disciplines: The impact of different data sources and citation metrics, 2010. Retrieved from http://www.harzing.com/data_metrics_comparison.htm.
6. M. Ley. DBLP: Some lessons learned. *Proceedings of the VLDB Endowment*, 2(2):1493-1500, 2009.
7. Y. Manolopoulos. A statistic study for the ADBIS period 1994-2006. In *Local Proceedings of the 10th East-European Conference on Advances in Databases & Information Systems (ADBIS)*, pages 247-266, 2006.
8. L. Page, S. Brin, R. Motwani, and T. Winograd. The PageRank citation ranking: bringing order to the web. 1999.
9. E. Rahm and A. Thor. Citation analysis of database publications. *ACM SIGMOD Record*, 34:48-53, December 2005.
10. C. V. Thornley, A. C. Johnson, A. F. Smeaton, and H. Lee. The scholarly impact of TRECVid (2003-2009). *JASIST*, 62(4):613-627, 2011.
11. T. Tsikrika, B. Larsen, H. Müller, S. Endrullis, and E. Rahm. The Scholarly Impact of CLEF (2000-2009). In *Proceedings of the 4th CLEF conference*, pages 1-12, 2013.
12. T. Tsikrika, A. G. Seco de Herrera, and H. Müller. Assessing the scholarly impact of ImageCLEF. In *Proceedings of the 2nd CLEF conference*, pages 95-106, 2011.