Database Publication Practices

Philip A. Bernstein, Microsoft Research
David DeWitt, Univ. Wisconsin
Andreas Heuer, Univ. Rostock
Zachary Ives, Univ. Pennsylvania
Christian S. Jensen, Aalborg Univ.
Holger Meyer, Univ. Rostock
M. Tamer Özsu, Univ. Waterloo
Richard T. Snodgrass, Univ. Arizona
Kyu-Young Whang, KAIST
Jennifer Widom, Stanford Univ.

Abstract

There has been a growing interest in improving the publication processes for database research papers. This panel reports on recent changes in those processes and presents an initial cut at historical data for the *VLDB Journal* and *ACM Transactions on Database Systems*.

1. Introduction

The database field is growing. There are more researchers, and hence more papers. The number of research papers submitted to the VLDB Conference has grown by 50% over the past 10 years (see Figure 1) [2]. SIGMOD has had similar growth, from 240 submissions in 2002, to 431 in 2004 (see Figure 1 in [1]). *VLDB Journal* (abbr. *VLDB J.*) and *ACM Transactions on Database Systems* (*TODS*) also have seen large increases.

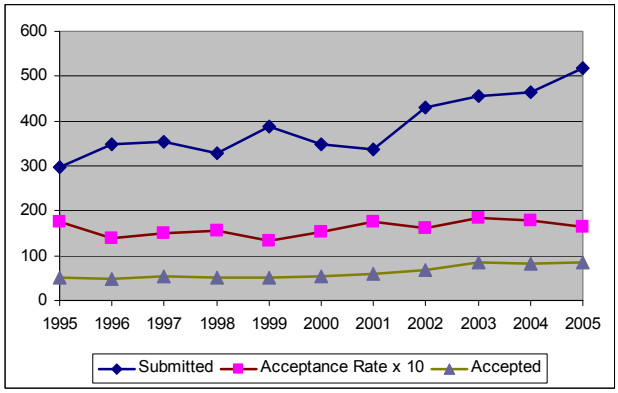

Figure 1 Research Papers in VLDB Conferences

Growth has led to problems: the management challenge of larger program committees (PCs), and in some cases larger reviewing loads. The increased breadth of the field and depth of many areas have made it difficult to obtain expert reviews. The long-standing problems of poor reviews and perceived randomness of PC decisions have become a greater irritation, as conferences have become the preferred prestige publication outlet.

Due to these issues, there has been ongoing discussion about ways to improve the publication process, by the VLDB Board of Trustees, a SIGMOD task force, the SIGMOD Executive Committee, the SIGMOD Advisory Board, journal editors, and a plenary panel at SIGMOD 2004. This panel will report on some of those discussions and on attempted solutions. Panelists may offer their own opinions about problems and solutions, and the audience will be encouraged to do the same. The panelists are Phil Bernstein (chair), Surajit Chaudhuri, David DeWitt, Zach Ives, Christian S. Jensen, and Kyu-Young Whang.

2. SIGMOD 2004 Panel

During SIGMOD 2004, the issue of how to improve the conference reviewing process was addressed in a panel organized by Mike Franklin and Jennifer Widom [1,3]. The core issue was that the current review process was not scaling. The sizes of PCs were increasing to accommodate the greater number of papers, and there was a general perception that the quality of reviews was dropping, to the point that good papers were being "lost in the noise."

Panel members suggested several issues that should be addressed. One was the lack of incentives to force paper authors to be discriminating in their submissions: there is no penalty to submitting papers that are clearly below threshold. A second was the belief that many papers rejected from one conference are quickly resubmitted to the next, often with insufficient effort to improve them. So many papers go through several reviewing cycles with different reviewers. Finally, the prestige of serving on a PC has diminished, so there are fewer incentives for reviewers to invest a great deal of energy in the process.

Several mechanisms were proposed to encourage authors to be more judicious in submitting papers; however, it was unclear that any of these could be implemented fairly. Another suggestion was to break the pipeline between conferences, by moving deadlines so

Permission to copy without fee all or part of this material is granted provided that the copies are not made or distributed for direct commercial advantage, the VLDB copyright notice and the title of the publication and its date appear, and notice is given that copying is by permission of the Very Large Data Base Endowment. To copy otherwise, or to republish, requires a fee and/or special permission from the Endowment

Proceedings of the 31st VLDB Conference, Trondheim, Norway, 2005

that rejected papers could not be immediately resubmitted. This would encourage authors to submit more polished work and choose the most appropriate forum. A third class of suggestions focused on improving the review process itself: ideas included open reviewing via the Web, mandatory physical PC meetings, reviewer evaluations, two-tier PCs, and more reviewers per paper.

One promising class of suggestions involved blurring the distinction between conferences and journals. Conferences might be given "institutional memory," so that resubmitted papers could be sent to the first-round reviewers, like journals. Alternatively, journals could be made more prestigious and desirable, with faster review times and best paper awards. Ultimately, the two types of forum might be merged, allowing paper submissions to the journal throughout the year, and choosing papers accepted for the journal to be presented at the conferences. This would avoid having the size of conference presentation programs limit the number of accepted papers. It could distribute reviewing load more evenly throughout the year. And more importantly, it might raise the standards for referees to check correctness and completeness of papers to a level that is commensurate with the increased prestige of top conference publications.

3. SIGMOD 2005 Experience

As PC chair, Jennifer Widom experimented with several changes to the PC process for SIGMOD 2005.

- The PC Meeting was mandatory. Everyone came!
- The PC was partitioned into 9 groups, roughly by area. Some cross-group reviewing was needed to fill in expertise. At the PC meeting, groups met independently to decide on most papers. Later, group leaders met to decide the borderline papers. Most PC members liked the groups, since they enabled more discussion.
- Authors could see reviews a week before the PC meeting and reply with written feedback. This affected some decisions. Many authors felt better in being able to talk back to the referees. It made reviewers more accountable, so perhaps more careful in their criticism.
- There was no quota on how many papers to accept. But the acceptance rate turned out to be the usual 15%.

Other processing changes included: a smaller PC and hence a bigger reviewing load; no abstract submission; a more detailed review form; monitoring of reviews, and a thank-you gift for PC members.

4. VLDB-SIGMOD Resubmission Process

The VLDB Trustees and SIGMOD Executive have agreed to the following process change: Some rejected papers from VLDB (resp. SIGMOD) may be revised and submitted to the next SIGMOD (resp. VLDB) *to the same referees* and a new referee. This process begins now, with about 10 papers rejected from VLDB 05 being offered this option when resubmitting to SIGMOD 06. A resubmission must be accompanied by an explanation of how criticisms by the first-round reviewers were handled. This process is viewed as a step toward evolving the two conference proceedings into an on-line journal.

5. Journal Statistics

5.1 Introduction

The editors of *TODS* and *VLDB J.* have collaborated to generate statistics based on common definitions. The data reported here is a first cut at these statistics. They are currently being analyzed and are therefore likely to change to fix inconsistencies and improve clarity. A detailed explanation of the statistics will be posted at the VLDB archive [2].

The following is a high-level summary of the statistics. There are 9 graphs. Graphs 1–4 and 8 report max and average, and min when meaningful numbers are available.

In graphs 1 and 2, a "round" is one of the following. A manuscript (original submission or revision) is submitted and an editorial decision is made (i.e., accepted or rejected or a revision requested) without going to reviewers. Or a manuscript is submitted and an editorial decision is made after gathering one or more external reviews.

1. First-round turnaround time: The time for the first round, measured from the manuscript's submission date to the journal to the date that an editorial decision is sent to the author(s). The X axis is the year of the submission.
2. Overall turnaround time: Same as First-round turnaround, but measured for *all rounds* that were initiated in a given year (i.e., for both original submissions and revisions). The X axis is the year the round was initiated..
3. Acceptance Time: The difference between the date of the accept decision and the date of initial submission. The X axis is the year of the initial submission.
4. End-to-end time: The difference between the date of the issue and the date of the initial submission. The X axis is the year of publication.
5. Submission rate: The Y axis is absolute number of submissions in each year.
6. Acceptance rate: The percentage of those manuscripts submitted that year that were ultimately accepted.

In graphs 7–9, the X axis is the volume.

7. Number of articles per volume.
8. Article length per volume: The number of formatted pages of an article, including bibliography and printed appendices.
9. Total page length per volume: The sum of the lengths of the articles of that volume/issue.

5.2 ACM Transactions on Database Systems

Graphs 1 and 2 are in Figure 2, graph 3 is in Figure 3, and graph 4 is in Figure 4. Some of these graphs are shown as tables due to lack of detailed data. For papers for which only the submission or acceptance *month* was known (all papers prior to 2001 and eight papers after 2001), the first day of the month was assumed. The data points in Figure 4 do not include about 23 papers during 1976–1998, about one a year, for which detailed data is not known.

Graphs 5 and 6 are in Figure 5. The acceptance rate for 2003 does not include one submission still in review (as a revision)—the acceptance rate for 2004 is not given, as there are many submissions still in review (all as revisions). Graphs 7–9 are in Figures 6–8. There was one volume per year (four issues, March, June, September, and December).

	First Round		Overall	
Year	Avg	Max	Avg	Max
2002	3.4	7.1	3.4	7.1
2003	2.9	6.5	2.9	6.5
2004	3.0	5.0	2.9	5.0

Figure 2 *TODS* Turnaround Time in Months

Year Initially Submitted	Min	Avg	Max
2002	5.5	12.8	31.7
2003	4.0	10.1	19.5

Figure 3 *TODS* Acceptance Time in Months

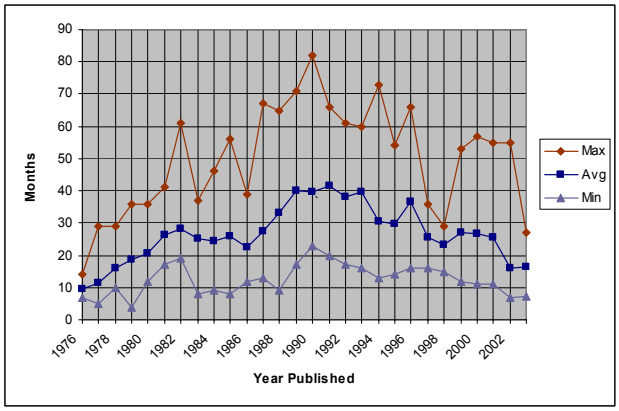

Figure 4 *TODS* End-to-End Time

Year	Number Submitted	Acceptance Rate
2002	60	38%
2003	72	35%
2004	79	---

Figure 5 *TODS* Submission and Acceptance Rate

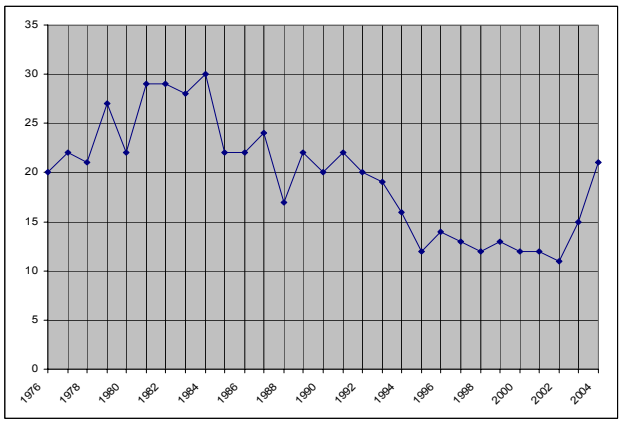

Figure 6 *TODS* Number of Articles per Volume

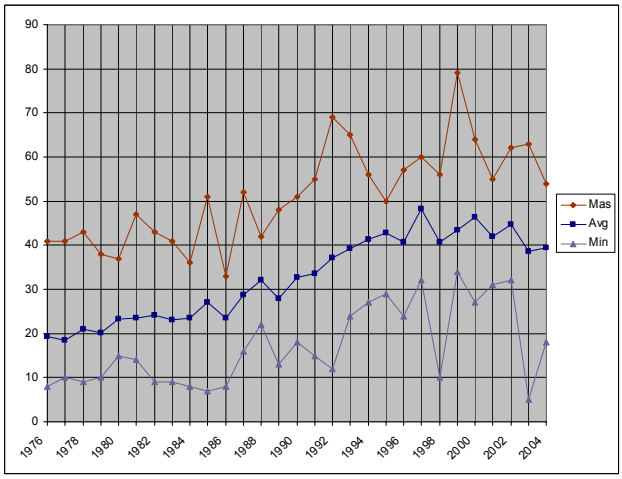

Figure 7 *TODS* Article Length per Volume

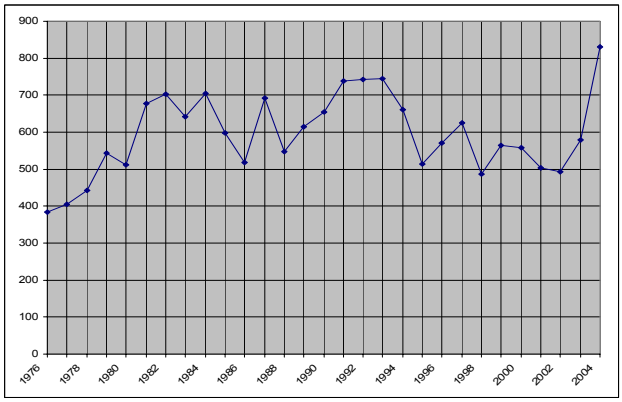

Figure 8 *TODS* Total Pages per Volume

5.3 VLDB Journal

Graphs 1–9 are in Figures 9–17 respectively.

- Turnaround times before 1995 are very incomplete and hence excluded from Figure 9–Figure 12.
- Figure 9–Figure 11 and Figure 14 exclude the year 2004 because many papers submitted in that year are still in review.
- The number of papers still in process, by year, are: 1999 (1), 2000 (6), 2001 (4), 2002 (6), 2003 (5).
- In 1996, *VLDB J.* moved from Boxwood Press (roughly *TODS* page format) to Springer-Verlag (in a larger format). We estimate the latter's page size as 1.86 times the *TODS* page size. Figure 16 and Figure 17 show curves normalized to the TODS format based on that factor.
- In Figure 15–Figure 17, years do not map exactly to volumes, e.g., for 1999 and 2000, when final issues of a volume were published late.

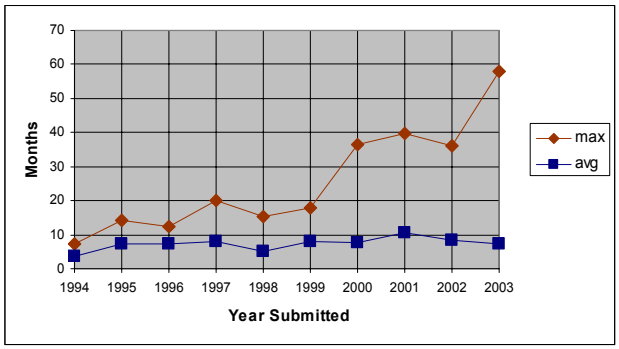

Figure 9 *VLDB J.* First Round Turnaround Time

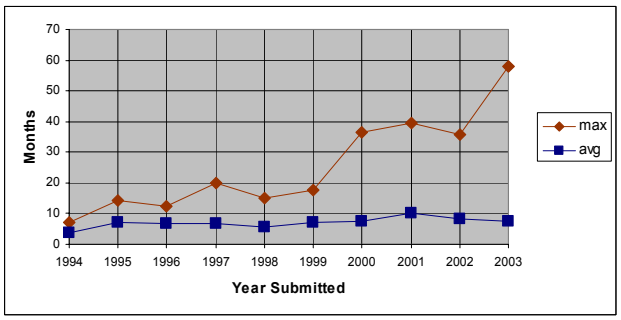

Figure 10 *VLDB J.* Overall Turnaround Time

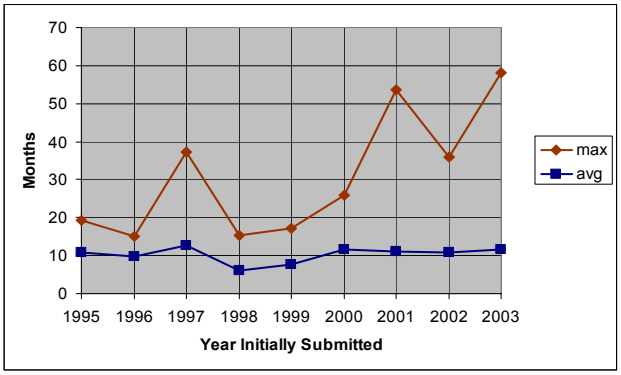

Figure 11 *VLDB J.* Acceptance Time

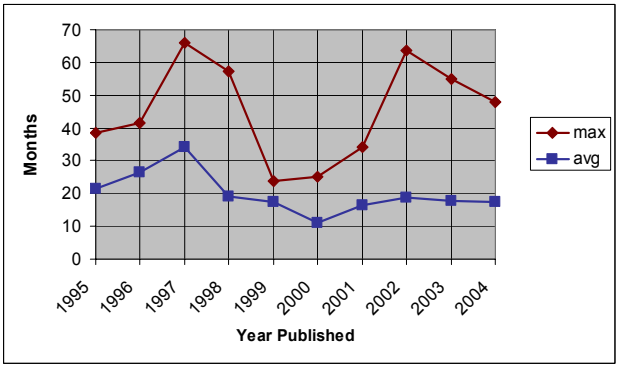

Figure 12 *VLDB J.* End-to-End Time

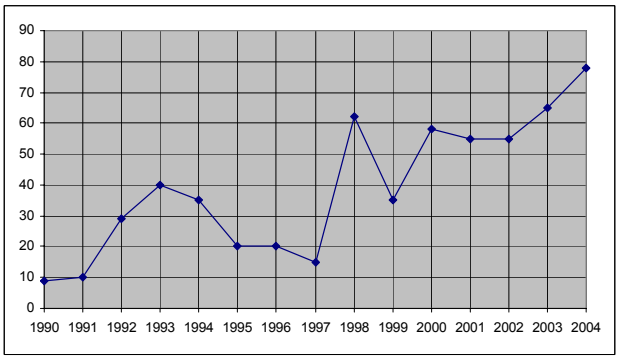

Figure 13 *VLDB J.* Number of Submissions

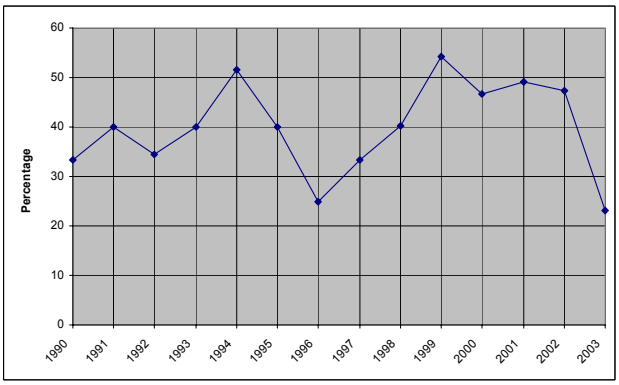

Figure 14 *VLDB J.* Acceptance Rate

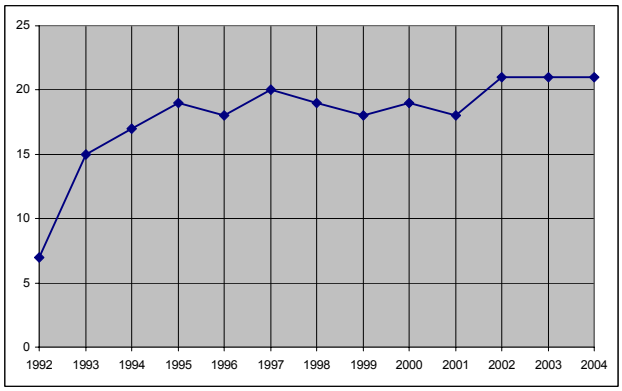

Figure 15 *VLDB J.* Number of Articles per Year

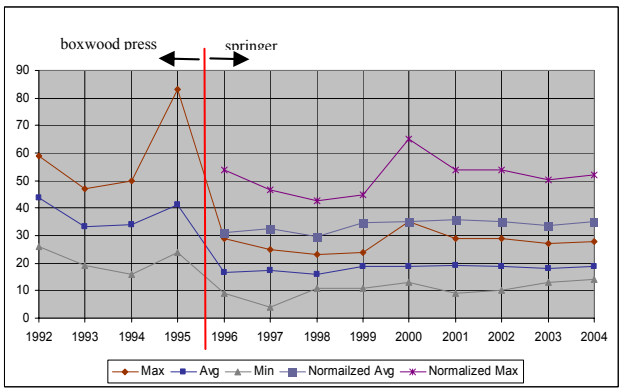

Figure 16 *VLDB J.* Article Length per Year

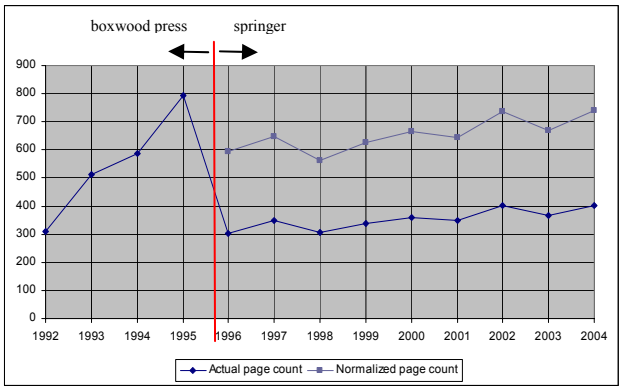

Figure 17 *VLDB J.* Total Pages per Year

References

1. M. J. Franklin, J. Widom, G. Weikum, P. A. Bernstein, A. Y. Halevy, D. J. DeWitt, A. Ailamaki, Z. G. Ives: Rethinking the Conference Reviewing Process - Panel. *SIGMOD 2004*: 957.
2. VLDB Archives. http://www-nishio.ist.osaka-u.ac.jp/vldb/archives/public/. See also www.vldb.org.
3. G. Weikum, P. Valduriez: Report on the 2004 SIGMOD Conference. *SIGMOD Record 33, 4*: 8-10 (2004).