Arguments for an Online PC Meeting for ASE 2018

ASE 2018 will make decisions about papers online, instead of the physical PC meeting the conference has used in the past. As PC chairs, we suggested the change for 2018, but it was clearly controversial. The ASE Steering Committee has eventually approved the change for 2018, with clear steps to evaluate the decision for future years. For transparency, with the remainder of the text, we share the key arguments behind this decision, as proposed to the steering committee. -- Gordon and Christian

Context:

Like many other conferences, ASE is facing scalability challenges. The number of submissions is consistently rising (see table), increasing the load on reviewers. ASE's ERP model has reduced the load somewhat, but without a substantial increase of the size of the PC, we are looking at assigning 20 to 30 papers to PC members. To enable a 18 paper load for a 6 week review period with the existing ERP+PC model, we would need 45 PC members and 45 ERP members, larger than any ASE PC before (50% larger than the 2017 PC). Several senior community members warned that physical PC meetings at this scale are very difficult to run, not least because of the need for a very large room and more difficult communication that comes with large distances between seats in such a large room (with experience shared from ICSE meetings with 50 PC members before ICSE changed to the program board model).

Year	Submissions (after desk rejects)	Nr. of PC members	Resulting review load for PC members
2014	324	39	16.6
2015	320	37	17.3
2016	333	36	18.5
2017	368	31	23.7
(projected) 2018	*390*	*45*	*17.3*

ASE has been at the forefront of exploring new approaches for evaluating submissions. It introduced and refined the ERP model and was the first major software engineering conference that switched to double blind reviewing. The latter step was accompanied with significant data gathering and was followed by both ICSE and FSE. We think it might be time to explore further changes.

Physical PC meetings have often been seen as a cornerstone of high quality conferences, enabling in-depth discussion and calibration across papers. It also provides a reputation

mechanism and a networking opportunity for PC members and a forcing function for PC members to submit high-quality reviews in time for the meeting. At the same time, PC meetings are associated with significant costs and problems:

- Ecological cost of 20-40 long distance flights (40 tons of carbon d.)
- Economic cost includes hotel and flight cost paid by PC members (800-2000$ for most PC members; ~65K$ for the PC meeting) in addition to 10K$ for the meeting room, catering and dinner covered through registration fees (about 50$[1] per registration). This adds up to roughly 500$ per discussed paper or 180$ per submitted paper.
- Time investment per PC member is typically 2 full travel days plus two full days for the meeting, on top of the reviewing work and the online discussion. Even when colocated with another event like ISSTA, many PC members travel solely for the PC meeting.[2]
- Unequal representation of ERP and PC members requires significant online discussion and may still give ERP members less of a voice in the decision.
- The increasing number of (accepted) papers leads to more papers being discussed, resulting in shorter discussion times at the PC meeting and longer and more rushed meetings. As there is an upper limit of papers that can be realistically discussed (independent of PC size), increasingly more papers are decided online before the meeting, requiring more online discussion, leading to more PC members traveling to the PC meeting to discuss only very few papers,[3] and resulting in questions about fairness of online-only decisions that are not as calibrated and potentially do not benefit from the expertise of the entire PC.

We discussed the issue of a significant increased PC size and the various associated problems with members of the ASE steering committee during the ASE 2017 PC meeting. There was general agreement that there is a problem, but little consensus of how to solve it. There were many suggestions, including:

- Splitting the PC and having 4 days of PC meetings, with each half of the PC meeting for 2 days (increased costs, increased noise).
- Experimenting with early rejects, reducing the number of overall reviews (fairness issues, requires longer review period, benefit uncertain). ICSE experimented with this model for a number of years.
- Using more ERP members (e.g., 2 ERP + 1 PC member) or using a PC board model (more focus on online decisions, fewer informed reviewers at PC meeting for decision making). ICSE uses a board model where only board members overseeing reviews of PC members meet.
- More reliance on video-conferencing, e.g. skyping in ERP members (logistic overhead of running the meeting). AOSD experimented with video-only PC meetings at a smaller scale.
- Journal first model, in which all papers are submitted to the ASE journal and all accepted papers in a year are presented at the conference (significant change of the entire

---

[1] Rough estimate based on 200 registrations.
[2] In 2017, only 9 out of 31 ASE PC members attended ISSTA.
[3] There have been cases recently of PC members flying cross-atlantic for discussing 3 papers.

> publication and reviewing model). This model was adopted in other communities including SIGMOD and SIGGRAPH.
> - No PC meeting; only online decisions (loss of benefits of face to face discussions). This model is used by many other communities including CHI and WSDM.

Given the current problems and alternatives, we propose to adopt an online-only model without a physical PC meeting. Without the physical meeting, we can experiment with other innovations that could compensate for some disadvantages, as discussed below. Similar to the introduction of double blind reviewing, we will monitor and evaluate the change to support decisions for future years.

The change will contribute to a reduction of the carbon footprint of the conference culture of computer science and helps to reduce the conference registration fee. The change would not be unprecedented. POPL 2018 changed to an online-only model.[4] Many large conferences in other fields have online-only models, including CHI and WSDM.[5]

Concrete Plan for ASE 2018:

We will invite 100-120 PC members (no ERP) to the ASE PC (i.e., larger than ASE 2017 with 31 PC and 34 ERP members, larger than ASE 2014 with 39 PC and 28 ERP members; slightly larger than the alternative of 45 PC and 45 ERP members planned for 2018 if a physical meeting is needed), using the same rigorous selection process as in past ASE conferences.

We will assign 10-14 papers to each PC member in a 6 week reviewing period.

We will conduct a 2 week discussion period.

Each PC member will be the discussion leader on 5 papers that they have not reviewed (program board style without the hierarchy;[6] PC chairs step in as needed).

We will require each PC member to be active in discussions (at least one response to the other reviews for each paper; otherwise removal from the PC like for physical PC meeting).

We will open discussions on all papers to all non-conflicted reviewers and actively solicit extra reviews or opinions on unclear cases. We will develop an explicit process guide for such cases,

---

[4] http://popl18.sigplan.org/track/POPL-2018-papers#Call-for-Papers: "As an experiment for POPL 2018, the program committee will discuss papers entirely electronically rather than at a physical programming committee meeting. This will avoid the time, cost and ecological impact of transporting an increasingly large committee to one point on the globe. Unlike in recent years, there will be no formal External Review Committee, though experts outside the committee will be consulted when their expertise is needed."
[5] Andrew Tomkins, Min Zhang, and William D. Heavlin. Single versus Double Blind Reviewing at WSDM 2017. arXiv preprint arXiv:1702.00502 (2017).
[6] This model has been successfully used by ICST to engage PC members.

for example, emailing the list of controversial papers to all non-conflicted PC members, or to subsets with relevant expertise.

We will hold a PC dinner the evening before the main conference starts to recognize the efforts of the PC, provide networking opportunities, and encourage more PC members to attend the conference.

We will remain lightweight double blind (reviewers do not know authors until the discussion period; reviewers know each other). We will continue to award distinguished reviewer awards.

Depending on eventual PC size and number of submissions, we will consider to assign 4 reviewers to each paper instead of 3.

Quality assurance

- We will run a survey among 2017 authors about their satisfaction with the review quality and the transparency of the decision. We will repeat the survey 2018 and compare results.
- We will run a survey among 2017 PC and ERP members about their experience and workload. We will repeat the survey 2018 and compare results.
- As required also with physical PC meetings, we will actively monitor review quality and discussion quality and intervene as needed. A larger PC provides more opportunities to requests extra reviews if needed.

Risks and mitigation

- Without the forcing function of a meeting, some reviewers may not participate in online discussions. This was already a problem with some ERP and PC members in previous years. Mitigation: clear expression of expectations; explicit responsibilities to respond to other reviews and moderate discussions on papers; threat of removal from the PC.
- Without the forcing function of a meeting and a larger PC, some reviewers may be late or provide low quality reviews. This problem existed before, but might be increased in an online-only meeting. Mitigation: clear communication of expectations, clear threat of removal from the PC, 4th review, ...
- Without the networking benefits community members might be less inclined to accept the PC invitation. Mitigation: Reduced load per reviewer; PC dinner at the conference. We expect this problem to be minor, as many invited PC members in the past voluntarily selected to serve on the ERP, and ERP acceptance rate was generally high.
- Less calibration across papers and fewer chances to include additional expertise in the PC in the discussion. Mitigation: active discussion moderation and additional reviews if needed. We argue that in the past only few decisions were strongly affected by these benefits (especially as the number of discussed papers increased and more decisions

were done online anyway). Given the general noise in reviewing,[7] we argue that the significant cost benefits of the model and the new opportunities can outweigh the slight increase in noise (notice that there is no reason to believe that this causes a fairness issue, where noise would affect certain kinds of papers more than others)

---

[7] There is much published evidence on this, including the NIPS experiment and the merit model discussed in *Thomas Anderson. Conference Reviewing Considered Harmful. ACM SIGOPS, Vol 43(2): 108-116 (2009)*