

The effect of conference proceedings on the scholarly communication in Computer Science and Engineering

Lior Shamir

Laboratory of Genetics, NIA/NIH

251 Bayview Boulevard, Baltimore MD 21224

Email: shamirl@mail.nih.gov

Conference papers have traditionally been a quick form of research communication, and an important source of information for scientists in addition to the standard journal papers. However, in the disciplines of Computer Science and Engineering, a vast majority of the peer-reviewed publications is communicated in the form of conference papers, and conference proceedings have become the primary channel of research communication in these disciplines. While this form of scholarly communication was effective for Computer Science as a young discipline, it introduces several limitations that make it non-optimal for a mature and established scientific field. These include the quality of the peer-reviewed work, selection of papers for publication, and also the efficacy of conferences as forums for expressing innovative and visionary ideas and providing opportunities for networking and meeting other researchers in the field. Here we review the differences between Computer Science and Engineering conference publications and the traditional journal publication used in other scientific disciplines, and discuss the effect of these differences on the scholarly communication in this field.

While in most scientific disciplines research results are normally reported in the form of peer-reviewed papers published in journals, in some disciplines the common method of reporting scientific results is through peer-reviewed papers presented at conferences, and published in conference proceedings. According to CiteSeerX (<http://citeseerx.ist.psu.edu>), the primary indexing engine in Computer and Information Science literature, 88 of the top 100 sources of peer-reviewed papers with the highest Garfield impact factor (Garfield, 1972) are conferences, while the remaining 12 are journals. While this list is generated automatically and may contain errors, it can reflect the importance of papers published in conference proceedings in these disciplines. In fact, unlike most other scientific disciplines, where oral or poster presentations typically require the submission of an abstract, presenters of oral or poster presentations at Computer Science and Engineering conferences are required to submit a full paper that is peer-reviewed by the conference program committee. These conference papers are in many cases considered prestigious, and are also highly important for promotion and tenure.

The conference proceedings are published as books or CDs handed to the conference attendees, but the papers are made available shortly after the conference to the members of the associations through fully indexed research databases such as *IEEE Xplore* and *ACM Portal*. Independent indexing engines such as *CiteSeerX*, *Google Scholar*, and *Microsoft Academic Search* also make these papers available to those who did not attend the conference. As a result, the accessibility of Computer Science conference papers to the scientific community is comparable to the accessibility of papers published in archived journals, providing them with an equal opportunity to be read and cited by other scientists. It should be noted, however, that according to CiteSeerX, among the top 100 most cited Computer Science and Engineering individual articles merely 22 are papers published in conference proceedings. That is, although conference papers are generally cited more often than papers published in journals, the top cited papers were not published in conference proceedings.

Historically, Computer Science and Engineering started to form as an academic discipline during the 50s and 60s (Denning, 2000). Unlike disciplines such as Biology, Chemistry or Physics, Computer Science was formed as a scientific discipline when commercial airlines were available, and provided scientists with an opportunity to travel and meet in different countries. As a young discipline it had the opportunity to shape itself in a fashion that was not dependent on traditional scientific disciplines, and conferences provided an effective method of scholarly communication for the relatively small number of scientists engaged in that discipline at the time. The explosive growth of the field in the following decades introduced the need for quick methods of peer-reviewed communication. Although Lawrence Roberts proposed and developed the idea of using ARPANET for research communication, on-line methods for publishing peer-reviewed work were not available at the time, and conference papers provided an effective form of research communication that was suitable for a fast-growing discipline.

While both journal papers and conference papers are formally considered peer-reviewed, there are substantial differences between them, which affect the peer-reviewed publications as well as the nature of conferences and meetings. One of the primary differences between the review process of papers submitted to conferences compared to standard journal papers is the time frame by which the reviewers and editors have to make their final decision about acceptance or rejection of a manuscript. While in journals the reviewers usually have several weeks to prepare their review, and in most cases (excluding the exception of special issues) the deadline for submitting the review can be extended, referees of conference papers have to review a pool of papers in a relatively short time of typically between three to eight weeks. Since reports received after the conference committee has met are useless, referees have to review numerous manuscripts within a relatively short time, and it can be reasonably assumed that this time constraint affects the amount of attention that the referee can dedicate to each paper she evaluates, and thus degrades the quality of the review (Cormode, 2008). Additionally, the conference chairperson is not able to give the personal attention to each paper in a fashion similar to the way journal papers are handled by the associate editor (Parberry, 1994), and in many cases the papers are selected based on numerical scores assigned by the referees (Smith, 1990). This limitation leads to rough criteria for evaluating the scientific work, and many authors often get the impression that the decision regarding acceptance or rejection of a manuscript is random (Mogul & Anderson, 2008). An experiment with a shadow program committee reviewing 209 conference papers showed that the shadow committee accepted just 11 papers of the 27 accepted by the actual program committee (Feldmann, 2005).

The relatively short timeframe of the review process of conference papers is also an obstacle to the communication between the authors and the reviewers. When a paper is submitted to a journal, the reviewers have the opportunity to interact with the authors and provide them with their comments and suggested revisions, and give the authors the opportunity to reply. Although this method of communication is slow and limited, it often leads to major improvements in the final accepted version of the manuscript. In conferences, however, the reviewer can score the papers or decide whether the paper should be accepted, but does not have the opportunity to review a revised version. This problem can lead to the publication of papers of marginal quality that the reviewers did not want to completely reject, and also to the rejection of strong papers that suffer from several weaknesses, which could have been easily explained by the authors in an interactive review process (Mogul & Anderson, 2008). Although some conferences allow a rebuttal, this limited method of interaction between the authors and the reviewers rarely leads to major improvements of the manuscript (Parberry, 1994).

Another major downside of the review process of conference papers is the quality of the referees. When a paper is submitted to a journal, the associate editor has the freedom to choose the reviewers, and can therefore select highly knowledgeable scientists in the specific topic of the article at hand, who the

editor believes are the most suitable people to review that type of work. In conferences, however, the large amounts of papers that need to be handled in a relatively short time often limits the chairpersons to a fixed pool of reviewers, and therefore the chairpersons might need to compromise on referees who are knowledgeable in the field, but perhaps are not considered the leading experts in the specific topic discussed in the paper. This compromise can negatively affect the quality of the review and the standards for rejection or acceptance of a manuscript.

The quality of the review becomes an important issue also in cases of less competitive conferences, where the pool of high-quality papers is limited. Since conference organizers need to attract a sufficient number of attendees, and since their registration fees (of typically \$400 to \$800) should cover the costs involved in the organization of the conference, in some cases the quality of the papers is compromised. A group of Computer Science graduate students at MIT, who suspected that some conference committees were willing to accept sub-standard papers just to collect the registration fees, submitted a meaningless randomly generated paper to several conferences. The paper was accepted for oral presentation and publication in the proceedings of the WMSCI (World Multiconference on Systemics, Cybernetics, and Informatics) 2005 conference (Ball, 2005). While this example cannot be considered representative, it can serve as a warning that in some cases sub-standard manuscripts can be published in fully citable conference proceedings (Ball, 2005).

The fact that most Computer Science and Engineering conferences require papers also inflates the total number of peer-reviewed papers published in each field. For instance, in the Computer Vision sub-discipline there are over 40 conferences that publish proceedings (<http://iris.usc.edu/Information/Iris-Conferences.html>). Each conference typically publishes between 150 to 400 peer-reviewed papers, which leads to thousands of peer-reviewed papers published every year. Even when considering only the three most prestigious Computer Vision conferences in 2009 (CVPR, ICCV, ICIP), these conferences alone produced a total of 1857 peer-reviewed papers. It can be argued that this high number of prestigious peer-reviewed papers may be inflated, and is not a reliable reflection of the true progress made in this sub-discipline in just one year.

While the review process of conference papers has a substantial effect on the type and quality of papers accepted for publication, it can also affect the conference itself as a method of scientific communication. The requirement of submitting a full paper that goes through a peer-review process might introduce an obstacle to junior scientists (e.g., graduate students), who wish to attend the conference and present their work, but might find it difficult to compete with papers submitted by established and experienced researchers. Obviously, students can attend the conference as listeners, but funding for graduate students for attending conferences is often limited, and in most cases they will not be able to attend unless they present. This is different from many of the largest meetings in Biomedicine (e.g., ASCB) or Astrophysics (e.g., AAS, IAU), in which graduate students have the opportunity to present their work, typically in the form of a poster presentation, and meet established scientists in their field.

The fact that an accepted full peer-reviewed paper practically serves as a “ticket” to the conference also introduces a downside to the authors of accepted papers who attend the conference. The reason is that when scientists spend their time and research funding for attending a conference, they often expect to meet as many researchers as possible, from which they can learn about their work and perhaps also collaborate. In that sense, a smaller number of attendees can make the conference less productive to those who attend, and the program committee practically not just filters the papers that will be published, but also the scientists that the attendees can meet at the conference.

Another downside of peer-reviewed conference papers is that in some cases the most innovative and inspiring ideas cannot be properly expressed in the form of a peer-reviewed paper due to their speculative and visionary nature. Therefore, the requirement for a peer-reviewed paper might keep some of the most creative ideas out of the conference. This can introduce a certain drawback, as one of the purposes of a fruitful conference is to stimulate discussions about the future and vision of the field.

The same problem also applies to reporting on work in progress. While one of the purposes of scientific conferences is to discuss the work before it is submitted to a journal, Computer Science and Engineering conferences attempt to become more prestigious and competitive by selecting just a small portion of the submitted papers, and conferences are often ranked by the acceptance rate of the submitted manuscripts. As a result, almost all papers accepted to Computer Science and Engineering conferences report on complete studies, and therefore scientists are not given the opportunity to discuss their work and share it with other researchers before it is submitted for publication. This leads to another issue, which is publishing scientific reports in journals after part of the study has already been published in conference proceedings. Publishing agreements of conference papers in many cases allow publication of the same work also in the form of journal papers, but require sufficient revisions and new content that justify republication. For instance, associations such as ACM and IEEE allow submitting conference papers to journals if the paper meets higher standards and justifies journal publication.

While publishing peer-reviewed papers in conference proceedings introduces several downsides as discussed above, it also has several important advantages. For instance, the review process of a conference paper is rapid, and the date of publication is known to the authors before the paper is submitted. This advantage is particularly important in Computer Science and Engineering, where papers submitted to journals are rarely rejected by the editor, and the initial reviews are normally returned to the authors within a period of between three to six months. The time that passes from the day of submission to the day of publication can be as long as two years, which introduces a significant delay in sharing the results with the scientific community. At conferences, on the other hand, the proceedings must be published before the conference starts, and therefore the publication date is well known to the authors.

Additionally, the peer-review process can set a standard for the quality of the work presented at conferences. It also allows the program committee to select speakers for oral presentations in a more knowledgeable fashion, based on the research and results described in their papers. This can be compared to the selection of oral presentations in other disciplines, in which the program committee makes its decisions based on abstracts. Practically, since the abstracts do not provide detailed information about the research, and since the review of the abstracts is not double-blind, in these conferences the primary criterion for selecting speakers for oral presentations is the reputation of the speaker, and not necessarily the scientific merit of the presented work. The conference papers, whether peer-reviewed or not, also allow the attendees to learn more about the work presented at the conference by looking up the specific paper of interest after attending an oral or poster presentation.

Another advantage of papers published in conference proceedings is that they provide an effective method for communicating scholarly work that might not be comprehensive enough to justify full journal publication, or might be of interest to a limited audience of scientists. Although bulletins and letters are designed for that purpose, conference proceedings provide an additional channel for this type of communication.

Perhaps one of the most important advantages of the prestige of Computer Science peer-reviewed

conference papers is that it allows the attendees to present their most recent work and their best results. This can be compared, for instance, to conferences in Life Sciences, in which many scientists do not present their most recent results because they feel that other researchers might replicate their work and publish it as a journal paper, which will practically give them full credit for the discovery. This discourages scientists from presenting their best and most recent work, and leads to a situation in which a substantial part of the research presented at Life Science conferences is either work that has already been published or has been accepted for publication elsewhere.

The approach of publishing scholarly work at Computer Science and Engineering conferences is different from most other disciplines. While the purpose of scientific conferences and meetings is to serve as a forum for scientists to discuss their research and ideas and to share their work with each other, Computer Science and Engineering conferences also serve as a method of publishing peer-reviewed papers. This method of research communication was effective for a young, small, and fast-growing discipline. However, since the number of active computer scientists has increased dramatically, and the current progression is arguably slower than the progression during the 80s or 90s, the effects of the many downsides of peer-reviewed conference papers start to become noticeable, and the discipline may mature and adopt the research communication methods that has been serving other disciplines for hundreds of years.

When the discipline was small, scientists could meet at conferences while also publishing their papers. However, the increasing demand for conferences that can serve as a channel for publishing scholarly work has been leading to a consequent supply of conferences in each sub-discipline, which results in thousands of papers published every year while often compromising on the quality of the contribution of the published work. Since not all papers are accepted, this form of publication also divides the researchers within each sub-discipline to different conferences, and practically does not allow all scientists active in a certain sub-discipline to meet.

When the progress of the field was fast, new ideas were not rare, and the need to express these ideas justified the publication of peer-reviewed papers at conferences. However, since the progress of the discipline has been slowing down, it is now also required to discuss work in progress and exchange visionary ideas before they become suitable for peer-reviewed publication. Although the existing Computer Science culture strongly supports the popular practice of conferences, it is possible that the academic needs described in this paper will put an increasing pressure on the discipline, and that in time, Computer Science and Engineering conferences will serve a purpose similar to conferences in other disciplines, while journal publications will become increasingly more dominant. Signs of this possible trend can be evident by emerging publication models in Computer Science, such as the one adopted by the Very Large Database (VLDB) endowment. This new publication model allows researchers in the field of databases to submit their work to the PVLB journal, and after a full peer-review process the authors of accepted papers are offered a slot in the VLDB conference.

Consequently to a possible gradual shift from conference to journal publication, conference committees might become less concerned about accepting just a small number of the submitted papers, and will work towards increasing the number of attendees and thus become more accessible to students and junior researchers who wish to discuss their ideas and meet other researchers in the field.

Acknowledgments

This research was supported entirely by the Intramural Research Program of the NIH, National Institute on Aging.

References

- Cormode, G. (2008). How NOT to review a paper: the tools and techniques of the adversarial reviewer, *ACM SIGMOD Record*, 37(4).
- Denning, P. (2000). *Computer Science: The Discipline*, in *Encyclopedia of Computer Science*, A. Ralston and D. Hemmendinger, Eds. George Mason University, Fairfax, VA.
- Garfield, E. (1972). Citation analysis as a tool in journal editing, *Science* 178, 471-479.
- Parberry, I. (1994). A guide for new referees in theoretical computer science, *Information and Computation*, 112, 96-116.
- Smith, A.J. (1990). The task of the referee, *IEEE Computer*, 23, 65-71.
- Ball, P. (2005). Computer conference welcomes gobbledegook paper, *Nature*, 434, 946.
- Mogul, J.C. & Anderson, T. (2008). Open issues in organizing computer systems conferences, *ACM SIGCOMM Computer Communication Review*, 38(3), 93-98.
- Feldmann, A. (2005). Experiences from the Sigcomm 2005 European Shadow PC Experiment. *SIGCOMM Computer Communication Review*, 35(3): 97–102.