

Viewpoint

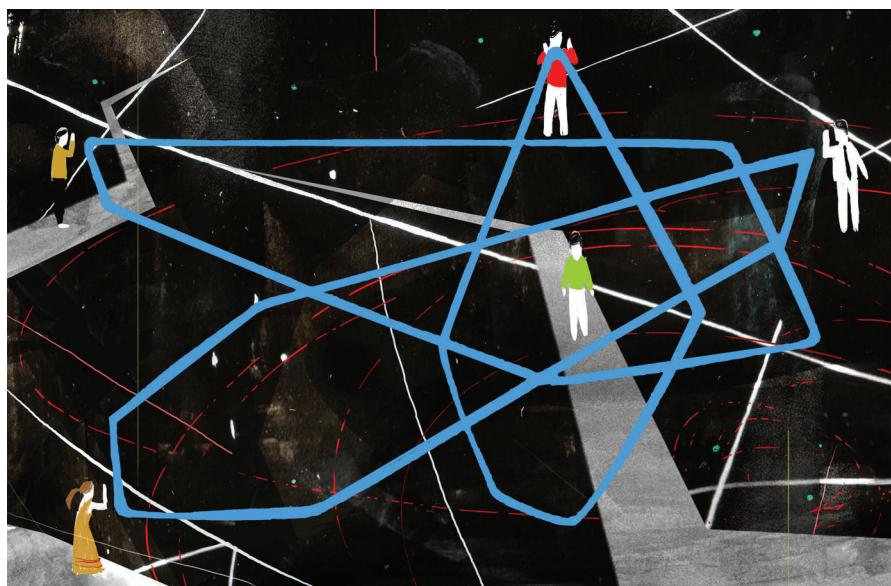
Scaling the Academic Publication Process to Internet Scale

A proposal to remedy problems in the reviewing process.

THE REVIEWING PROCESS for most computer science conferences originated in the pre-Internet era. In this process, authors submit papers that are anonymously reviewed by program committee (PC) members and their delegates. Reviews are typically single-blind: reviewers know the identity of the authors of a paper, but not vice versa. At the end of the review process, authors are informed of paper acceptance or rejection and are also given reviewer feedback and (usually) scores. Authors of accepted papers use the reviews to improve the paper for the final copy, and the authors of rejected papers use them to revise and resubmit them elsewhere, or withdraw them altogether.

Some conferences within the broader computer science community modify this process in one of three ways. With double-blind reviewing, reviewers do not know (or, at least, pretend not to know) the authors. With shepherd-ing, a PC member ensures that authors of accepted papers with minor flaws make the revisions required by the PC. And, with rollover, papers that could not be accepted in one conference are automatically resubmitted to another, related conference.

Surprisingly, the advent of the Internet has scarcely changed this process. Everything proceeds as before, except that papers and reviews are submitted online or by email, and the paper discussion and selection process is conducted, in whole or in part, online.



A naive observer, seeing the essential structure of the reviewing process preserved with such verisimilitude, may come to the conclusion that the process has achieved perfection, and that is why the Internet has had so little impact on it. Such an observer would be, sadly, rather mistaken.

Problems with the Current Review Process

We believe the paper review process suffers from at least five problems:

- ▶ A steady increase in the total number of papers: Because the number of experienced reviewers does not appear to be growing at the same rate, this has increased the average reviewer workload.
- ▶ Skimpy reviews: Some reviewers do a particularly poor job, giving numeric

scores with no further justification.

- ▶ Declining paper quality: Although the best current papers are on par with the best papers of the past, we have found a perceptible decline in the quality of the average submitted paper.

- ▶ Favoritism: There is a distinct perception that papers authored by researchers with close ties to the PC are preferentially accepted with an implicit or overt tit-for-tat relationship.

- ▶ Overly negative reviews: Some people enjoy finding errors in other people's work. But this often results in reviews that are overly negative, disheartening beginner authors.

These problems are interrelated. The increase in the number of papers leads, at least partly, both to a decline in paper quality and a decline in the

Mechanism goals.

A1	Authors should not submit poor-quality papers
A2	Authors should become reviewers
R1	Reviewers should submit well-substantiated reviews
R2	Reviewers should not favor their friends
R3	Reviewers should not denigrate competing papers

quality of reviews. It also leads to an ever-increasing variance in paper quality. Similarly, as the acceptance rate of a conference declines, there is a greater incentive for reviewers to write overly negative reviews and favor their friends.

The Paper Publishing Game

Paper reviewing and publishing can be viewed as a game. There are three players in this game, who are assumed to be rational, in the usual economic sense, and who have the following incentives:

► Authors want to get published, or, at least, get detailed, but not necessar-

ily positive, reviewer feedback on their work. They also don't want to be induced into becoming reviewers.

► Reviewers/PC members want to minimize their work (for instance, by giving scores, but no justifications), while trying to reject papers that compete with their own papers, and accepting papers from their friends. They want to reject unacceptable papers that would embarrass them. Finally, they want to get the prestige of being in the PC.

► Chairs/TCP/Research Community stakeholders want to have the highest quality slate of papers, while trying to include fresh ideas, and providing

some sense of coverage of the field.

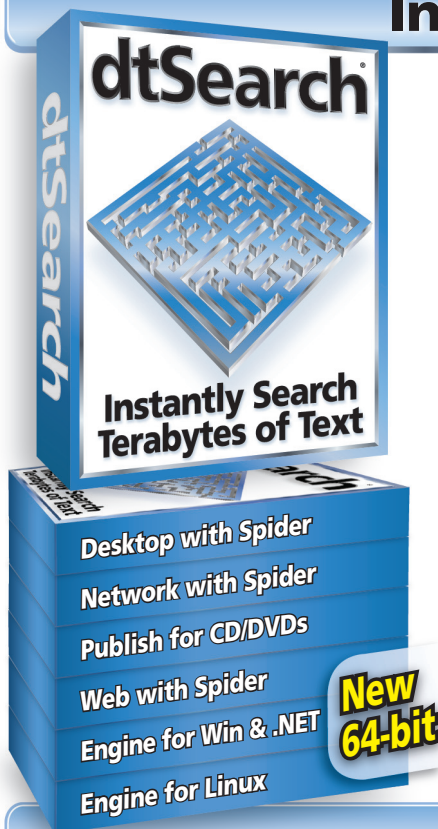
Interestingly, the problems outlined here arise because the existing paper reviewing process does not explicitly address these contradictory incentives. There is no explicit incentive for authors to become reviewers or for authors to limit the number of papers they submit, or to submit good-quality papers. There is no check on reviewers who write skimpy reviews,^a are overly negative, or play favorites. No wonder the system barely works!

Mechanisms for Incentive Alignment

Our goals, illustrated in the table here, involve designing mechanisms such that it is incentive-compatible to do the right thing. Here, we describe some mechanisms to achieve these goals (correlated to the A1, A2, R1, R2, R3 labeling scheme established in the table). Our proposals include some steps that have been tried by some brave conference PC chairs. Others that are

^a Other than a slight risk of embarrassment at the PC meeting.

Instantly Search Terabytes of Text



- ◆ Dozens of indexed, unindexed, fielded data and full-text search options (including Unicode support for hundreds of international languages)
- ◆ File parsers/converters for **hit-highlighted** display of all popular file types
- ◆ Spider supports static and dynamic web data; **highlights hits** while displaying links, formatting and images intact
- ◆ API supports .NET, C++, Java, databases, etc. New .NET Spider API

The Smart Choice for Text Retrieval® since 1991

- ◆ "Bottom line: dtSearch manages a terabyte of text in a single index and returns results in less than a second" – *InfoWorld*
- ◆ "For combing through large amounts of data," dtSearch "leads the market" – *Network Computing*
- ◆ dtSearch "covers all data sources ... powerful Web-based engines" – *eWEEK*
- ◆ dtSearch "searches at blazing speeds" – *Computer Reseller News Test Center*

See www.dtsearch.com for hundreds more reviews, and hundreds of developer case studies

Contact dtSearch for fully-functional evaluations

1-800-IT-FINDS • www.dtsearch.com

novel and would need experimentation and experience.

Author Incentives. Our first mechanism addresses A1 using peer pressure. It requires the conference to publish not only the list of accepted papers, but also, for each author, the author's acceptance rate for that conference. For example, if an author were to submit two papers and none were accepted, the conference would report an acceptance rate of 0, and if one was accepted, the author would have an acceptance rate of 0.5. Because no author would like to be perceived to have a low acceptance ratio, we think this peer pressure will enforce A1.

Our second mechanism addresses A2 by raising the prestige of reviewing. For example, conferences can have a best reviewer award for the reviewer with the best review score^b or give them a discount in the registration fee.

A more radical step would be to solve A1 and A2 simultaneously by means of a virtual economy, where tokens are paid for reviews, and spent to allow submission of papers.^c Specifically, assuming each paper requires three reviews on average, reviewers are granted one token per review, independent of the conference, and the authors of a paper together pay three tokens to submit each paper. We recognize that this assumes all conferences expect the same level of reviewing: one could pervert this scheme by appropriate choice of reviewing venues. We ignore this fact for now, in the interests of simplicity. Continuing with our scheme, authors of accepted papers would be refunded one, two, or all their tokens depending on their review score. Authors of the top papers would therefore incur no cost, whereas authors of rejected papers would have spent all three of their tokens. Clearly, this scheme forces authors to become reviewers, and to be careful in using the tokens thus earned, solving A1 and A2.

We note that we obviously need to make tokens non-forgable, non-replicable, and perhaps transferable. E-cash systems for achieving these goals are

^b See the subsection *Reviewer Incentives* for details on review scoring.

^c We have been informed that this scheme was first suggested by Jim Gray, though we cannot find a citation to this work.

The goal would be to have a standard way for members of the community to review and rank papers and authors both before and after publication.

well known^d—they merely need to be adapted to a non-traditional purpose. We recognize that regulating the economy is not trivial. Over-damping the system would lead to conferences with too few papers, or too few reviewers. Underestimating the value of tokens would only slightly mitigate the current problems, but would add a lot of expensive overhead in the form of these mechanisms. Moreover, it is not clear how this system can be implemented. Indeed, even if it was, it would not be obvious how it can be bootstrapped, or whether it would have unintended consequences. One possible technique would be to start by publishing signed reviews and rely on technologies such as Citeseer and Google Scholar as we describe here in more detail.

Reviewer Incentives. We first discuss dealing with R1 and R3. We propose that authors should rate the reviews they receive for their papers, while preserving reviewer confidentiality. Average (non-anonymized) reviewer scores would then be circulated among the PC. No PC member wants to look bad in front of his or her peers, so peer pressure should enforce R1 and R3 (PC collusion will damage the conference reputation). Note that we expect most authors to rate detailed but unfavorable reviews highly.

An even more radical alternative is for reviews to be openly published with the name of the reviewer. The idea is that reviewers who are not willing to publish a review about a paper

^d For example, David Chaum's seminal work "Blind signatures for untraceable payments," *Advances in Cryptology Crypto '82*, Springer-Verlag (1983), 199–203.

Calendar of Events

January 15–16

The 3rd International Conference on Ubiquitous Information Management and Communication
Suwon,
Sponsored: SIGKDD,
Contact: Won Kim,
Phone: 512-329-6673,
Email: wonkimtx@gmail.com

January 18–24

The 36th Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages
Savannah, GA, USA
Sponsored: SIGPLAN,
Contact: Zhong Shao,
Phone: 203-432-6828,
Email: shao-zhong@cs.yale.edu

January 19–21

International Conference on Agents and Artificial Intelligence
Porto, Portugal
Contact: Joaquim B. Filipe,
Phone: 351-91-983-3996,
Email: jfilipe@insticc.org

January 19–22

Asia and South Pacific Design Automation Conference
Yokohama, Japan
Sponsored: SIGDA,
Contact: Yutaka Tamiya,
Phone: +81-44-754-2663,
Email: tamiya.yutaka@jp.fujitsu.com

January 20–23

The Eleventh Australasian Computing Education Conference
Wellington, New Zealand
Contact: Margaret Hamilton,
Phone: 613-992-52939,
Email: mh@cs.rmit.edu.au

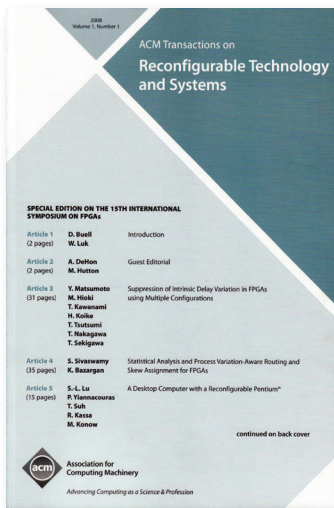
January 23–24

International Conference on Advances in Computing, Communication and Control
Mumbai, India
Contact: Srija Unnikrishnan,
Phone: +919869005457,
Email: srija.unni@gmail.com

February 4–6

6th International Conference on Computer Graphics, Virtual Reality, Visualization and Interaction in Africa
Pretoria, South Africa
Contact: Alexandre Hardy,
Phone: 27-83-267-9841,
Email: alexandre.hardy@gmail.com

ACM Transactions on Reconfigurable Technology and Systems



This quarterly publication is a peer-reviewed and archival journal that covers reconfigurable technology, systems, and applications on reconfigurable computers. Topics include all levels of reconfigurable system abstractions and all aspects of reconfigurable technology including platforms, programming environments and application successes.

www.acm.org/trets
www.acm.org/subscribe



Association for
Computing Machinery

We believe the academic community as a whole desires such a system. However, we also realize such a system can also be subverted.

are perhaps inherently conflicted and therefore should not be reviewing that paper. Of course, there is a danger that public reviews will be too polite, but this will no doubt sort itself out over time. The advantage of using true identities (verinym) is that this handles R1, R2, and R3. Alternatively, reviews could be signed with pseudonyms, where the pseudonyms could persist across conferences. Once pseudonyms will protect the nervous but prevent building reputation. There is a fundamental balance between anonymity and credibility that we cannot hope to solve.

A Grand Unified Mechanism

A deeper examination of the incentive structure suggests that perhaps the real problem is that too much of the work of submitting and selecting papers is hidden. What if the entire process were made open, transparent, and centralized? The goal would be to have a standard way for members of the community to review and rank papers and authors both before and after publication, in a sense adding eBay-style reputations to Google Scholar or arXiv. All papers and reviews would be public and signed, with either pseudonyms or verinym. This system, would, in one fell swoop achieve many simultaneous goals:

- Readers can draw their own conclusions (and tell the world) about the quality of papers published by an author. This would encourage authors not to submit bad papers (achieving A1).
- Community members who publish often and review rarely would be exposed, achieving A2.
- We would see the reviews and the

names of the reviewers alongside the paper, addressing R1, R2, and R3.

- We get to see whose opinions correlate well with our own to help decide what papers to read.
- There is a good chance that very good papers that end up as technical reports or in smaller, less well known conferences, are raised to the top by popular acclaim.
- The system would allow continued discussion and feedback about papers even after they have been published (1) to help others (busy senior people, and new people not knowing where to start), and (2) to provide an opportunity for others to participate in the discussion and debate.

We believe the academic community as a whole desires such a system. However, we also realize such a system can also be subverted. As with e-cash, the hardening of reputation systems to resist collusion and other attacks is well known, and we merely need to import the appropriate machinery and techniques.

Conclusion

We have identified the underlying incentive structure in the paper publishing process and shown where these incentives lead to poor outcomes. These insights allow us to propose several mechanisms that give incentives to authors, reviewers, and the community to do the “right thing.” We accept that there has been much altruism in the past, but in today’s resource-scarce world, it may not be fair to rely on this any longer. We recognize our work is preliminary and leaves out many important details but nevertheless hope these ideas will serve as the foundation of a fundamental rethinking of the process. We hope at least some of our proposals will make their way into future conferences, workshops, and publications. □

Jon Crowcroft (jon.crowcroft@cl.cam.ac.uk) is the Marconi Professor of Communications Systems at the University of Cambridge, Cambridge, U.K.

S. Keshav (keshav@uwaterloo.ca) is a professor and Canada Research Chair in Tetherless Computing in the Cheriton School of Computer Science University of Waterloo, Ontario, Canada.

Nick McKeown (nickm@stanford.edu) is an associate professor of electrical engineering and computer science at Stanford University, Stanford, CA.

An earlier version of this material was published in *Proceedings of the Workshop on Organizing Workshops, Conferences, and Symposia for Computer Systems (WOWCS 2008)*.