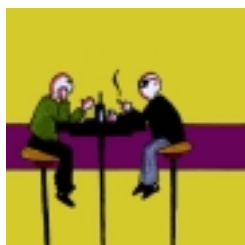


# How to *Increase* the Acceptance Ratios of Top Conferences?

Graham Cormode<sup>1</sup>, Artur Czumaj<sup>2</sup>, S. Muthukrishnan<sup>3</sup>

## Abstract



In the beginning was the pub. This work was triggered by a pub conversation where the authors observed that many resumés list acceptance ratios of conferences where their papers appear<sup>4</sup>, boasting the low acceptance ratio. The lower the ratio, better your paper looks. We decided to lampoon rather than lament such claims: wouldn't the world be better if we could encourage only high quality submissions and so run top conferences with very high acceptance ratios? This paper captures our thoughts, and it is best consumed in a pub (and in color).

## 1. Introduction

We are good citizens in the Computer Science community. We are active participants in conferences as authors, program committee (PC) members (PCMs) and organizers. We will describe certain problems with our conferences from the perspective of the PCMs and propose solutions. We really want to improve everyone's lives, and not waste everyone's time. We are also idealistic, so we do not always pay attention to the fine details of implementing our proposals or their social cost and implications.

**PROBLEM:** We have noticed that the numbers of submissions to conferences have gone up over recent years<sup>5</sup>. This puts an increasing burden on the PCs of these conferences. PC members have to read far too many papers; yet an overwhelming

---

<sup>1</sup> Center for Discrete Mathematics and Computer Science, Rutgers University, Piscataway, NJ USA [graham@dimacs.rutgers.edu](mailto:graham@dimacs.rutgers.edu)

<sup>2</sup> Computer Science Department, New Jersey Institute of Technology, Newark, NJ USA [czumaj@oak.njit.edu](mailto:czumaj@oak.njit.edu)

<sup>3</sup> Division of Computer and Information Sciences, Rutgers University, Piscataway, NJ USA [muthu@cs.rutgers.edu](mailto:muthu@cs.rutgers.edu)

<sup>4</sup> See <http://wwwhome.cs.utwente.nl/~apers/rates.html> for some detailed figures on acceptance rates of various conferences.

<sup>5</sup> We are by no means the only people to have noticed this problem. It is highlighted in an editorial in SIGACT News (Vol. 35, No. 1, March 2004, page 2). There will be a panel devoted to the subject in SIGMOD 2004: <http://www.sciences.univ-nantes.fr/irin/SIGMODPODS04/panelcamera.pdf>.

fraction of them are rejected. Sometimes one feels that all this effort is pointless. But still, sub-standard papers have to be read and referee reports written. This leads to problems. We feel jealous, since it seems that other PC members get better papers. We start to worry that this is because we are expert in a particularly poor area. Maybe our own research is less than stellar if the rest of committee thinks we are best equipped to referee the chaff. Ultimately we begin to look forward to being on PCs with fear<sup>6</sup> and loathing. A big disappointment when reviewing a poor paper is that it will not die. Once a paper is written it will be revised and resubmitted over and over (causing pain to all involved) until it finds its correct level somewhere in the conference food-chain. We need to find ways to reduce the work of the PC, and perhaps also the authors. Several of our solutions below either directly or indirectly raise the acceptance rates of conferences, hence the title of our work.



We considered existing strategies that are used to alleviate the load on the program committee, and note their principle failings:

- A. Increase the size of the PC (but this creates *more* work for *more* people. We dub this the systems solution: as theoretical computer scientists who once took a class in systems, we know that the systems solution to the problem of too much data is to throw more resources at it).
- B. Delegate reviewing to others, i.e., one's grad students (but this means we have to become *managers*, ugh).
- C. Accept (almost) all submissions (this ignores the main problem).
- D. Accept only invited submissions (this stifles creativity and new talent).

We also compared to similar situations that arise in filtering spam email messages and in the university admissions process. These approaches are explored in more detail in the full version of this paper<sup>7</sup>; in general they, like all analogies, fail and we need to look for new approaches to address the problem.

## 2. New Conference Procedures

From the preceding discussion it is clear we must eschew extant methods and embrace radically new suggestions to solving our problem. Our work is predicated on the following. The problem is too many papers are submitted, and then recycled after

---

<sup>6</sup> And this is a bad thing, since we all know that fear leads to anger, anger leads to hate, hate leads to suffering, and this leads us to the *dark side*, e.g. giving up academia to work in Seattle.

<sup>7</sup> Available as DIMACS Tech Report 2004-12, <http://dimacs.rutgers.edu/TechnicalReports/abstracts/2004/2004-12.html>

being rejected. It is not that the case that papers are all good and we have a tough time choosing which deserving papers must be dropped. If this were true, then instead we would be worrying about having to increase the number of accepted papers and the consequent organizational problems. Rather, it is the fact that we face many papers that are not appropriate for the conference, but that it takes us too much time and effort to discover this.

We propose a few approaches to reduce sub-standard submissions and reduce the load on PCs:

- Reduce the number of papers submitted
- Reduce the work for the PC by filtering techniques
- Using different submission structures
- Economic approaches
- Information-based approaches

For each of these, we outline our recommended approaches, and list some of the research problems that emerge from them.

## **A. Reducing the Number of Submitted Papers**

We can consider various ways to restructure our conferences that will guarantee a low number of submissions. For example, we could only accept the last 100 papers to arrive, or the first 100 papers to arrive. In practice, neither of these is practical: in the latter case, it will just ensure that any automated submissions server is brought down under a barrage of submissions as the process comes down to a question of who can click ‘submit’



the closest to the deadline (or write a program to automatically submit the paper as close to the deadline as possible). In the former case, then we would find the same effect at the official “start” date for submissions, and so we succeed only in moving the deadline earlier. Or, we could find some other way to fix the numbers, perhaps by randomly deleting all but 100 of the submissions. Some conference management systems have this feature built in already, but it is currently considered to be a bug, not a feature.

Sometimes individual authors are responsible for a large fraction of the work of a PC. Some authors regularly submit several—as many as five or ten—papers to a single conference. We propose limiting the number of papers that any author can submit to

a conference—perhaps to as few as one. Such restrictions apply in certain grant writing scenarios, such as the NSF ITR, where it is assumed that working on one grant will consume all of one’s available time. This approach could give rise to interesting combinatorial decision problems for multi-author papers, and would also require some authors to commit to papers. This would punish the most prolific authors... or at least slow them down a bit so that we can more reasonably compete with them. Since they can no longer deliver half a dozen papers in a single conference, the prolific would have to travel more, thus keeping them tired and jet-lagged and hopefully reducing their output and hence their overall burden on the community in terms of reading and reviewing their interminable work. To succeed, the limit may have to be set quite low. For example, in the recent SODA 2004 conference, there were only two authors with 3 or more papers out of the 135 accepted submissions. But setting the limit at one submission would have affected 55 authors who had two or more accepted papers.

**Problem 1:** *We must evaluate the effect of limiting each author to one submission on paper submission rates. Would it really reduce load on PCs? This needs detailed data analysis at various conferences across multiple years.*

**Problem 2:** *Given a set of co-authored papers and a limit of at most submission one per author, how would authors maximize the number of submissions? Can this be done without the authors discovering who else is submitting to the conference?*

## **B. Reducing PC Work by Filtering**

It is often said that getting Computer Science papers into conferences is in part a *beauty contest*. To an outsider visiting a CS conference, this remark would be quite surprising. What is meant is that the most attractive (or popular) papers on hot topics or containing many buzzwords stand the best chance of being accepted. We propose to formalize this system.

All submissions are listed and paraded. Each PCM ranks each paper according to her/his interest in it, *based only on a cursory inspection*. Perhaps each PCM will have a limited budget of total points to allocate. The papers with the highest popularity would be accepted to the conference. To reduce bias and vote-rigging, we would take the median score rather than the average.

In fact, we believe that this ranking would give a very good *prediction* of the output of a typical refereeing process without the tiresome bother of actually reading the papers. This process would also accept papers that are most interesting for the PC

members, and so lead to conferences with the most interesting papers. The downside is that this process certainly is not aimed at selecting the best papers, and some excellent papers on unpopular topics would never make conference publications. This also could lead to promoting known-names only. The next step is to automate this process: use machine-learning and data-mining methods to rank papers based on previously seen decisions. It may then be desirable to keep the rules that are found a secret, since otherwise this gives a recipe of keywords and phrases to use to increase chance of acceptance.

**Problem 3:** *We must implement and test computerized filtering based on, e.g., keywords in the abstract and author names. How does this contrast to the rankings (scores) produced by the PC? What are the secrets to getting a paper accepted based on history and textual prowess?*

### **C. Multi-resolution Papers**

For our next model, we take inspiration from procedure used for some grant proposals. Now authors must submit **two** versions of the paper: a (say) 12-page **long paper** containing both general presentation of the results and more detailed or complete analyses, and a 2-page “long introduction” **short paper**. The work of the PC is performed in two steps. First, the PC would read only the short paper and made the first selection on this basis alone. In this way, a large portion of all submissions would be rejected with a significantly smaller effort from the PC. In this phase, the PC members would have to read only motivation, comparison to prior work, and the statement of the results<sup>8</sup>. With enough caring brutality, half or more of the submissions could be rejected in this phase, significantly reducing the work of the PC. This way also allows a dynamic resource allocation approach: if in the first phase the PC decides that a given paper should be accepted unless its analysis is incorrect, then we need only 1-2 experts in the area to verify the soundness of the result without further troubling the PC.



We argue that, if the result cannot be succinctly summarized in two pages for a general CS audience, then the paper, in its current shape, is not ready for publication. In general, good papers read well, and so we expect few good papers would be pulped

---

<sup>8</sup> Some PC members admit privately that they already apply such a system informally, by only reading the introduction of most submissions.

in this way. The rejections at the first stage can be given a positive spin: this is a faster option than waiting for complete referee reports, and avoids keeping the fate of the paper up in the air for months on end. How does this affect acceptance rates? It increases them! Suppose a conference accepts 25% of submissions. Then, under this two-stage system, 50% are accepted by stage one, and after detailed study, 50% of those are accepted at stage two. Thus the conference now has  $50\% + 50\% = 100\%$  acceptance rate.

**Problem 4:** *The two-stage process is just one possibility. Design and analyze the optimal structure and acceptance rates of a reviewing process so as to minimize PC effort and maximize quality of accepted papers. Given  $n$  submissions, how many stages does the optimal solution need?*

## D. Economic

The global capitalist system tried to reduce the workload of Computer Science conferences, by starting the Internet Economy. This had the effect of redirecting the attentions of those researchers who thought that there perfect hash functions could be turned into perfect



cash functions. Unfortunately, this effect was not felt uniformly by all areas and it did not last. Before we set about beginning a new Internet bubble we should also admit that it will probably cut down the good submissions, but not the bad. So, instead we will come up with some more principled methods for using economic ideas to reduce poor submissions. The obvious approach is to design schemes of payments and penalties, to punish authors who write poor papers that waste the PC's time. This is open to criticisms that authors from poor institutions would be affected, while those from rich institutions or with big grants could more easily pay any fees levied. Instead, we look to less-direct applications of Economics.

We advocate outsourcing paper refereeing to India. It has a trained population that is well versed in recent results in Computer Science. This marks a natural progression in the market forces. This will not hinder employment in other countries. In fact, this will lead to academics in US and EU taking up higher value jobs. It will also be an opportunity to learn another language as researchers learn the nuances of "Indian English", putting commas arbitrarily, deleting articles carelessly, and long-winding

sentences that abruptly...We will have conference call centers (pun intended) in India and change the social fabric there.

**Problem 5:** *Another approach is to tap information markets: a system where value is attached to each paper, so that the best papers attain the highest value. Effectively, this means that the PC starts betting on which papers or topics are accepted. How could such a system be designed and operated effectively? What would the pay-offs be? What about futures (predicting that X will write a paper in the next 6 months on cache-oblivious algorithms) and derivatives? If we correctly predict all the accepted papers, can we make enough money to retire to Barbados?*

## E. Information-based



There is a great deal of information generated in the process of reviewing and evaluating papers. One possibility is to use this information as a weapon against authors who submit substandard papers. At the extreme is the possibility of making all reviews, of both accepted and rejected papers, fully public. One could even envision an Amazon-style system where each paper is listed along with reviews (from referees, and from casual readers) and given five-star ratings. We acknowledge that such a course of action is unlikely to prove popular, and given some of the reviews our past papers have received (from reviewers who must have somehow failed to appreciate their greatness), we think this might not be a good idea.

Indeed, even just the information about the titles and authors of rejected papers would seem to be too sensitive to make public, specially in the case of authors for whom, when given the title of one of their unpublished papers, it is possible for one to completely reconstruct their claimed results over the course of an afternoon. We propose that it is reasonable to release the (multi)set of names of all people who submitted a paper to the conference, in no particular order, in addition to the list of accepted papers. This gives a mild disincentive for people to submit papers to conferences if they are less than sure that it will be accepted.

This effect can be amplified if these “participants lists” are made publicly available for many conferences. Then we can pursue a baseball analogy, and begin to compute various statistics on performance. For example, the basic batting average, given by  $(\text{number of acceptances})/(\text{number of acceptances} + \text{number of rejections})$  gives an individual’s personal ‘acceptance ratio’. This number could be quoted on their resume, in addition to the acceptance ratio of the conferences they were accepted to.

Such statistics could find official use, being an additional criterion to consider in addition to awards, jobs, service, PCs served on and so on. Do not underestimate the power of statistics in baseball or real life.



The rejected papers are the ones that really create the most work for the PCs. A paper once rejected is often re-submitted, possible many times. This is certainly a natural situation, because of different standards of various conferences. Nowadays we even set up conference deadlines to catch as many as possible papers rejected from stronger conferences! To reduce the work of the PC we could

maintain a central confidential database for PCs eyes only. Every reference report or a report from TCS conference would be stored there and could be sent on the request of a PC chairman of any other conference to which a paper has been re-submitted. Thus, the PC could reuse this information and so reduce their work.

**Problem 6:** *How to manage and index the database of reviews? We must design a privacy-preserving scheme to manage reviews from all TCS conferences and make it available to PCs.*

There is another subtle information-driven approach, coming from comments we received from Rich Martin. We accept all papers, but not equally: some are accepted as full papers with conference presentations and appear in the proceedings; yet others are accepted as posters in the conference and appear in the proceedings as two page descriptions; *all the rest* are accepted as web presentations and appear as one page presentations. Thus all the submissions are officially accepted; submissions, once made, cannot be recanted after certain period, and once accepted, cannot be resubmitted. But the ignominy of making only the web presentation or standing in front of the poster awaiting audience will suffice to condition the authors to ensure they only make submissions that will make it to the elite class with high probability. This approach puts a lot of responsibility on the PCMs to assign the suitable bin to each of the papers, but such a push-back should be welcome by PCs.

### **3. Experimental Evaluation**

In order to evaluate the efficiency of our proposals in increasing the acceptance rates of conferences, and reducing the burden on program committees, we suggest that a detailed set of experiments be carried out. Ideally, our proposals should all be

adopted immediately for all conferences, but we acknowledge that this may cause some upset. Instead, we plan to carry out a series of experiments on past conferences to analyze the effect of our changes. Due to the increasing computerization of conference deliberations, from papers submissions, PCMs “bidding” for papers to read, discussions and decisions, then each conference potentially generates a very large amount of data that can be analyzed and interpreted. Our next step will be to obtain such logs and subject them to detailed analyses. This leads us to pose a ‘meta-problem’:

**Meta-Problem 7:** *To design and carry out experiments on large amounts of conference server logs to see how our different approaches interact. For example, to understand the effect of using simple filtering rules to accept or reject papers, study how many papers that initially receive a lukewarm review eventually get accepted.*



This manifesto for data compilation and analyses about the PC process should cause some researchers to salivate. In recent times, rather than actually building new complex systems, the community seems to encourage analyzing data and extracting minutiae from them.

## 4. Conclusions

There are many problems with the conference system beyond those focused on paper submission discussed here and we have in mind solutions for *all* of them. For example, the US National Science Foundation would like to encourage diversity in schools of higher education among students and faculty. We suggest that less published a school faculty is the more chance we accept their papers, thereby inducing graduates to prefer jobs in Dakota or Dubai rather than Massachusetts. The same broadening initiative applies to women and minorities. International coalitions should also be rewarded.

Another constituency that is currently discriminated against by conferences is that of cranks: the acceptance ratio for papers that prove Goldbach’s conjecture or  $P=NP$  is virtually zero. We would suggest setting up a special Conference for Rejected and Abandoned Papers or Journal of Unexpected and Novel Knowledge to nurture and divert this community.

There are other aspects of conference organization that needs fixing. For example, how should the members of the program committee be selected in a fair manner?



Since this is service to the community, perhaps it should be done by random selection, like jury service. Note that in the United States, the prosecution and defense (the program chair) can strike off jury members thought to be poor decision makers. If this whole process is televised, much good will come of it. This will be explored in a future production.

By now, we hope to have convinced the readers that the problem we have addressed is very important. We conclude with one final thought from our pub conversation about conference *publications*. There are no jokes that begin “A Computer Scientist goes into a pub and...”. We hope that this work will begin to remedy this situation.

## Acknowledgements

We have been greatly influenced by the works John (“Jack”) Daniels.

## References

References available on request.

## Picture Credits

Page 1 Bar Conversation <http://www.creaweb.it/grafica/illustrazione/giallo/giallo3.htm>

Page 2 <http://choson.lab.net/images/ds-bc1.jpg>

Page 3 <http://www.pages.drexel.edu/~kd35/bar.htm>

Page 5 <http://www.daleamburnfineart.com/Pages/conversation.html>

Page 6 <http://www.public.coe.edu/~pthomps/bars3.html>

Page 7 Icon from <http://www.btinternet.com/~donnypubs/>

Page 8 [http://www.portlandartmuseum.org/artwork/exhibitions/weegee/catalog/fullsize/sleeping in a bar.jpg](http://www.portlandartmuseum.org/artwork/exhibitions/weegee/catalog/fullsize/sleeping%20in%20a%20bar.jpg)

Page 9 <http://www.public.coe.edu/~pthomps/bars3.html>

Page 10 <http://www.slickrockdesigns.com/artgallery.htm>