

SIGMOD 2018 Program Committee Chair's Report

Philip A. Bernstein
Microsoft Research
philbe@microsoft.com

ABSTRACT

This paper reports on the program committee process for SIGMOD 2018, including statistics, trends, and changes from previous years. Some highlights are:

- Submissions to SIGMOD 2018 were down 6% from 2017. The acceptance rate of research papers was 20%, which is in line with recent years.
- Reviewers showed a strong bias to reject borderline papers rather than giving authors the opportunity to revise. By being biased in the second round in favor of offering authors the opportunity to submit a revision, we increased the acceptance rate significantly from the first to second round. We strongly recommend that future PC chairs adopt this bias.
- To help ensure high quality reviews, we gave PC members a light reviewing load and ensured that 95% of review assignments went to PC members who bid Eager or Willing to review the paper. Nevertheless, approximately 20-25% of reviews are unacceptably shallow. We need to do better.
- The main changes in 2018 were (i) to reduce the number of parallel sessions, (ii) include tutorials during the main conference (Tuesday – Thursday), and (iii) return to clustering industry presentations into separate industry sessions rather than grouping them with research papers on the same topic. To enable (i) and (ii), we shortened the standard presentation time to 20 minutes and offered only 10 minutes to 40% of the papers. Anecdotal evidence is that attendees were happy with these changes.

The paper closes with comments about my previous experiences as PC chair for SIGMOD 1979 and VLDB 2002 and with the evolution of PC processes.

1. SUBMISSIONS

Submission statistics are summarized in Table 1. The acceptance rate for research papers has been constant for the past few years at 20%. However, the absolute number has been declining. There was a big jump in the research submissions in 2016, presumably due in part to the attractive location of the conference, San Francisco. Since then, the number has been declining, but is still 11% higher than 2015.

For the 2018 industry track, we reverted from 2017's invitation-only approach back to the tradition of evaluating unsolicited submissions. The acceptance rate was 38%, and there were two invited papers. For the demonstration track, the acceptance rate was 33%. Seven of

Table 1 Submission-Acceptance Statistics

		2015	2016	2017	2018
Research	Submitted	413	569	489	458
	Accepted	106 (25%)	116 (20%)	96 (19%)	90 (20%)
Industry	Submitted	18	50	0	40
	Accepted	18	21	0	17
	Invited	0	4	4	2
Demo	Submitted	86	126	90	108
	Accepted	30	31	31	36
Tutorial	Submitted	11	24	16	14
	Accepted	4	10	13	6

14 tutorial submissions were accepted, and two of the seven were asked to merge into a single 3-hour tutorial.

Figure 1 shows the number of research submissions with a given number of coauthors. Only 1 of 21 papers with one author was accepted. Most submissions had 2-6 coauthors, with acceptance rates of 18% to 22%. The acceptance rate was much higher for papers with 7-9 coauthors and dropped to zero after that.

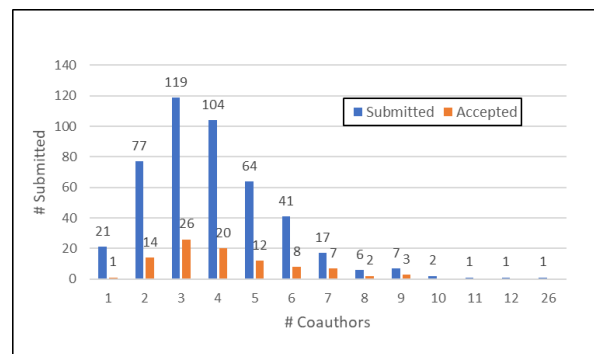


Figure 1 Number of research submissions with a given number of coauthors

As shown in Figure 2, most authors submitted just one research paper. The authors who submitted a lot of papers are professors who were coauthors of many of their students' submissions.

About 20% of submitted abstracts do not result in a submitted paper. In the future, if you are an author of such a paper, please withdraw your paper before or immediately after the submission deadline. This saves work for the PC chair, who has to search for such papers and delete them manually.

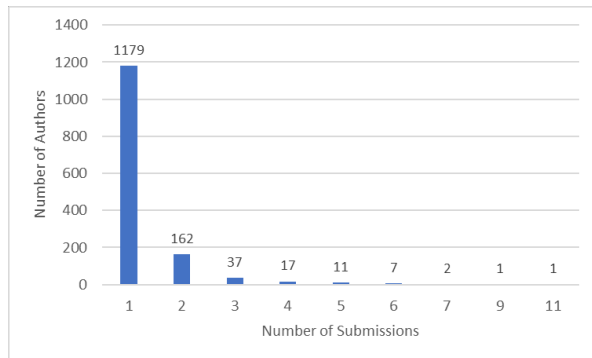


Figure 2 Number of authors with a given number of research submissions

2. REVIEWING PROCESS

The quality of the research program depends heavily on the quality of reviewing. It is therefore important to maximize the expertise of reviewers of each paper.

In each of the two submission rounds, PC members and group leaders read abstracts to guide them to bid for papers they wanted to review: Eager, Willing, In-a-Pinch, or Not-Willing. After running the automatic assignment algorithm, I did manual fix-up to improve the result. Overall, 95% of reviewing assignments went to PC members who bid Eager or Willing. (To be precise: 54% Eager, 41% Willing, 4% In-a-Pinch, 0.4% Not-Willing.) All cases of In-a-Pinch or Not-Willing assignments were papers that had an insufficient number of Eager or Willing bids. I manually chose those reviewers.

I assumed that Eager/Willing reviewers would be very knowledgeable about the paper’s topic. This didn’t always turn out to be true. It can happen because a PC member misjudges the submission’s technical focus based on the abstract. I suspect (but cannot prove) that sometimes a PC member will bid Eager/Willing in order to learn about the topic. Please don’t do this! It’s unfair to the authors and causes extra work for the PC chair, who receives low-confidence reviews and then has to get additional reviews under time pressure.

As in recent years, we used a large PC to ensure a light reviewing load. In each round, most PC members had four papers to review in a 4-week reviewing period, for a total of eight papers for the two rounds. A few had more, either because we had too many submissions in their areas of expertise or because we called on them for a 4th review of a borderline paper (see Figure 3). A few external reviewers reviewed one paper, and a few PC members were recruited for round two to cover topics for which we received more submissions than expected.

Each paper was also assigned to a group leader, who read the paper, guided the discussion after the reviews were in, recommended a decision, and wrote a meta-review. Group leaders also escalated borderline cases to my attention. If reviewers disagreed or they all thought

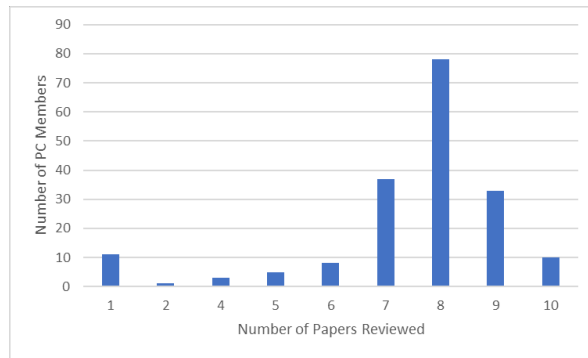


Figure 3 Reviewing Load: Number of PC members with a given reviewing load

the paper was borderline, we pushed them to discuss it on-line. The average number of comments was the same in 2018 and 2017; see Figure 4 where the number of 2017 submissions was multiplied by 458/489 to normalize for the larger number of submissions than in 2018.

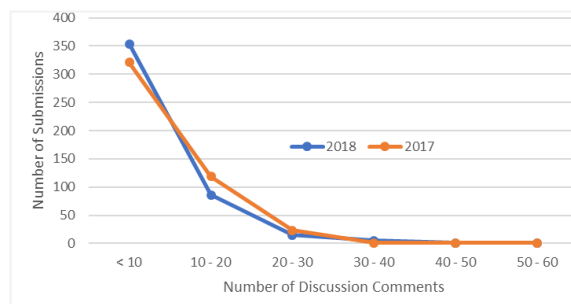


Figure 4 Number of submissions with a given number of discussion comments

3. DECISIONS

Of the 458 research submissions, 20 were desk rejected either because they were very weak or out-of-scope. Of the 210 round-one submissions, we accepted 4 (2%) without revision and asked for a revision of 39 (19%), of which 33 were accepted (85%). I received feedback on the first round from systems researchers that papers they thought were excellent were rejected. The authors speculated this was due to reviews by PC members who weren’t systems researchers and didn’t know how to evaluate such papers. Since 95% of reviews were by PC members who bid Eager or Willing, I was skeptical of this diagnosis. Therefore, I investigated it by re-reading a lot of reviews and came to a different conclusion: in borderline cases, PC members are strongly inclined to reject rather than offering the authors an opportunity to revise. This inclination seemed to be equally true for all topics, not just for systems papers.

In round one, I tried to mitigate this tendency by encouraging group leaders and PC members to ask for more reviews of borderline papers. Nevertheless, in the common case that all three reviewers believed themselves to be relatively knowledgeable about a paper, they pushed

to reach a decision without additional reviews. They took me up on the offer for only 10% of the round-one submissions. I found many of the 4th reviews helpful. In retrospect, I should have requested many more 4th reviews even if the reviewers and group leader did not. I apologize to authors of rejected papers that I might have been able to save with more effort.

Given my experience in round one, in round two I read the abstract and reviews of most papers. In cases where I thought the reviewers were insufficiently generous, I added discussion comments to try to move the consensus toward a revision. For example, I pushed negative reviewers to be specific about the improvements they would like to see for the paper to be acceptable. I also looked for papers where another review or two might help, and sometimes asked for the reviews even if the PC members and group leader thought it was unnecessary. As a result, we requested revisions of 63 (25%) of the 248 submissions, of which 48 (76%) were accepted. With the 5 papers accepted without revision, that gave us an acceptance rate of 21.4% — significantly higher than round one.

Based on this experience, I recommend that we be more generous in giving authors an opportunity to revise their papers. Often, PC members think that the authors can't make the required changes in the one-month revision period we offered. I believe they should not make this judgement. Authors are highly motivated and will go to great lengths to improve a paper under time pressure, if they have a list of specific improvements that are required. Let *them* decide if they have enough time. I still advocate that we give them a list of specific required changes, and not simply ask the authors to do their best with the reviewers' criticisms, as is often done with journal submissions. However, we should not reject a paper that has a potentially good idea just because the list of changes seems too long.

4. REVIEWING QUALITY

I read hundreds of reviews. Most reviewers do a very good job. They think about the ideas presented in the paper, consider whether the ideas have merit, check that the paper justifies its conclusions, compare the paper's contributions to prior work, and write a review that conveys all this to the authors, the other reviewers, and ideally in enough detail that the PC chair (who probably won't read the paper) can understand it. Great reviews aren't always long, but they are always insightful.

Last year's PC Chair, Dan Suciu, introduced a "Distinguished PC Member Award" to recognize reviewers who did a great job, ~10% of the PC. This year, the vice chairs, group leaders, and I continued this new tradition by recognizing the following PC members: Ashraf Aboulnaga, Manos Athanassoulis, Sebastian Breß, Graham Cormode, Sudipto Das, Khuzaima Daudjee,

Aaron Elmore, Ada Fu, Michael Hay, Yuxiong He, Yannis Katsis, Alexandra Meliou, Dan Olteanu, Andrew Pavlo, Peter Pietzuch, Lucian Popa, Semih Salihoglu, Ryan Stutsman, Yufei Tao, & Alexander Thomson.

I regret to report that my impression from reading so many reviews is that 20%-25% of them were unacceptably shallow and sketchy. Previous PC chairs told me this was consistent with their experience. In these cases, the reviewer clearly didn't think hard about the paper, and the review reflects it by offering just a few cheap shots, e.g., the motivation is weak, some sections are hard to understand, more experiments are needed, there are too many typos, and some references are missing. Probably all of that was true, but what did the reviewer think of the ideas? Are they good or bad ideas? Why? What is weak about the justification for the proposed innovation? What did you expect to see that's lacking?

Weak reviews are not a measure of the strength of the reviewer as a researcher. Many were from people with excellent publication records. A few PC members consistently produced weak reviews. Some weak reviews were from reviewers who also wrote very good reviews.

Given that PC members were asked to review only four papers in each round, the problem cannot be the reviewing load. I suspect the following scenario is common: PC members are busy, and they treat reviewing as their lowest priority task. A PC member reads a paper once. Without thinking hard or spending much more time on it, he or she writes a few obvious criticisms, plus a Weak Accept or Weak Reject rating, depending on the PC member's first impression. Task complete.

Unfortunately, writing a good review is time-consuming. To do it, I usually have to read a paper three times. My first reading is to understand the main ideas and get an overall opinion of the work. On the second pass, I start writing the review while I'm reading, commenting on each section as I go. But that writing activity usually makes me question some of my criticisms, which requires a third pass to sharpen my arguments. The hours usually add up to a full working-day, sometimes more, over several days. Despite the effort, I usually find the time well spent because it forces me to think deeply about a topic, even if the submission is quite weak.

With a 170-person PC, it's a statistical certainty that some reviewers will have an unexpected problem with work, family, or health that prevents them from investing enough time to do a competent review. But that does not account for a quarter of the reviews being unacceptably shallow. We need to improve. Which means that some of you reading this article need to improve.

A weak review leaves a lasting negative impression on other reviewers of the paper. It certainly lowered my opinion of the technical depth of some PC members. I'd have thought that PC members who submit superficial

reviews would be embarrassed when they see their colleagues' substantial reviews. Apparently not.

Moreover, weak reviews are hugely unfair to authors. We all know the weeks, or more often months of effort it takes to write a paper, even a weak one. Authors have a right to expect reviewers to spend enough time to give substantial feedback.

I don't know of any silver bullet to fix this problem. Here are some suggestions I've heard that might encourage people to write better reviews:

1. Have each PC chair ask PC chairs of the last few database conferences to review the proposed list of PC members, to identify reviewers they would not recommend. Not all PC chairs do this, and often only for the preceding conference, so it misses some weak reviewers.
2. Ask authors to vote on whether each review of their paper was well done—thumbs up or down. This isn't a vote on whether the author agrees with the review—only on whether the review reflects a serious attempt at understanding the work and thinking about its novelty, importance, and correctness. We should report it as an aggregate, only to PC members who reviewed enough papers that would make the feedback anonymous.
3. Ask PC members to vote on the quality of other reviews of papers they reviewed. Again, report it as an aggregate only to PC members who reviewed enough papers that would make the feedback anonymous, perhaps merging it with the result of (2).
4. PC chairs and group leaders could ask some PC members to strengthen particular reviews. I did this in a few cases where I thought a gentle push would get the desired effect. In general, it is hard because it is a direct criticism of a PC member's work, which everyone involved finds uncomfortable.

5. THE CONFERENCE PROGRAM

Since 2005, SIGMOD's policy has been to accept all papers that passed the PC's quality bar, with no budget for the maximum number of presentation slots. Then the program schedule was adjusted to give all papers a presentation slot. This has led to a conference schedule of 5-6 parallel SIGMOD sessions, plus a PODS session on Tuesday and Wednesday. Table 2 reports the number of parallel sessions, including industry, tutorial, and demo sessions.

As an experiment, SIGMOD 2016 organizers compressed the program into 3 parallel sessions. To fit all papers into the schedule, they gave each paper a 15-minute slot, which meant about 12 minutes of presentation plus 3 minutes of Q&A, and they moved tutorials to Friday after the conference. Instead of separate industry sessions

Table 2 Number of parallel SIGMOD sessions

Year	Tuesday	Wednesday	Thursday
2009	5	6	5
2010	6	6	6
2011	6	6	6
2012	5	5	5
2013	5	5	6
2014	5	5	6
2015	5	5	6
2016	3	3	3 (tutorials on Friday)
2017	5	5	5 (tutorials on Friday)
2018	4	4	4

sessions, they grouped presentations of industry papers with research papers on the same topic. In 2017, the conference reverted to 5-way parallelism and compensated by introducing plenary teaser talks. It kept the tutorials on Friday and again mixed industry and research papers in the same sessions.

I thought the 2016 experiment was largely successful, as did many attendees who responded to the survey distributed after the conference. However, I found the 15-minute slot too short for many papers, and longer than necessary for others. Therefore, with the approval of the 2017 SIGMOD officers, I decided to try the compressed program again with three modifications: (i) different lengths of presentation slots, 20 minutes or 10 minutes, depending on the paper, (ii) tutorials as a 4th parallel session on Tuesday – Thursday, and (iii) industry presentations placed into separate industry sessions.

There were 54 long presentations and 39 short ones. The decision of long vs. short had several phases. During the reviewing process, PC members were asked to recommend whether each paper, if accepted, should be offered a full presentation slot. Then research PC group leaders made a recommendation for each accepted paper they supervised—definitely 20 minutes, 20 minutes if there's time available, no preference, or definitely 10 minutes—based on reviews, reviewer discussions, and their own judgment, without knowing the identity of authors. Their recommendation is not necessarily a quality metric. They recommended “definitely 10” for some papers highly-rated by reviewers, because the topic was narrow, could be explained in 10 minutes, or couldn't be explained in 20 minutes so extra time wouldn't help. All papers rated definitely-20 or 20-if-there's-time were given long slots, plus some of the no-preference ones. For the latter, the final decision was based on many factors, such as topic diversity, institutional diversity, and the time available in the relevant session.

The industry PC chairs, Sam Madden and Neoklis Polyzotis, were given a free hand in choosing industry

submissions and scheduling them in four sessions. Most of the papers were chosen with the help of the Industry PC. There were also two invited papers. I made the mistake of not asking the industry chairs to align the lengths of talks with the 20- and 10-minute boundaries of research talks. This interfered with session hopping, which didn't occur to me until it was too late to change.

Due to recent changes in U.S. immigration policy, some authors have been unable to attend a conference to present their papers. To accommodate these authors, the SIGMOD officers and VLDB Endowment agreed to allow papers that could not be presented in a SIGMOD or VLDB conference to be presented instead at the next such conference. As a result, three research papers and one demonstration paper from VLDB 2017 were presented at SIGMOD 2018.

Authors of ACM TODS papers can present their paper as a poster at the next SIGMOD conference after the paper's publication. One such TODS paper was presented as a SIGMOD 2018 poster.

I considered compressing the program into three parallel sessions: 2 parallel paper sessions, plus 4 industry sessions, 4 tutorials, 2 demo sessions, and a panel. Using the same schedule as 2018, there would be 960 minutes of presentation time for research papers. With 93 presentations, all of them would have only 10 minutes. To offer 20 minutes for some presentations, some papers would get only a 3-minute teaser talk. I seriously considered doing this for 2018. But when I saw that with three parallel papers sessions we could give everyone at least a 10-minute talk, I dropped that plan.

At SIGMOD 2017, in place of traditional keynotes, there were invited plenary talks on hot topics by database researchers. In my opinion, the talks were excellent, and the concept of invited database research talks is worth repeating. However, I would make space by reducing the number of long paper presentations and reserve plenary slots for high-profile invited speakers who would not ordinarily attend SIGMOD.

6. PROCEEDINGS AND BOOKLET

The proceedings and booklet handout were prepared in parallel with the conference program. There is manual effort in preparing both of them, which requires great care to avoid inconsistencies. Many authors make last-minute changes to their paper's title and author affiliations, and to schedule constraints that affect session assignment. Maybe someday this will all be generated automatically from a single database. Until that nirvana arrives, authors should do their best to notify the PC and proceedings chairs as early as possible of such changes. They should also strive to submit their non-technical material on time, e.g., photo and bio for tutorial speakers, to assist in the booklet preparation process.

7. HISTORICAL NOTES

This section summarizes my experiences as PC chair of two earlier major database research conference.

7.1 SIGMOD 1979

The last time I was SIGMOD PC Chair was for SIGMOD 1979. For the amusement of younger readers (nearly everyone, I guess) and to capture a bit of history, let me describe what that activity was like. Unfortunately, I no longer have a written record about the PC process, so my foggy memory will have to suffice.

In those days, the Internet was a research project, not a utility. And the World Wide Web was still about 13 years in the future. Therefore, everything was done via hard copy and the postal service. The latter slowed down the process a lot. The schedule had to leave enough time for coast-to-coast mail delays of 4-5 days.

People on the SIGMOD mailing list received a hard-copy call-for-papers in the mail. The call-for-papers was usually published a year in advance, so it could be distributed at the previous SIGMOD conference.

Except for researchers at a few wealthy labs who had access to fancy printers, most authors prepared their submissions with a text editor and impact printer. (My first access to a laser printer came 5 years later.) If the paper had a lot of fancy math, then it might have required using a typewriter (i.e., no computer). Authors had to mail five photocopies of their submission to the PC Chair (i.e., me), ensuring I would receive it before the submission deadline. There were about 75-80 submissions to SIGMOD 1979, the vast majority of which were from U.S. universities and research labs.

I knew each PC member well enough to assign papers that were within their areas of expertise. After doing the reviewer assignments, I mailed a package of papers to each PC member, with copies of the review form for them to fill out. That left two copies of the paper in my file, one for me and one for an extra reviewer if needed.

During the reviewing period, I read all the submissions. I'm told this wasn't common practice for SIGMOD PC chairs, but I doubt I was the first or last to do so. About a quarter of them were so weak that after a half-hour of reading it was obvious they would be rejected, so I didn't have to dig deeper. I read the rest more carefully, but even so, it was a manageable load for a ten- to twelve-week reviewing period.

PC members mailed (mostly hand-written) reviews to me, to arrive before the reviewing deadline, which was a week before the face-to-face PC meeting, which all PC members attended. I produced a list of all the papers in order of their average review score. Unlike today, we had a quota of how many papers to accept, about 24, to fill a two-and-a-half-day single-track program. We started at the top of the list and accepted papers until we

hit a controversial one. Then we switched to the bottom of the list and rejected papers until we hit a controversial one. Most of the discussion happened on the papers in the middle. We talked about them one-by-one until we converged on a final list. Then it was dinner time, the reward for a day of intense discussions.

7.2 VLDB 2002

I was overall PC chair for VLDB 2002. In response to a request by the VLDB Endowment Board to strengthen the trend of broadening the database field beyond database engines, we had two program committees, one for Core Database Technology (chaired by Raghu Ramakrishnan) and another for higher levels of the stack, called Infrastructure for Information Systems (IIS) (chaired by Yannis Ioannidis). The committees accepted 38 of 209 and 31 of 222, respectively (16% overall).

VLDB continued splitting the research PC into two tracks until 2011. By then, it was agreed that the community sufficiently welcomed IIS papers that they no longer needed a separate PC to obtain a fair hearing.

The PC process was similar to SIGMOD 2018, and very different from SIGMOD 1979. Like today, everything was done on-line: submissions, reviews, and reviewer discussions. If I recall correctly, this was the first VLDB that required submissions to be in camera-ready format, to avoid arguments about whether a submission exceeded the length limit. A big difference from today's conference was that we did not offer authors an opportunity to revise a paper and resubmit it for a second evaluation. That is, every paper was accepted or rejected.

At that time, there were growing complaints that PC decisions were too random. As PC chair, I was on the front-line listening to those complaints. As a result, after the conference, I started lobbying to improve the process by having an on-line journal with the same structure as a PC but including a revision cycle. I was not alone in promoting change. Rick Snodgrass, then Editor-in-Chief of ACM TODS, worked to speed up turnaround time to make TODS as appealing to authors as conferences. In 2003, he and I proposed to the VLDB Endowment Board and SIGMOD Executive Committee that borderline rejected papers from one conference could be revised and resubmitted to the next one with the same reviewers, plus one new reviewer for the receiving conference. This process started in 2005 and ran for a couple of years. We also suggested this evolve into an on-line journal, but it was viewed as too radical and rejected. Over the next several years, I presented versions of that concept at CIDR 2003, in panel sessions on PC processes at SIGMOD 2004 [1] and VLDB 2005 [2], and at annual VLDB Endowment Board meetings in 2003-2005. There were many other proposals, some

presented in [1] and [2] and some discussed privately at SIGMOD and VLDB Board meetings. I was insufficiently persuasive to get either organization to agree to the change. However, after I rolled off the VLDB Endowment Board in 2006, H.V. Jagadish got approval for a related proposal: changing VLDB to an on-line journal, PVLDB, with monthly submissions year-round. I believe the approval was helped by his agreement to serve as its first editor-in-chief, something I was not willing to do. There is widespread agreement that PVLDB has been a big success, which was one of many contributions for which Jagadish received the 2013 SIGMOD Contributions award.

An aside: The VLDB 2002 general chair, Fred Lochovsky, and I pushed for approval to publish the proceedings only in electronic form. The VLDB Endowment Board declined our request and insisted on a printed copy, which ended up as a weighty tome of 1050 pages. Old habits die hard. It took a few more years before hard-copy proceedings were abandoned.

8. FINAL REMARK

Our community has been at the forefront of changes in the PC conference reviewing process for many years. Today's processes are imperfect, and we should continue to strive to improve them. If I could wave a magic wand to get only one improvement, it would be that all PC members invest enough time to give substantial thought to every paper they review and write a detailed evaluation. That would go a long way to increase author satisfaction of the processes that we currently use.

9. ACKNOWLEDGMENTS

Assembling the SIGMOD 2018 program was the work of over 200 people: PC members, group leaders, and track chairs. I'm enormously grateful for their help. I would especially like to thank the two PC vice-chairs, Luna Dong and Mohamed Mokbel, who did a huge amount of work, collaborating closely with me on all aspects of the program. They deserve a lot of credit for the best aspects of the program.

10. REFERENCES

- [1] Michael J. Franklin, Jennifer Widom, Gerhard Weikum, Philip A. Bernstein, Alon Y. Halevy, David J. DeWitt, Anastasia Ailamaki, Zachary G. Ives: "Rethinking the Conference Reviewing Process - Panel." SIGMOD Conference 2004: 957
- [2] Philip A. Bernstein, David J. DeWitt, Andreas Heuer, Zachary G. Ives, Christian S. Jensen, Holger Meyer, M. Tamer Özsu, Richard T. Snodgrass, Kyu-Young Whang, Jennifer Widom: "Database Publication Practices." VLDB 2005: 1241-1246.