The ACM SIGCOMM 2009 Technical Program Committee Process

Konstantina Papagiannaki
Intel Research Pittsburgh

Luigi Rizzo
Università di Pisa

ABSTRACT

Selecting a technical program for a conference, and running the process so that decisions are well received by authors and participants, is a challenging task. We report our experience in running the SIGCOMM 2009 Technical Program Committee (TPC). The purpose of this article is to document the process that we followed, and discuss it critically. This should let authors get a better understanding of what led to the final acceptance or rejection of their work, and hopefully let other colleagues in charge of similar tasks make use of our experience.

Categories and Subject Descriptors

A.m [**General literature**]: Miscellaneous

General Terms

Documentation

1. INTRODUCTION

Sigcomm is a single track conference that usually attracts a high number of submissions, around 300 in recent years. The acceptance rate varies by year between 8% and 12%. Despite the single track format, the conference addresses a relatively broad set of topics. This calls for a unified evaluation of all papers by a single set of reviewers, which in turn impacts the size of the TPC and the workload of each member. Other conferences with a similar or larger number of submissions sometimes partition the overall conference in a number of tracks and form a number of PCs to decide the technical program of each track.

Receiving many submissions is both a challenge (scaling and management issues) and a resource (it leads to better quality published papers). Sigcomm has the great fortune to be regarded very highly by the research community in systems and networking. As a result, researchers tend to enthusiastically agree to participate in the TPC despite the associated workload. The TPC chairs and the committee as a whole are then faced with the usual challenges: implementing a fair review process, returning adequate feedback to all authors, dealing with strict turnaround times, and deciding where to "set the bar" for papers to be accepted.

In what follows, we discuss how we addressed those questions while preparing for ACM Sigcomm 2009. The document itself is organized in a number of sections outlining the several steps involved in organizing and running the TPC.

Due to space limitations, this article does not discuss every aspect of the reviewing process. Interested readers can refer to a growing body of literature on this topic, including [8, 4, 1]. People interested in running their own analysis will find suitably anonymized data from the ACM Sigcomm 2009 HotCRP site at:

`http://www.sigcomm.org/about/conference/the-process-of-acm-sigcomm-2009/`

Please be very careful in drawing conclusions from this dataset: many aspects of the review and decision process cannot be captured by numbers, and especially when it comes to scores for individual papers, the sample size is really small and the numbers have large variance.

1.1 Traditions

As with every long standing event, Sigcomm comes with an established tradition on the format of the conference, its procedures, and expectations on the content of the final program. Deviations from these traditions must be carefully evaluated because they may not be received well. Changes should also be advertised loudly, so that they reach also authors or reviewers who do not "read the instructions".

In summary, the tradition is to have a single track, three day conference with a double blind reviewing process that keeps authors and reviewers anonymous to each other. There is only one type of submission, namely full length papers, with a request to register title and abstract one week in advance. The areas of interest are nearly any topic for which the authors can prove relevance to networking.

The TPC chairs are generally appointed by the ACM SIGCOMM Executive Committee, whereas TPC members are appointed by the chairs. All TPC members are requested to do the reviews themselves, and some of them, called the "heavy TPC", are required to take part in the TPC meeting. There are some well established rules (within this community; other communities may have different rules) on what constitutes a conflict of interest that prevents a reviewer from reviewing a specific paper (see also Section 3.2). Part of the tradition is also the expectation that authors will receive multiple and very detailed, technically accurate, reviews of their papers.

1.2 Dealing with people

A good part of running the TPC process has to do with dealing with people. The chairs interface with authors, indirectly when preparing the Call for Papers and submission instructions, directly when managing non-conforming submissions, and even more directly when dealing with com-

plaints from authors of rejected papers. Chairs also interact heavily with reviewers throughout the selection process. Finally, reviewers "interact" with authors through the content of their reviews.

In all these phases there is a delicate balance between "politeness" and effectiveness of the communication. For instance, it could be seen as inappropriate for us to give too detailed instructions on how the reviewers should perform their task [9, 3]. On the other hand, not having a common set of explicit reviewing guidelines does not allow for the direct comparison of reviewers' scores.

Similarly, while authors may not want to be told how to write their paper and support their findings, giving a bit of explicit advice might have avoided some of the many cases, frequently reported by our reviewers, of papers overstating their contributions, providing meager evidence in support of their claims, and omitting discussion on weaknesses or limitations of their proposals.

Other common communication issues arise in the reviews: encouraging words from the reviewers are nice, but not always useful if politeness hides the message that should be delivered. Obviously, offensive or demeaning tones (both in reviews and rebuttals) should be absolutely avoided.

Our first lesson, acting as the interface between people, was that it was tremendously valuable to have two people sharing the role of the TPC chair; when it comes to reading between the lines and fine tuning any written communication with the TPC itself, the authors or the executive committee, a second perspective was critical. Second, we learned that all information and associated assumptions need to always be explicitly stated. Even though we knew this, in some cases we failed to apply the principle.

As an example: several authors registered placeholder titles and abstracts for their papers. While they may have followed the instructions to the letter, they certainly missed the point, which is mainly to allow reviewers to express their preferences for papers they would like to review, while the submission is being finalized. The requirement for abstract registration is also a checkpoint for quality control. One could argue that if the authors do not have a clear idea about the title and abstract of their paper a week in advance of the submission, then they are probably not ready.

2. THE CALL FOR (SHORT) PAPERS

The two main features of the call for papers are the list of conference topics, and the type of submissions accepted. We stayed within tradition here: the CFP solicited full papers with a list of topics that only marginally deviated from last year's. We do believe that an open list of topics is beneficial as it permits submissions on newly emerging areas. The downside of this choice is that i) the TPC must have expertise on a wide set of topics, and ii) it encourages authors to submit just about anything, even when the relation with networking is extremely limited.

A few years ago, Sigcomm did not only call for full submissions but also position papers. It was seen as a way to get new, potentially controversial ideas in front of the Sigcomm audience for feedback. The experiment was dropped shortly thereafter, and short papers have not been solicited recently.

We do see a problem with requesting only 14-page submissions, and want to use this document to convey a position we frequently heard from TPC members as well: *short papers could be an interesting addition to future CFPs.*

In a conference with such a low acceptance rate as that of Sigcomm, "small contributions" naturally find themselves at a disadvantage. The issue is that the description and evaluation of an ingenious but small idea certainly takes less than 14 pages. The same is true for work that improves on prior work or presents tools with important practical use but small technical contribution. The common practice for authors of such work is to pad the paper up to full length, often with negative effects on paper quality: either the padding brings in a lot of background material, which makes the paper feel more and more incremental, or it makes the presentation redundant. Both approaches are perceived negatively by reviewers.

We believe (and on this there was probably some consensus also in parts of the TPC) that the conference could benefit from soliciting "small contributions", whose submission format could be limited to shorter papers. Past experiments of this kind have not been very successful, perhaps because authors think of them as "B class" contributions and are more interested to compete for the premier league. A good specification of the intent of short contributions, and suitable rules for submission, will be central in encouraging short paper submissions. However, picking short papers as leftovers from the main selection is not a good way to go: authors in the first place should be aware of the level of their contribution and submit accordingly.

3. SELECTING THE TPC

The mandate for selecting the Sigcomm TPC has been quite open, with only a suggestion to implement diversity of the TPC under a number of different metrics (gender, expertise, geographic areas, institutions, age and so on). In our view, diversity comes naturally when trying to achieve the true goal of a fair and high quality program, able to deal with various topics and within the constraints imposed by conflicts of interest.

3.1 Light, heavy and senior TPC

With the current set of submissions, the requirement to do all reviews within the TPC (with only a few exceptions) either calls for a large TPC, or imposes a great load on TPC members. We estimated around 900 reviews for this year, and in fact ended up with 826 reviews.

For a few years Sigcomm has used a two-tier TPC: "light" members are asked to do 10-15 reviews and are not invited to the meeting; "heavy" members are enrolled for 20-25 reviews plus mandatory participation in the TPC meeting. The split is mostly motivated by keeping the meeting of manageable size.

This year we introduced a third tier, called "senior" for lack of a better name. The role we intended for senior TPC members was a mix of conflict solver, an additional voice during the TPC meeting, helping hands for last minute reviews, and possibly carriers of a different perspective that could put the submitted research under a more positive light. We asked the seniors to get involved in the last three weeks of the reviewing process, and read through 10-15 papers (and their reviews) among those still under discussion at that stage. Note that this type of involvement is more time consuming than simply having to read a paper and provide a review, because it implies engaging into discussions, comparing papers, and interacting with the chairs.

The final TPC comprised 27 light, 25 heavy and 6 senior members. The choice of people to be invited in the various roles was driven by their expertise in their respective areas, but also by the existence of some kind of "chain of trust" between us and the candidates. We felt that this chain of trust was a necessary requirement to improve the chance of a smooth and successful operation of the TPC, especially considering that we have no authority on TPC members to demand that they complete their heavy workload on time and with the desired quality. Part of this chain of trust was the knowledge of a good track record at reviewing in prior TPCs. We also tried to extend the team beyond the set of former Sigcomm TPC members or authors, to avoid inbreeding while covering a breadth of topics. For some specialized topics we reached critical mass by engaging a small number of external reviewers.

3.2 Conflict of interest

ACM SIGCOMM defines a conflict of interest between an author and a reviewer if the two have worked together in the past two years. Students and advisors are considered "conflicted for life" and of course any institutional or private relationship between the author and the reviewer instantly qualifies as a conflict.

In our double blind review system, conflicts are declared by the authors. For the first year, we had the paper management system (HotCRP [7]) modified to store the reason why each TPC member was declared as a conflict. In addition to the above categories, we added an "other" category to capture situations that could not be easily explained (e.g., sometimes individuals do not get along well), even though we knew that this category could be abused, e.g. to avoid that a paper is judged by certain reviewers.

The categorization allowed us to verify the correctness of conflict statements, correct a couple of instances in which authors misunderstood the extent of conflicts, and identify one case where authors, unhappy for the treatment their paper received in a past conference, declared a conflict with the overlapping part of the PC between that previous conference and Sigcomm. We consider this motivation to be really borderline, and we believe that the problem should have been addressed by talking directly with the chairs. Nevertheless, it did not influence the review of the paper or impair the selection of reviewers.

3.3 Chairs' conflicts

To ensure the same treatment (in terms of double blindness) to the papers that the TPC chairs had authored or had a conflict with, we created a second instance of the paper management system, to which the chairs had no access, and that was managed by one of the senior TPC members who was constantly informed of the process followed on the main site. That site held 10 papers in total: those in conflict with the TPC chairs, and a number of randomly selected papers from the overall submission pool to further obfuscate the identity of the chairs' papers from the reviewers.

In retrospect, we feel that splitting the systems was not a good choice. A big part of the review process involved comparing papers, scores and reviews, and having papers on two different systems was limiting this comparison. We now believe it would be better to keep all papers on the same site, and let reviews for one chair's paper to be handled by the other chair, trusting the chairs' integrity and professionalism. To further inhibit abuse one could even consider keeping the HotCRP logs for verification.

4. THE REVIEW PROCESS

Our goals in designing the review process were to i) guarantee fair treatment to all papers; ii) keep the workload manageable for all people involved; iii) achieve an early detection of problems such as missing, low confidence or conflicting reviews; iv) enable an extensive discussion before the TPC meeting; and v) ensure that each paper could be discussed by at least 3 people during the TPC meeting.

Fairness is obviously a meta-goal, and hopefully the result of the various strategies we put in place. The reduction of workload was addressed by allocating reviews to papers as deemed necessary, described in the following paragraphs.

To minimize the impact of late or otherwise problematic reviews, we organized the process in three non-overlapping rounds with a few days for discussion and management between them. In reality, late, declined and low-confidence reviews caused some overlap among the various rounds. Discussion was open and solicited at all times, and subsequent reviews were allocated to heavy or senior reviewers to give enough coverage at the meeting.

Throughout the process, the chairs monitored all reviews and scores as they were coming in, started discussions on the papers, solicited clarifications from reviewers, and made assignments of extra reviews in all cases where the existing reviews were in disagreement, had low confidence or the reviewer themselves asked for more opinions. All data entered by reviewers (reports, scores and comments) could always be modified during the process, even though only a few reviewers made use of this option.

4.1 The three rounds

During the first round, each paper was allocated one heavy and one light reviewer, with a deadline of four weeks. The discussion between the reviewers and the chairs started as soon as the reviews were in.

The second round had one or more extra reviews allocated for all papers for which there was not a clear and unanimous consensus for a reject between the chairs and the reviewers. The deadline for the second round reviews was three weeks later, and the allocation targeted expert reviewers able to address specific concerns raised during the first round.

In the third round we allocated extra reviews in the same way as in the second round. In this round we also solicited 10 external reviews (out of 826 reviews in total) when the required expertise was not captured in any TPC member that had not already reviewed that paper. Senior members got engaged in this round. They were assigned 10-12 papers each, and asked to make a "short review", aimed at resolving conflicts or clarifying areas of uncertainty.

From the beginning of the third round, access to papers, reviews and discussion was open to all TPC members on all papers, of course subject to conflicts. Having access to the reviews can be thought as introducing some form of bias, but we felt that the benefits would largely exceed the risks, if any: third round reviews were normally the fourth or more for each paper, and at this stage we really needed to focus on the issues that required the most attention. By the end of this process some papers were reviewed by up to 7 people (plus the two chairs), and this led to about 70 papers brought up for discussion at the TPC meeting.

4.2 Deadlines

Setting a deadline (in this case for reviews) means nothing if you don't have a way to enforce it; and we didn't. Splitting the process in multiple rounds was an attempt to achieve more control on the timeliness of reviews, and it worked, but only to a certain degree. Even though we planned a small grace period (about 1 week) at the end of the first two rounds, about 20% of the reviews were missing at the end of the first round and the grace period; this slightly delayed the subsequent allocation and the handling of some papers. The accumulation of missing reviews by the end of the second round required us to use our best flattery to convince our colleagues to take on late reviews.

The way we dealt with missing reviews was first to allocate extra reviewers, but also avoid assigning too many extra papers to the late reviewers. A slightly larger PC and the variable allocation of reviews to papers naturally allowed for such actions.

4.3 The review form

A review form usually contains multiple text fields, and one or more numeric fields used to sort papers according to some criteria. We kept the number of text fields as small as possible to avoid the confusion that arises when it is not clear what information should go where: we had "Paper summary", "Comments for PC" (hidden from authors) and "Comments for Authors". "Paper summary" was meant to contain an executive summary of the paper, which in turn should describe what is the actual contribution of the paper according to the reviewer: while authors should usually include this information in their submission, the authors' view is not always present or accurate.

As for numeric fields, in addition to the standard "Overall Merit" and "Reviewer Expertise", we proposed (and this was a mistake) a few extra fields to capture the technical accuracy of the work, its longevity, its novelty and its community interest. These extra fields were not really useful during the process, for at least two reasons: 1) some of them refer to the topic and not to the paper, so they are not useful for selection, but perhaps only for statistical analysis of the submissions; 2) for some fields the available choices did not fill the full spectrum of possibilities, so reviewers were forced to pick random values. We would advise future chairs to omit fields from the review form if they have not a clear idea on how to use the information they carry.

We mostly used overall score and reviewer confidence during our many scans of the database. In particular, throughout the process, we used various functions (average, min, max, variance) of the "overall score" field as sorting keys to detect papers which did or did not require more attention. Finally, we always asked for clarification when the textual fields and numerical scores did not seem to match.

4.4 Scores and ranks

Overall merit was structured as a 5-level log-like scale, with 1 meaning that the paper is in the bottom 50% of submissions, 2 for the top 25% to 50%, 3 for the top 10% to 25%, 4 for the top 5% to 10%, and 5 for the top 5%. As expected, reviewers were reluctant to use the extremes of the scale: the top score was used only 4 times, the bottom score was used only in 25% of the reviews. The remaining categories mostly followed a log-like scale, as shown in Fig. 1.

The observation that reviewers tend not to use the extreme scores is common to all TPCs we have been part of, and a big problem with the selection process, especially as the number of submissions grows. Related to this is the fact that reviewers have different behaviors, and their perception of the quality of papers differs. One interesting proposal is to ask reviewers to rank papers *within their batch* [5] and then spend some time before or during the meeting trying to harmonize the various partial orders into a global one (at least approximate). The partial ranking is a practice that several reviewers already use. The "harmonizing" phase would be explicit and applied to all reviewers, so it should not be viewed negatively.

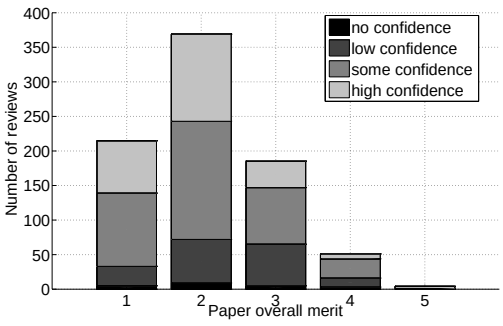

Figure 1: Number of reviews versus paper overall merit, and associated reviewer confidence level.

4.5 Reviewer confidence

Fig. 1 also shows the distribution of reviewer confidence for each review score. We are glad to see that the vast majority of all reviews were given with the top two confidence categories. Not surprisingly, high reviewer confidence tends to be correlated with lower scores. This is a well known phenomenon that might deserve a correcting factor when averaging scores with different confidence levels.

On the other hand, this also means that low scores (possibly leading to rejects) usually carry high confidence, and so reject decision are not taken lightly. The lack of a high confidence opinion in the evaluation of a paper always triggered the solicitation of an expert review throughout the process.

One common criticism on the Reviewer Confidence field is that it is self-assessed and as such not very reliable. We find the criticism a weak one. In fact, we have seen that a number of TPC members tend to use the second highest confidence level despite their expertise, probably due to their modest character. Having said that, the weight given to a review does not only come from the Confidence field, but also from its content and the reviewer's ability to convince the other reviewers and TPC chairs about the validity of their point.

4.6 How many reviews

A common expectation from a top-level conference is that every paper gets reviewed by at least three experts, and is given substantial feedback on how the authors can improve or extend their work. In recent years, to deal with the large number of submissions, the minimum number of reviews returned by Sigcomm has been two. We followed that same tradition, still trying to preserve the amount of feedback returned to authors. Clearly, the more reviews one paper

receives, the better it is. But guaranteeing a third review to all papers would have increased everyone's load, and also the duration of the first round of reviews. In turn, this would have reduced the time for discussion and our chances for a prompt handling of delays or other problems.

Number of reviews	1	2	3	4	5	6	7
Number of papers	1	135	58	30	31	13	4

Table 1: Number of reviews per paper. Only one paper received a single review because it was out of scope.

4.7 Review allocation and management

Especially for papers receiving only two reviews, it was important that the initial allocation would give a good match of reviewers to papers. Since no allocation method is error proof, it was even more important to make sure that the reviews returned were analysed critically, so that extra reviews could be allocated promptly if needed.

For the initial allocation we relied on the categorization given by authors, and on the bids (on topics and on specific papers) made by reviewers. This info, together with conflicts, was processed using the HotCRP paper allocation algorithm, and the only manual adjustments we made were to ensure that each paper had at least one heavy reviewer (so it the reviewer would be present in the TPC meeting).

We felt that, in the first round, an automatic allocation step was at least as good as what we could have done manually, and having a well specified algorithm trying to optimize the allocation basing on available information would remove any bias that could arise in a manual allocation process.

After the first round, all review allocations were done manually: at this stage, the key information for allocating reviews was sitting in the reviews themselves, and not in a suitable form for automatic processing.

Controls on the quality of reviews were exercised very carefully by the chairs, who read all reviews as they were coming in, and asked for clarifications or allocated additional reviews in case the existing ones had unsatisfactory content, reviewer confidence, or were conflicting with each other. The majority of the TPC provided very thorough and timely reviews ensuring no last minute stress. Some reviews were indeed late, but thankfully without signs of sloppiness.

There is no good and precise metric for review quality. We can only report that the average review length was around 3200 characters, independently of the individual overall score. Several reviews were even much longer, counting 10,000 characters and above. Most of the longest reviews have been provided for score 2, which was shown to also correlate with high confidence.

5. THE TPC MEETING

This year we extended the TPC meeting to one-and-a-half days. The extra half day allowed us to dedicate sufficient time to each paper without having to rush decisions at the end of the meeting. This change is certainly one we would recommend to future chairs.

5.1 Selecting papers to discuss

The goal of the various review rounds was to trim the set of submissions so that the meeting could focus on the subset containing the best submissions. The trimming occurred continuously throughout the various rounds, as information was available and unanimity was reached between the reviewers and the chairs. No reject was made lightly, and no reject was indeed final until the last day of the TPC meeting. The perception of the "acceptance bar" was obviously changing over time, and during our many many scans of the papers we frequently revisited those marked as reject, possibly bringing them back into discussion if we felt unsure about previous decisions. Reviewers were also encouraged to submit requests for resurrecting papers if they felt we made a wrong decision. A small number of papers were indeed brought back into the discussion.

Eventually we brought up for discussion approximately 70 papers, whose list was prepared by the chairs using all input available.

5.2 Discussion order

As in most conferences (see [2], Fig.1), the confidence intervals for the scores of papers being discussed heavily overlapped. So, rather than discussing papers by average score, we structured the overall discussion in thematic entities: papers were grouped by topic, and discussed in random order within each topic. After a brief summary of the paper with its strengths and weaknesses, the TPC discussed the paper and made a decision on whether it should be rejected or ranked within its topic.

The grouping by topic was meant to amortize the differences in maturity of the various areas (hence the expectations of the reviewers), while the ranking within each topic was meant to preserve the state of the TPC decisions for the end of the meeting, when the final decisions had to be made. Having ranked lists would (in theory) allow the TPC to derive a program of any target size as long as the number of ranked papers exceeded that target. We dedicated two hours at the end of the TPC meeting to merge the ranked orders of the different topics, ensuring that quality was the main factor in the decision. The TPC meeting was attended by two PhD candidates of the hosting university that scribed notes on the decision rationale, returned to the authors with the notification email.

Unfortunately, and despite our best intentions, only 30 papers managed not to get rejected (our maximum target for acceptance was 34-35), and some of them were still perceived to have unacceptable technical flaws. In the end, the TPC accepted 27 papers in total, some of them conditionally, making the program the same size as in 2005, and slightly shorter than that of the past two years. Whether paper quality is declining (see [4]), authors are squeezed between too many deadlines, or reviewers are too demanding is open for discussion.

5.3 Shepherding

A number of the papers discussed at the TPC meeting were conditionally accepted with "shepherding": a TPC member was assigned to oversee that the authors would properly address certain issues that were identified during the reviewing process.

There were a few things that surfaced as concerns in need of shepherding. 1) The paper did not describe its methodology in a clear enough way that it could be reproduced. 2) The paper overstated its contribution and did not contrast fairly to related work. 3) The paper did not study the pro-

posed scheme to the extent that it was able to define its area of applicability - a limitations section was missing.

As of this writing, the shepherds are still working with the authors to resolve those issues. The final acceptance of the work will be based on how well the authors have addressed the relevant issues, either in their camera-ready, or through communication with the shepherd and the chairs in which they justify their disagreement with the proposed changes.

5.4 Rebuttals

We discussed implementing a rebuttal phase, and decided against it based on the following reasoning. On the positive side, rebuttals could address those (hopefully rare) cases where the reviewers completely misunderstood a paper. On the negative side, implementing rebuttals would impose more deadlines in the review process, and require a lot of discipline for authors and reviewers to avoid rebuttals becoming a second submission [6, Sec.2].

Having said that, we did follow up on four specific complaints we received after the decision notification. In three out of four cases, we could not find faults in the process or in the reviews (in one such case one of the chairs even supplied an extra review). The complaint made on the last case regarded the style of the review that was perceived as rude, for which the reviewer apologized. Reviewers are only human and sometimes break under pressure as well. We should also mention that some of the authors of those complaints still remain with their opinion, and that we decided to stop some discussions when comments started being made on the competence of the reviewers.

Given our experience with these exchanges we tend to believe that at least in our case a rebuttal would bear limited gains while having a substantial overhead.

5.5 Handling Malpractice

There are a number of incorrect practices (both moral and technical) that may surface in this process, involving authors, reviewers and possibly also your trustworthy TPC chairs. It is necessary to be prepared to handle them, but the countermeasures should not cause more harm than good. Examples include plagiarism, self plagiarism, violations of anonymization or double submission policies, abuse of the conflicts of interest mechanism, reviewers' bias in favor or against people or topics, authors overstating their contribution or providing dubious data. We believe it is up to each year's chairs to decide the way to deal with them.

We encountered some (fortunately a small number) of these problems and addressed them on a case-by-case basis. In particular, we identified two cases of double submission. In cooperation with the chairs of the other conferences, we rejected the papers from both venues without reviews.

6. LESSONS LEARNED

We conclude with a brief summary of lessons learned:

- include short papers in the CFP, but be extremely clear on what they are and how they differ from full papers;
- plan well ahead of time on what scoring/ranking strategy you want to use;
- keep human factors into account. Expect delays and plan accordingly;
- do not put fields into the review form for which you do not have a clear use;
- be extremely explicit when giving instructions to authors and reviewers. Do not be afraid to state the obvious;
- keep the three tier committee, perhaps clarifying better the role for senior reviewers;
- have a day and a half TPC meeting;
- use a good conference management system that allows you to make various data analyses. No single ranking scheme will solve all problems so you will need to look at the information from different points of view;

Finally, remember that a program is as good as its TPC. Select TPC members that are well respected, timely and able to articulate and back up their position.

7. ACKNOWLEDGMENTS

We would like to thank the ACM SIGCOMM executive committee for giving us the opportunity to act in the capacity of TPC chairs, all the reviewers involved in the SIGCOMM 2009 conference for their hard work, and Prof. Eddie Kohler, who developed and maintains the HotCRP conference management system, which was irreplaceable in helping us run the process. Eddie did not only explain to us the full potential of HotCRP but also implemented some features upon our request.

8. REFERENCES

[1] SIGCOMM Program Committee BCP. `http://www.sigcomm.org/about/policies/sigcomm-program-bcp/sigcomm-pc-bcp/`.

[2] T. Anderson. Towards a model of Computer Systems Research. In *Proceedings of the Workshop on Organizing Workshops, Conferences, and Symposia for Computer Systems*. USENIX Association, 2008.

[3] G. Cormode. How NOT to review a paper. The tools and techniques of the adversarial reviewer. *SIGMOD Record*, 37(4), December 2008.

[4] J. Crowcroft, S. Keshav, and N. McKeown. Scaling the academic publication process to internet scale. *Communications of the ACM*, pages 27–30, 2009.

[5] J. Douceur. Paper rating vs. paper ranking. In *Proceedings of the Workshop on Organizing Workshops, Conferences, and Symposia for Computer Systems*. USENIX Association Berkeley, CA, USA, 2008.

[6] P. Francis. Thoughts on improving review quality. In *Proceedings of the Workshop on Organizing Workshops, Conferences, and Symposia for Computer Systems*. USENIX Association, 2008.

[7] E. Kohler. Hot Crap! In *Proceedings of the Workshop on Organizing Workshops, Conferences, and Symposia for Computer Systems*. USENIX Association, 2008.

[8] J. Mogul. WOWCS'08: The Workshop on Organizing Workshops, Conferences, and Symposia for Computer Systems. `http://www.usenix.org/event/wowcs08/`, San Francisco, CA, 2008.

[9] A. Smith. The task of the referee. *IEEE Computer*, 23(4):65–71, 1990.