You Must Be Joking...
A Historic Open Reviewing at Global Internet '07

Michalis Faloutsos
UC Riverside
Dpt. of CS
Riverside, CA
michalis@cs.ucr.edu

Reza Rejaie
U. Oregon
Dpt. CS
Portland, Oregon
reza@cs.uoregon.edu

Anirban Banerjee
UC Riverside
Dpt. of CS
Riverside, CA
anirban@cs.ucr.edu

ABSTRACT

We did it! A few CCR issues back, this column argued in favor of an all open review. At the time, most people thought that it was a joke. Quite frankly, we meant it as a joke. It is a crazy world we live in. No, seriously, it was not so much of a joke, as a wild attempt to think outside the box. Or else, desperate times call for desperate measures. Something like that. Apparently, we were so desperate that we actually tried it. And it worked. Beautifully. But then again, we may be biased.

> *"To say it was easy would be a lie."*
> *R. Rejaie and M. Faloutsos, early 21st century*

1. INTRODUCTION

Here is a recap for the non-fans[1] of the column in three statements: (a) getting quality reviews is a problem, (b) reviewers have too much power and protection due to their anonymity and lack of memory in the system, (c) let us make reviews and reviewer names public to increase accountability. At the time, many people said: "easier said than done". And they were absolutely right. It was much easier to write that column than to actually do it.

The facts: Global Internet 2007, chaired by Rejaie and Faloutsos, had an open review process: the names of the reviewers were revealed to the authors and, in addition, reviews of accepted papers[2] are published in a website for anyone to see.

How cool is that?

The rest of this column attempts to distill what we learned from this experiment.

Seriously now, we did ask for feedback and most people found the process reasonable if not better and they would like GI to continue with the open review format. We will discuss the feedback later.

Based on some initial feedback, we would like to clarify the following regarding the scope and intention of this article:

---

[1]Quite frankly I don't see a reason why you should not join the other two fans of the column. Furthermore, I don't see any reason why you should not send monetary support in the thoughtful way that people do for freeware. Sharing wealth is a beautiful thing. Think about it.

[2]We thought of doing the same for the rejected papers, but that could create a lot of commotion: exposing the idea early, while at the same time blurring the line of accept and reject to some extent. However, we think it may be worth considering it in the future. If your paper is rejected unfairly, you probably want the reviews to be made public...

1. *For real or for laughs?* Despite the hopefully humorous tone of this article, the conclusions, observations, and suggestions are meant to be serious.

2. *Science or fiction?* The article below presents data and facts, and our interpretation of them, along with some personal opinions. Other interpretations and opinions are possible.

2. THE LONG AND WINDING ROAD

Reza and Michalis started by discussing that it would be really cool to actually try the open review approach.

Michalis likes tinkering with things, as can be attested by his wife: "Yes, he always breaks things around the house". Does he also fix things? "Not so much".

Reza was surprisingly quick to get on board with the idea. Luckily, he was also more reasonable and organized than Michalis.

Getting the green light from the Steering Committee. The start is half of the battle. And in this case, the most critical battle was the start. Apparently, there was a lot of concern among the Steering Committee that this would flop: none would want to join the TPC, none would want to send papers, none would want to mention the letters G and I in this order.

All jokes apart, we can't blame them for being concerned, and we appreciate their courage in trusting us in this effort. Many people over the years have worked hard to bring GI where it is now. We happen to be among the people that love GI. We think it fills in the niche of what a conference really should be about: a quick way to share preliminary ideas, get feedback, and put a time-stamp on an idea. We argue that it is the best 6-page workshop for general wireline networks.

The rules of conduct. Defining how the reviewing would be done was more involved than one can think. This is what we decided to do, and we find that it worked well. This is pretty much the email we sent to reviewers when explaining the process:

" 1. The reviewers have the option to not see the authors names. This is for the "protection" of the reviewers[3], but they can forgo this right and learn the names of the authors before doing the reviews.

---

[3]In fact, we have to give credit to our colleague Srikanth Krishnamurthy for this idea. However, let it be known though that Srikanth was against the open review effort, although he is Michalis' academic twin brother.

2. Each reviewer provides a technical review along with a **ranking** for each assigned paper. The ranking simply shows the overall quality of a particular paper within the batch that is assigned to a reviewer. Therefore, reviewers do not suggest a specific outcome (e.g. reject, accept) for a particular paper, but, if they want, they could suggest it in their review.

Note that each reviewer has to use distinct ranking numbers for each paper starting from 1 for the best paper of the batch. i.e. we should not have papers with the same ranking as this alters the meaning of ranking and breaks the consistency among different reviewers. (Allowing for non-strictly monotonic ranking would not work well. For example, consider: a ranking 1,1,1,2,2,2 3,3,3,3 the worst paper here has a ranking of 3, which would be compared to the top 3rd of another reviewer. Normalizing would not work well either. Trust us...)

3. The selection process: the selection process is roughly similar to that of other conferences. The average ranking is used to order the papers. The Chairs upon inspection will define some thresholds for clear accept and clear rejects, while reviewers can move to resurrect or reject any paper. Large variance of rankings will prompt attempts to resolve the differences through reviewer discussion, and TPC chairs intervention (adjusting for relative quality, additional reviews).

4. Reviews, ranks and **reviewer names** are returned to the authors. This information may also be published in the GI web site, unless the authors indicate they do not want the reviews to be made public (strong request from the GI Steering Committee). Note that the reviewers cannot veto this. "

The nightmare on TPC street. Inviting TPC members was quite the experience. First, many people, some of which are our friends, replied and pretty much said that this is crazy and we are stupid, without using these exact words, though. We are exaggerating a bit: several people (roughly 5-7) expressed concerns about whether this was going to work and doubted that this is the right approach, primarily fearing the retaliation of the authors. Well-intended criticism is always good and it was helpful.

The main concern here is to not invite assistant professors, and in general fresh PhDs, who are most vulnerable to retaliations. This is clearly one of the shortcomings of this approach, and it is hard to overcome as long as people give in to pettiness when they are asked to evaluate others. As a result, we would encourage assistant professors to not give in to the temptation, even if at the time it sounds like a good idea. Having said that we did have some less-senior people who volunteered to join that TPC.

To our joy and happiness, we managed to collect an absolutely awesome TPC committee (see list at the end). You can check out the list at the GI website. However, we won't claim all the glory: there were many of the TPC members that are GI supporters, and they would have joined with or without the open review. However, they were others that joined because of the open-review novelty in at least 2-3 cases.

What really gave us nightmares was the leisurely replies of the otherwise awesome TPC. Really: we lost sleep over this. Especially, with the novelty of the effort, we were afraid that our failure would be easily associated and attributed wrongly to the open-review approach.

So, our advice to the next chairs: invite TPC members earlier than the time you think you should. In fact, you should probably start inviting them at this very time you are reading these lines (assuming this issue is published on time).

Reviewing. Overall, the actual reviewing went much smoother than anticipated. One mixed blessing was the low number of submissions – approximately 42 (coincidence?). As a reference point, the number of submissions in previous offerings were between 60-80.

Doing the paper assignments needed some attention. We wanted every reviewer to have approximately the same number of papers, and a fair number to exclude statistical variations due to randomness. Thus, for example, we could not easily integrate external reviews in a straighforward way. For this, we assigned more reviews per paper than usual. Is that bad? Not at all, the more feedback the better.

The rest was very smooth. Most people did their reviews on time, the email discussions were passionate in a good way, many decisions were made by the TPC, the final decisions were made by the chairs, after reading the reviews and the papers to the extent needed. Nothing out of the ordinary.

For full transparency, we would like to report the following: Co-chair Rejaie was part of a submitted paper. An initial attempt to treat this like any other paper was met with some reluctance by some reviewers. To alleviate any concerns, co-chair Faloutsos decided on the following course of action: he identified an extremely competent and experienced TPC member who was charged with providing a "yes or no" answer for that paper after soliciting reviewers of his/her own choice. The reason for this approach was to remove the effect of either of co-chairs on the fate of this paper. Perfect solution? Not really. Reasonable? Absolutely. It is worth mentioning that Reja regretted having been involved in that paper and inadvertently creating this issue. However, note that this issue is neither unique nor caused by the specifics of the reviewing process. It would have been an issue and potentially handled in the same way by Michalis under any reviewing process.

Here is something worth noting. His Majesty Sir Jon Crowcroft nearly declined to be in the TPC claiming that he over-comits himself, but he did anyway partly tickled by the open review process. Interestingly, when discussing the final list of accepted papers, Jon asked to quickly look over all the to-be-accepted papers, which he did, and confirmed that the list contained papers that were at least reasonable. This confirmed beyond doubt that Jon is crazy, which is exactly why we love him.

3. FEEDBACK AND OBSERVATIONS

From the very beginning, many people insisted that the "experiment" would not be complete unless we got feedback. And so we did. And it was very nice. We explain everything below.

Before we go into the numbers, we outline outcome of our feedback:

1. Authors that did submit papers seemed to be motivated to participate in the workshop because of the review process.

2. The reviewers that did join the TPC were encouraged to participate because of the open review. However, there were a few that declined because of that reason.

3. Authors saw an improvement in the reviews compared to other same-format conferences and workshops (sample list was provided in the question).

4. Neither authors or reviewers are particularly keen on establishing contact and potentially collaborating.

5. Finally, authors and reviewers alike would like GI to continue with the open review process.

Note that one could argue that the reduced number of submitted papers may have been caused by the open review process. It may be. We highly doubt it. Is there an author in the history of science who got reviews and never wondered: I wish I knew who wrote this review? We doubt it. Note also that the authors had the right to not let their papers and reviews be posted, as wisely suggested by our Steering Committee. So why would an author be intimidated by our open review is mystery to me. The low number of submissions is a fact but it could correlate with other things, such as the proximity to other major deadlines, (two NSF deadlines and the SIGCOMM deadline in January, when the GI deadline was Feb 12). One could even speculate that the existence of the INFOCOM mini-symposiums may have taken away from submissions to GI. However, we do not have hard evidence for either explanation.

In our feedback form, we defined the following range for the feedback for both authors and reviewers: [-2, 2], with -2 strong no, -1 weak no, 0 neutral, 1 weak yes, 2 strong yes
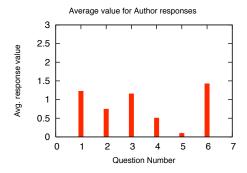

Figure 1: Summary of authors' feedback.

Feedback from the authors. We asked the following questions to the authors of both rejected and accepted papers and got 29 replies out of approximately 120-140 authors. Note however, that typically there is one person per paper that will reply.

Q.1. Did the idea of open review encouraged (+) or discouraged (-) you from submitting to GI 2007?
Q.2. Did you see an improvement in the reviews compared to last year's GI or other 6-page forums such as NOSSDAV, NETWORKING and GLOBECOM?
Q.3. Is knowing the name of the reviewers helpful (to value and trust the reviews, understanding the context of the comments)?
Q.4. Would it be useful to contact reviewers in order to clarify any issue or continue a technical discussion related to the review/paper?
Q.5. Do you feel motivated and interested in contacting the reviewers?

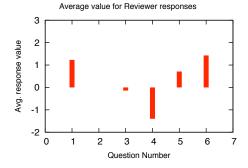Q.6. Would you like to see GI maintain the open review process next year?

Figure 2: Summary of reviewers' feedback.

Feedback from the reviewers. We asked the similar questions to the reviewers, with almost a one-to-one mapping. We got 7 replies out of the 31 reviewers.

Q.1. Did the idea of open review encouraged (+) or discouraged (-) you from participating at GI 2007?
Q.2. Did you put more effort in providing substantial reviews?
Q.3. Did this affect the tone in your reviews?
Q.4. Did you find that you hesitated to say what you wanted in the reviews?
Q.5. Would you welcome a contact from the authors to clarify, improve or collaborate on their paper?
Q.6. Would you like to see GI maintain the open review process next year?

The results for both types of feedback are shown in figures 1 and 2.

Interestingly, authors saw improved reviews while reviewers said that the open review did not affect their reviewing in a major way. This is interesting, and reminiscent of the results of a poll on how many lovers one had in a particular country (the exact details escape us): men reported an average of 7-8 lovers, while women reported 3-4. Assuming equal gender populations, things obviously do not add up. The explanation is that sometimes psychology and expected social norms can alter our reporting of facts. Something like the "Lake Wobegon" effect, where "all children are above average" (see prairiehome.publicradio.org and, of course, wikipedia). Anyway, we are digressing here.

It is interesting note that not all reviews showed extra attention. There were several reviewers that provided very short reviews. In most cases, these reviews were not seriously considered in the discussion phase, unless the reviewer could substantiate his/her concerns. However, given that reviews are publicly available, subsequent TPC chairs can easily evaluate the diligence of each reviewer before inviting them at the TPC for GI or any other conference.

The conference. These lines are being written from the airport in Anchorage, where the workshop took place. The workshop happened rather uneventfully, which is good. There were approximately 30-35 people in the room attending the workshop during the morning sessions. Obviously, the afternoon sessions had fewer attendees than that. Note that 30 people for 15 papers is two attendees per paper,

which is not bad.

Interestingly, the chairs had the pleasure of receiving positive oral feedback from authors and others people. The main feedback was that the reviewers were more "gentle" and "polite". For people to come out and say this spontaneously, it means something. In addition, if we believe that this was not an outlier, making reviews more "polite" is a big thing.

In addition, we were told that people were intrigued to submit papers because of the open process.

Take-away Lessons. Here we list our personal account of what we think we learned from this.

1. Open reviewing is doable. This is a major and indisputable lesson. It is like one of those proofs by example: there exists a workshop where open reviewing was used. The fact that there were no assassination attempts or civil wars caused by it is a bonus.

2. Reviews seem to improve with the open review according to the authors of both accepted and rejected papers.

3. Although we claim victory, time will tell the long-term implications of open reviewing.

Unsolicited advice to the new chairs of GI.

1. Choose the paper submission dates carefully. Our submission date came after a long stretch of the SIGCOMM and two NSF deadlines (after trying to avoid being sandwiched by them).

2. The review records of this years' TPC are yours to use. In fact, this applies to the chairs of any conference.

3. Sticking to a one day event may be a good idea (instead of the one and a half day). We did because we accepted only 15 papers, but we think that an extra half day would have been a stretch for many people to attend. Note that the other two infocom workshops were also for only one day.

4. Please consider adopting the open review for GI. We find that GI could become a unique forum. First, it would be good for the community to have this "different kind of reviewing". Second, it would be good to see the long-term impact of this, as the community becomes aware of this transparent process.